Structural conformational analysis of molecular dynamics of a P-glycoprotein homology model and generated intermediate structures

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Melanie Salome Hafner

aus Marburg an der Lahn

Bonn, 2018

Angefertigt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen
Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Michael Wiese
2. Gutachter: PD Dr. Anke Schiedel

Erscheinungsjahr: 2019
Tag der Promotion: 23.07.2019

Allwissend bin ich nicht;
doch viel ist mir bewusst!

Goethe (1749-1832)

Abstract

The major reason for failure of chemotherapy in cancer is multi-drug resistance. A mechanism by which human cancers develop multi-drug resistance is overexpression of certain ATP-binding cassette (ABC) transporters. P-glycoprotein (P-gp) is one prominent member of the ABC transporter family. One of the strategies to overcome resistance due to P-gp, is the investigation to inhibit its function. However, the basis of drug recognition, as well as its transport has not been clarified. To date, several conformations of bacteria, nematode and murine P-gp could be crystallized.

Aim was to generate different human P-gp conformations the transporter might pass through its transition from an acitve to an inactive form. In this work, a homology model was build using a murine crystal structure. This homology model, which represents P-gp in a conformation with the nucleotide binding domains open (NBD-open) and a homology model of P-gp with the nucleotide binding domains dimerized (NBD-closed) were used to generate 18 transition states of P-gp. In addition to the two homology models, three generated P-gp conformations were taken to perform molecular dynamics simulations (MD), to investigate possible flexible structures within the P-gp. Further, three substrates and four inhibitors were docked into the high affinity P-gp NBD-open conformation to examine the influence of a ligand. Additionally principal

component analysis were performed.

All five P-gp conformations showed a good quality comparable to the crystal structures. The performed MDs showed, that a ligand enhances the closure of the nucleotide binding domains, and therefore induces a conformational change of P-gp. By flexibility studies, three putative hinge regions within the transporter were characterized. The hinge regions appeared in all conformations independent of the presence of substrate. Correlated motions, of connecting helices and their counterparts within P-gp were identified.

The results obtained from this study are promising for the ongoing research to understand the function of P-glycoprotein.

Contents

1 Introduction **1**

1.1 Cancer and multidrug resistance 1
1.2 ATP binding cassette (ABC) transporter 2
1.3 P-gylcoprotein (P-gp / ABCB1 / MDR1) 5
1.3.1 Molecular structure 5
1.3.2 Function 6
1.3.3 Transportcycle 8
1.3.3.1 ATP-switch model 9
1.3.3.2 Nucleotide occlusion model 11
1.3.4 Drug-binding site 13
1.3.5 Substrates and inhibitors 14

2 Homology Model **17**

2.1 Introduction 17
2.1.1 Crystal structures 17
2.1.1.1 Bacterial ABC transporter SAV1866 . . 18
2.1.1.2 Mouse P-gylcoprotein 18
2.2 Building the homology model 19
2.2.1 Template selection and alignment 19
2.2.2 Backbone generation, loop modelling and side-chain generation 22
2.3 Refinement and validation 23

2.3.1 Ramachandran Plot 23
2.3.2 Superposition of the homology model and the crystal structure 4M1M 26
2.3.3 Z-score 27
2.4 Further Analysis 30
2.4.1 Location of the transmembrane sequences . . . 30
2.4.2 Comparison with mutagenesis studies 32
2.4.3 Drug-binding site 34

3 Morphing **45**
3.1 Introduction 45
3.1.0.1 Targeted molecular dynamics 46
3.1.0.2 Linear interpolation 46
3.1.0.3 Normal mode analysis 46
3.2 Morphing by GROMACS 48
3.3 Morphing by UCSF Chimera 50
3.3.0.4 Chimera cluster 51
3.4 Analysis of the morphing structures 53
3.4.1 Ramachandran plot 55
3.4.2 Drug-binding sites 58
3.4.2.1 Drug-binding site NBD-1/4-closed structure 58
3.4.2.2 Drug-binding site NBD-1/2-closed structure 60
3.4.2.3 Drug-binding site NBD-3/4-closed structure 61
3.5 Normal mode by NMSim Server 63
3.5.1 Analysis of the generated structures 67
3.5.1.1 Rigid cluster decomposition 67
3.5.1.2 Superposition of the generated structures 71

3.5.1.3 Joints in the structures 72

4 Molecular Dynamics **75**

4.1 Introduction 75
4.1.1 Theory 75
4.1.2 Force Fields 76
4.2 Moclecuar dynamincs run 79
4.2.1 Building the simulation system 79
4.2.1.1 Inserting P-gp into a membrane 79
4.2.1.2 Periodic boundary conditions of the simulation box 81
4.2.1.3 Solvation of P-gp embedded in the membrane 82
4.2.1.4 Minimization 83
4.2.2 Equilibration of the system 85
4.2.3 Production run 85
4.3 Inserting a ligand by molecular docking 86
4.3.1 Docking preparation 86
4.3.1.1 Validation 86
4.3.1.2 Defining the drug-binding site of the P-gp homologymodel 87
4.3.2 Inserting the ligand into the protein 89
4.3.2.1 Docking with GOLD 89
4.3.2.2 Preparing the ligand for molecular dynamics 89
4.3.3 Equilibration 91
4.3.4 Production run 91
4.3.5 Analysis 92
4.3.5.1 RMSF 92
4.3.5.2 DSSP 92
4.3.5.3 Data merging 93

5 Analysis of the P-gp conformations 95
5.1 NBD movement of the different P-gp conformations . 97
5.2 Origin of structure flexibility 102
5.2.1 Unwinding of helices 102
5.2.1.1 Comparison with rigid cluster results of Normal mode approach by NMSim Server 112
5.2.2 H-bonds of helices 3 and 4 as example for indication of flexibility of P-gp 113
5.3 Rearrangements of helices during the simulation . . . 117
5.3.1 Helix 2 and Helix 3 119
5.3.2 Helix 3 and Helix 4 122
5.3.3 Helix 2 and Helix 4 125
5.3.4 Helix 4 and Helix 11 128
5.4 Gate of TM4/TM6 132
5.4.1 Occupancy calculation by VMD volmap plugin 134
5.4.2 Distance measurement of TM4 to TM6 137
5.4.3 TM4 Kink 142
5.5 Cross-linking 143
5.5.1 Distances of TM3 (L175) to TM9 (N820) 147
5.5.2 Distances of NBD residues P517 (NBD1) to I1050 (NBD2) 150
5.5.3 Distances of ECL1(A80) and ELC4 (R741) . . . 153
5.5.4 Dimensions of the drug-binding site 156

6 NBD-open conformation in complex with other ligands 161
6.1 Distances of NBDs 164
6.2 Ligand movement during the simulation 168
6.3 Drug-binding sites of ligands 173
6.3.1 Colchicine 173
6.3.2 Daunorubicin 178

6.3.3 Hoechst 33342 180
6.3.4 Elacridar 182
6.3.5 Tariquidar 185
6.3.6 Cis- and trans-flupentixol 189
6.4 Gate of TM4/TM6 196

7 Principal component anylsis 205
7.1 Residue-wise loadings 207
7.1.1 NBD-open conformation 207
7.1.1.1 PCA without NBDs 212
7.1.2 Morphing structures 214
7.1.3 Substrates 218
7.1.4 Inhibitors 221
7.2 Dynamic cross-correlation maps (DCCM) 223
7.2.1 Morphing structures 227
7.2.2 Substrates 231
7.2.3 Inhibitors 234

8 Conclusions 237
8.1 Outlook 243

9 Crystal structure of human P-gp 245

Appendices 249

A List of abbreviations 249

B PCA movement 260

C DCCM plots 265

Chapter 1

Introduction

1.1 Cancer and multidrug resistance

Millions of deaths worldwide are caused by cancer. A major part of cancer treatment is chemotherapy, of which the efficacy is one of the main challenges in achieving lasting remission and a definitive cure. However, the emerge of cancer resistance by tumors, that initially respond to chemotherapy, but later show a strong resistance to the original treatment, is challenging. Drug resistance is accountable for treatment failure and tumor relapse. A multidrug resistance (MDR) of tumor cells consists a cross-resistance towards structurally and mechanism diverse chemotherapeutic drugs and is a serious problem in cancer chemotherapy by becoming untreatable by cytostatics [1]. This multidrug resistance can occur through an increased efflux of chemotherapeutic drugs, leading to reduced intracellular drug levels. An increased expression of the ATP binding cassette (ABC) transporter family, which export a variety of chemotherapeutic drugs out of the cell, is a well-established cause of MDR [2].

Multidrug resistance contributing proteins are mainly localized in human tissue of the brain, lung, breast, kidneys, liver, ovaries, prostate, placenta and the pancreas. This transporters play an important role in the protection of normal cells [3].

The most prominent transporters, that are expressed in the majority of drug resistant tumors are ABCB1 (also known as MDR1 or P-glycoprotein, P-gp), ABCC1 (multidrug resistance-associated protein 1, MRP1) and ABCG2 (breast cancer resistance protein, BCRP).

1.2 ATP binding cassette (ABC) transporter

The ATP-binding cassette (ABC) transporters, firstly introduced 1992 by *Chris Higgins* [4], represent a ubiquitous superfamily of membrane proteins, classified into seven gene subfamilies (designated ABCA-G), that transport diverse substrates across membranes under ATP (Adenosine triphosphate) consumption [5].

In both, eukaryotes and prokaryotes, ABC transporters function as exporters, expelling toxins, drugs and lipids from cells, while importers appear to be present exclusively in procaryotes, bringing nutrients and other molecules into cells [6]. To these two ABC protein superfamily groups (exporters and importers), ABC proteins can also act as ion channels or channel regulators and additionally act as DNA repair and translation proteins [7].

To date, 49 members of the ABC transporter family have been discovered in human tissues. These are divided depending on the similarity in gene structure, sequence homology or order of the domains into seven subfamilies from ABC-A to ABC-G [8].

Most of them, are membrane transporters (subfamilies ABCA-D,

and ABCG) that export molecules outside of cells against a concentration gradient.

To achieve this export ABC transporters require a minimum of four domains: two transmembrane domains (TMDs) are embedded in the membrane bilayer, form the ligand binding sites and provide specificity, while two in the cytoplasm located nucleotide-binding domains (NBDs) bind and hydrolyze ATP to drive the translocation of the ligand [9].

This two NBDs are homologous throughout the ABC transporter family and have characteristic and highly conserved motifs including the Walker A and B motifs, which are found in all ATP-binding proteins, and others like the ABC signature motif and the D, H, P and Q loops, which are unique to this family [4, 10]. This ABC signature motif of the *cis*-NBD in the NBD dimer sandwiches together with the Walker A motif of the *trans*-NBD the γ-phosphate moiety of the ATP and contributes to the formation of a composite catalytic site. In addition the A-loop provides with its aromatic residues an essential contribution to the nucleotide affinity of the NBDs. Also the Q-loop thought to play a central role in the conformational coupling between the NBDs and TMDs by its contact with the TMDs and the γ-phosphate of ATP [10].

In contrast to NBDs, the TMDs are variable and reflect the diversity of the translocation pathway of the differing substrates. Despite the structurally heterogeneous sequence appearance, ABC exporters feature the two TMDs forming a dimer interface with a total of 12 α-helices (six per domain) [6]. For importers, however, the number of the helices forming the TMD varies from ten up to 20. By their different TMD folds, the ABC transport protein im-

porters can by divided into Type I and Type II, which transport substrates and a Type III importer, also called energy coupling factor transporter [11]. As mentioned, importers are found only in prokaryotes.

Beyond the TMDs and NBDs additional regulatory elements may be present. [12].

The TMDs in ABC exporters feature two intracellular (ICL) and two extracellular (ECL) loops, of which the ICL may form short "coupling helices" that make contact with the NBDs and are suggested to play a role in the interdomain communication between TMDs and NBDs [13].

The nucleotide-binding and transmembrane domains can be encoded on different polypeptides or as a single polypeptide [14]. To function the transporter needs to be a complete protein with two NBD and TMD domains (full transporter) or a dimer of two half transporters. The half transporters assemble to form a homo- or heterodimer. While homodimers have all structural elements twice, pseudo-symmetry is present in heterodimers and full transporters as P-Gylcoprotein [7].

1.3 P-gylcoprotein (P-gp / ABCB1 / MDR1)

In 1976, *Juliano* and *Ling* discovered and characterized P-glycoprotein (P-gp) in multidrug resistant Chinese hamster ovary cells [15]. Since then, the first member of the ABC subfamily B (ABCB1) became the most studied ABC transporter. The plasma membrane efflux pump is primarily expressed at tissue barriers such as skin, intestine, blood brain barrier and placental barrier [16]. The role of P-gp is likely to be cellular detoxification by expelling a large numbar of chemically unrelated hydrophobic substrates out of the cytosol. However, in anti cancer treatment, the overexpression of P-gp in tumor cells lowers the intracellular concentrations of many anti cancer agents to a subtherapeutic level [17].

1.3.1 Molecular structure

Due to its physiological role and clinical relevance in multidrug resistance, understanding the structure and mechanism is of great interest. Human P-gp consists 1280 amino acid residues, which are organized as two tandem repeats. Each repeat contains a N-terminal transmembrane domain (TMD), consisting six α-helices, followed by a nucleotide-binding domain (NBD). The two repeats are joined by a linker region with a length of 75 amino acid residues (figure 1.1) [18]. This linker segment is presumably very flexible and cleavage slightly increases ATP hydrolysis [19].

The apparent molecular weight ranges from 130-180 kDa, depending on the species and cell type in which it is expressed [20].

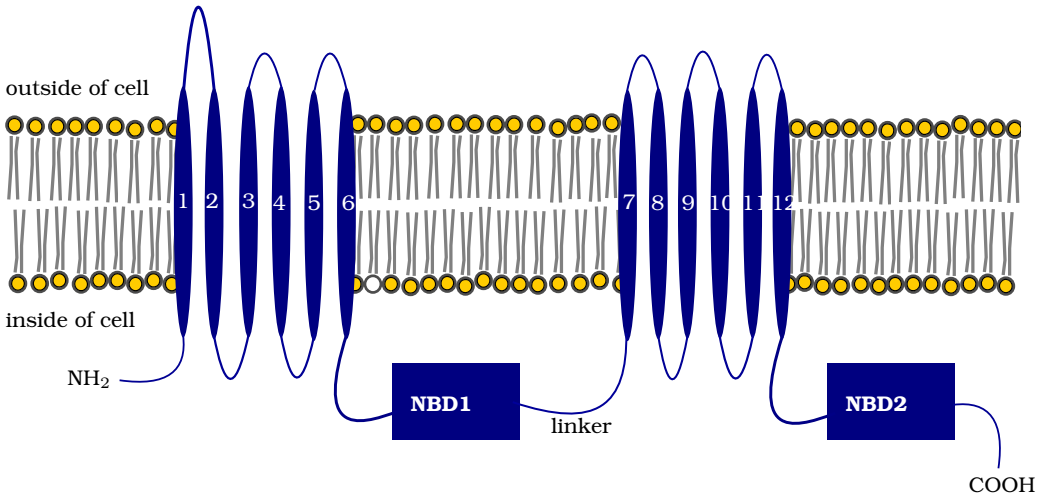


Figure 1.1: Topology of P-glycoprotein.

1.3.2 Function

Different conformations of crystal structures have shaped the perception of P-gp alternating between an inward-facing (NBD-open; NBDs far apart and drug-binding cavity facing the cytoplasm) and outward-facing (NBD-closed; NBDs close together and drug-binding cavity facing the outside of cell) conformation to bind a substrate on the one side of the membrane and release it to the other. This substrate translocation is energized ATP binding and hydrolysis.

To date, crystal structures from bacteria (e.g. Sav1866 [13]), nematode (caenorhabditis elegans [21]) and mouse (mouse ABCB1 [22, 23]) have been used for homology models to identify the structure of P-glycoprotein as well as to detect binding sites for substrates and inhibitors. However, the molecular basis of substrate recognition, the pathway substrates route through the TMDs [24]

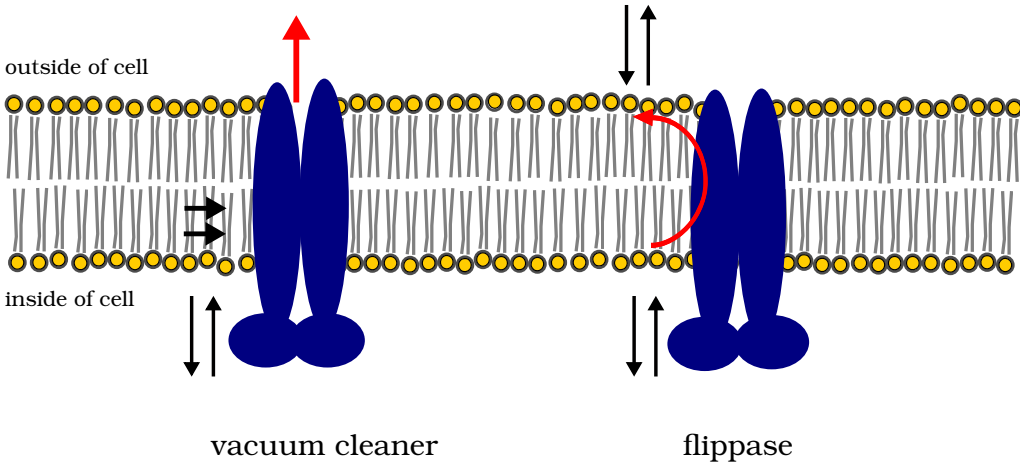


Figure 1.2: Vacuum cleaner and flippase model modified from [25] and [26].

into the extracellular space, as well as coupling between ATP binding/hydrolysis and substrate transport has still not been clarified.

Observations of P-gp substrates being typically lipophilic or amphiphathic led to the suggestion, that the transporter was a "hydrophobic vacuum cleaner" (figure 1.2). The substrates accumulate in the hydrophobic core within the lipid bilayer [25], from where they interact with P-gp and are effluxed to the extracellular space. A similar suggestion was made by *Higgins* and *Gottesman*, where P-gp moves the substrates from the inner to the outer leaflet of the membrane (flippase model, figure 1.2) [26]. After being transported into the outer membrane leaflet, substrates diffuse passively but fast into the extracellular space, rather than moving slowly back into the membrane. The rate of spontaneous returning into the inner membrane leaflet is lower that the transportation

by P-gp, and thus a substrate concentration gradient across the membrane occurs [27]. Since both suggestions reach the same equilibrium state, it is experimentally not yet possible to distinguish between these two models.

1.3.3 Transportcycle

Although there are several suggested P-gp models for the catalytic cycle, that can be divided into two main categories,

1. Complete NBD separation
2. Asymmetric nucleotide binding, associated with NBDs remaining in contact throughout the transport cycle

all models implicate a tight coupling of NBDs and TMDs. This coupling of the domains includes ATP-binding and hydrolysis by the NBDs, as well as the ADP and inorganic phosphate release and nucleotide dependent structural changes, which are linked to a conformational switch between the substrate high-affinity and low-affinity state of the P-gp [7]. ATP hydrolysis occurs by dimerization of the NBDs, in which the NBDs are positioned in a head-to-tail arrangement (ATP sandwich dimer) [28]. The transmitting signal and communication from TMDs to NBDs is still unclear.

1.3.3.1 ATP-switch model

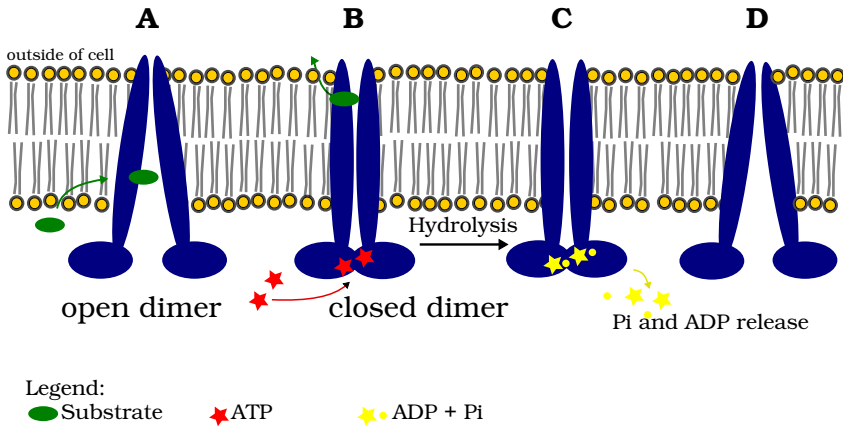


Figure 1.3: ATP-switch model. The substrate binds to the open dimer (**A**) and facilitates ATP binding (**B**). A following ATP hydrolysis to ADP and Pi destabilizes the transporter (**C**), which returns to its initial state (**D**).

The transport is driven by an on-off "switch" between a closed NBD dimer with two ATP molecules sandwiched at the NBD interface, and its dissosiation to an open NBD dimer, as a results of ATP hydrolysis and ADP/Pi release. This *switch* faciliates the conformational changes of the TMDs to transport substrates across the membrane and concentration gradient [29].

When a substrate binds to the high affinity binding site on the TMDs of P-gp, the transport cycle is initiated (figure 1.3 A). The substrate causes a conformational change, which is transmitted to the NBDs, facilitating ATP binding (figure 1.3 B). This enhanced ATP binding is suggested to lower the activation energy for forming a closed NBD dimer [30].

Translocation of the substrate across the membrane requires a reorientation of the binding site. Therefore, the high affinity bind-

ing site, accessible from the inner leaflet, is converted into a low affinity site at the extracellular face of the membrane. The key energy needed for this conformational changes is provided by ATP binding and by NBD dimerization rather than by ATP hydrolysis [28, 30]. However the precise structure interactions which trigger the changes remain unclear.

Following, ATP hydrolysis destabilizes the NBD dimer and initiates a resetting of the transporter to is basal state (figure 1.3 C and D). The exact resetting of the transporter is contentious. Two possibilities for are discussed. Firstly, substrate release from the TMDs may trigger hydrolysis by a conformational change, which is transmitted to the NBDs. Secondly, the ATP hydrolysis is an automatic reaction of the closed NBD dimer. In either way, Pi and ADP need to be released before P-gp is restored to its basal state and to start a new catalytic cycle.

1.3.3.2 Nucleotide occlusion model

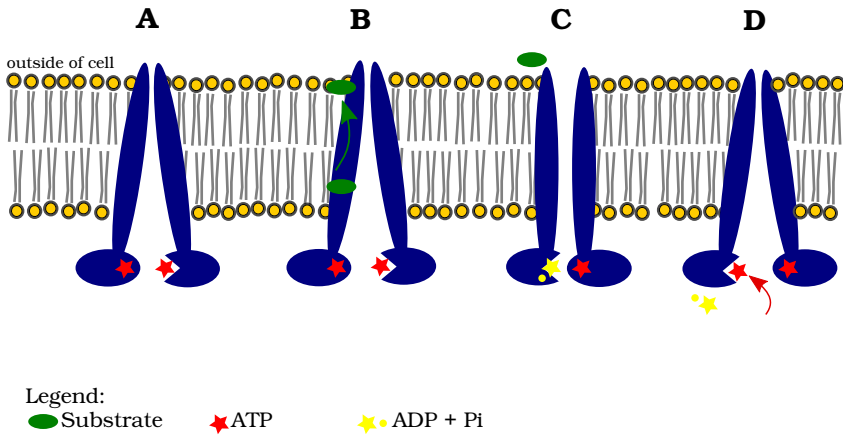


Figure 1.4: Nucleotide occlusion model. One of two ATPs gets tightly bound **A**, followed by hydrolysis and drug binding (**B** and **C**). The drug is transported and the dimer interface opens again (**C**). ADP and Pi can be replaced by a new ATP molecule (**D**).

The nucleotide occlusion model is a site switching model, in which the NBDs alternate in hydrolysis of ATP in the catalytic cycle. By this asymmetry, with a closed (high affinity) active site and a open (low affinity site) and a simultaneous switching of the affinities a memory can be preserved between two cycles (figure 1.4).

Before the catalytic cycle starts, P-gp lies in a symmetric state, which consists two ATPs bound with low affinity in both halves of the dimer interface open.

The transition state starts with one of the two ATPs getting tightly bound (occluded), thus closing the dimer interface (figure 1.4 A). The following ATP hydrolysis and drug binding may occur randomly [31] with no order (figure 1.4 B and C). When the tightly bound ATP is hydrolysed, the affinities of the two catalytic sites

of ATP are switched. Additionally, the energy for drug transport across the membrane, from the binding pocket within the transporter to the extracellular space, is provided (figure 1.4 C). This drug transport probably involves the conformational change of P-gp, from an inward-facing (NBD-open) to an outward-facing (NBD-closed) conformation.

Next, the NBD dimer interface opens again, allowing ADP and Pi to dissociate, while the catalytic site of the other NBD switches from low affinity to high affinity and binds (occludes) the second ATP molecule tightly. By binding this ATP, the dimer interface of the second NBD closes (figure 1.4 D). The dissociated ADP and Pi can be replaced by a new ATP molecule and achieving an asymmetric state again [32].

1.3.4 Drug-binding site

To date, the location of the binding site of the many P-gp substrates belonging to different chemical classes has not been identified. *Loo* and *Clarke* have performed extensive mutagenesis studies, characterizing clusters of residues, that were implicated in binding and/or transport of substrates. The investigated binding locations of verapamil, colchicine, rhondamine and vinblasitine showed a spatial proximity, which suggests P-gp containing a large non-specific binding pocket [33, 34].

In contrast to these biochemical results, pharmacological studies differentiated multiple substrate binding sites by competitive binding studies of different substrates. *Shapiro* and *Ling* defined a H- (Hoechst) and R- (Rhodamine) site by fluorescence monitoring of transport kinetics of the substrates rhodamine 123 and Hoechst 33342 with the result, that colchicine and quercetin stimulated rhodamine 123 transport and inhibited Hoechst 33342 tranpsort [35, 36].

In addition, radioligand-binding studies indicated that interactions between drugs were either competitive at a common site, or non-competitive at distinct sites. Therefore, this investigations suggest multiple drug binding sites, rather than a large non-specific pocket [37].

1.3.5 Substrates and inhibitors

P-gp transports a wide range of structurally diverse compounds out of the cell. These substrates belong to different chemical classes including cytotoxic drugs, HIV protease inhibitors, calcium channel blockers, cardiacs, central nervous system drugs, steroids, antibiotics and peptides [17, 38]. In order to overcome MDR induced by P-gp, several efforts have been undertaken in the past to develop potent inhibitors. As the substrates, inhibitors belong to a variety of chemical and pharmacological classes [39].

In general, inhibitors of P-gp can be classified in three generations [40]. Examples of *first generation* inhibitors include verapamil, nifidipine, cyclosporine A and quinidine. These inhibitors share a unique property of being already existing therapeutic agents, but exhibited high toxicity profiles at their required concentrations to attenuate P-gp. This lack of efficacy, severe toxicity and side effects led to the development of *second generation* P-gp inhibitors, which were majorly structural analogues. Inhibitors of the second generation as valsopar (PSC-833), a cyclosporine A analog, or elacridar (GF120918) exhibit lower toxicity and are more potent, but unpredictable pharmacokinetic interaction limits drug clearance and leads to toxic plasma concentrations [41].

Third generation inhibitors were designed based on structure-activity relationships favoring low pharmacokinetic interactions, and a minimal effect on other members of the ABC transporter family [42]. These inhibitors as tariquidar (XR9576), which binds with high affinity to P-gp, demonstrate MDR reversal at nanomolar range. Although tariquidar is one of the most promising third generation inhibitors, most studies have been discontinued, due to non-achievement of improved response rate or a benefit risk ratio

[43]. Because of the non-specific toxicity associated with the use of P-gp inhibitors, most of clinical trials have been terminated or were not successful. Due to a high variability of P-gp expression, the clinical studies showed a high variability of the response rate.

Besides directly inhibiting P-gp, alternative ways and techniques are being explored. An inhibition of the expression of P-gp, may offer more specific mechanisms [44]. Also MDR gene inactivation, peptides and monoclonal antibodies are an alternative to overcome MDR at molecular level [39].

Chapter 2

Homology Model

2.1 Introduction

To study the function of a protein, the structure is of great assistance, but there is not always a protein structure available. Homology modelling can be used to predict the protein structure and is based on the observations that two proteins of the same family with similar amino acid sequences have similar 3D structures. The protein with the known structure is the template on which the homology model of the target protein is based on. The higher the sequence identity of the template and the protein, the better the resulting model.

For the investigation of conformational changes of P-gp, homology models with different conformations are needed.

2.1.1 Crystal structures

There are presently two crystallized conformations of P-gp. The NBD-closed, outward-facing conformation PDB ID: 2HYD (bacte-

rial ABC transporter SAV1866 from *Staphylococcus aureus*) and the NBD-open, inward-facing conformation PDB ID: 4M1M (*mus musculus P-gp*), 4AYT (*human ABCB10 transporter*) and 3QF4 (ABC transporter TM287–TM288 from *Thermotoga maritima*).

2.1.1.1 Bacterial ABC transporter SAV1866

The SAV1866 structure represents the ATP-bound state with the NBDs in close contact and a central cavity formed by the TMDs. This cavity is shielded form the inner leaflet of the lipid bilayer and cytoplasm, but exposed to the outer leaflet and the extracelluar space. The bottom of the cavity reaches beyond the intracellular membrane boundary, but no connection to the cytoplasm exists. The polar and charged residues suggest little or no affinity for hydrophobic drugs [13].

2.1.1.2 Mouse P-gylcoprotein

This crystal structures have the NBDs separated by approx 30Å and a large internal cavity open to both the cytoplasm and the inner leaflet exist. Two portals formed by TM4/TM6 and TM10/TM12 allow access for entry of hydrophobic drugs directly from the membrane. The crystal structures of mP-gp with co-crystallized QZ compounds (PDB ID: 4M2S and 4M2T) demonstrate that the NBD-open inward-facing conformation may represent an initial stage of the transport cycle that is competent for drug binding [22, 23].

2.2 Building the homology model

2.2.1 Template selection and alignment

The first challenge in homology modelling is the selection of a suitable template. Since 1971 the protein data bank provides as single repository information about 3D structures of proteins and nucleic acids [45]. To find suitable templates for the P-gp homology model, the Basic Local Alignment Search Tool (BLAST [46]) is used. BLAST finds crystal structures in the Protein Data Bank similar to the given protein sequence. In the following table the first five search results are listed.

Sequence identity	PDB ID	Species	Resolution in [Å]	Literature
87 %	3G60, 3G5U	mus musculus	4.4, 3.8	[22]
87 %	4M1M 4M2S, 4M2T	mus musculus	3.8 4.4, 4.35	[23]
46 %	4F4C	caenorhabditis elegans	3.4	[47]
37 %	4AYT	homo spapiens	2.85	[48]
34 %	3QF4	thermotoga maritima	2.9	[49]

Table 2.1: Listings of the BLAST results.

The chosen crystal structure is the mouse P-glycoprotein (mP-gp) structure PDB ID: 4M1M with a sequence identity of approx. 86%. This high sequence identity is an ideal basis for the following alignment of template and target structure.

In the alignment each residue in one sequence is paired with either the same or similar residue of the other sequence or with a gap. By taking into account the gaps, the identical and similar residues an alignment score can be calculated. The values of

identities and amino acid replacements are pre-calculated numbers (amino acid substitution matrix, e.g. BLOSUM62 [50]). Gaps take penalties.

Usually the sequence identity is by far lower (e.g. Sav1866 structure) and a multi sequence alignment is needed.

The starting (1-33) and ending (1275-1280) residues, as well as the internal linker (residues 631-687) of P-gp were cut out, due to lack of structural information in the accessible crystal structures.

The sequence alignment was done with the *Protein Align* of MOE [51] using the default settings. The alignment was additionally checked with the topology, structure and family domains of the Universal Protein Resource (UniProt) database [52]. The UniProt database contains structural and functional information of proteins.

Chain A (representing TMD1 and NBD1) and chain B (representing TMD1 and NBD2) were aligned separately. Considering the high sequence identity of mPgp and Pgp, there are only small residue differences. Unfortunately there is a "bigger" 20 residue gap/insertion in TM1 and TM2, especially spanning the first extracelluar loop. Loops are considered as the most variable regions of a protein, for this reason additional alignments with known structures PDB ID: 2HYD, 4AYT, 3QF4 were performed to model this loop and especially the residues 83-104. After the multi-alignment with the additional crystal structures the same procedure as with the whole protein was done (see section 2.2.2).

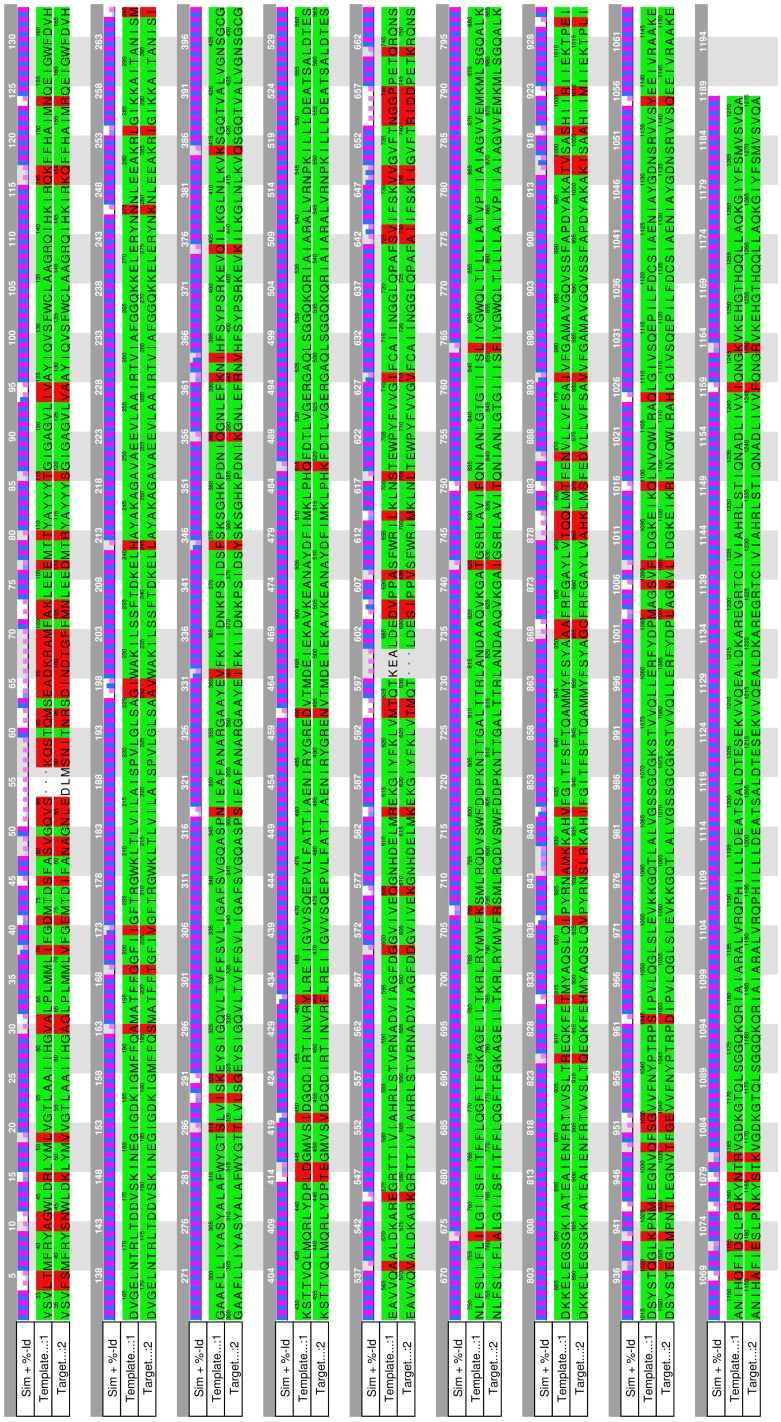


Figure 2.1: Alignment of human P-gp sequence to template 4M1M sequence. In *Green* identical residues and in *red* residues that differ from each other are shown. On top of the template chains is the residue similarity (*blue*) and residue identity (*magenta*).

2.2.2 Backbone generation, loop modelling and side-chain generation

The next step was the backbone and side-chain generation. This was done using the default MOE [51] *homology model* tool settings. Every homologous part of P-gp was generated separately, but since P-gp is a homodimer the environment function was used to include the possible interactions and clashes of the neighbour chain into the calculation. While generating the first chain (chain A), the second chain (chain B) of mP-gp was used, for generating chain B, chain A of the homology model was used as environment.

For each homologous part 100 models were generated and the best one of these models was chosen for further investigations. The MOE final model is based on score of the best generated intermediate structures by RMSD value (see section 2.3.2), electrostatic solvation energy, residue-packing or effective atomic contact energy.

The crystal structure has a "broken" helix segment between residues 990-994. Besides having the priority only on a good packing score and low number of outliers of the final model, also this helix was controlled for inconsistencies.

The ECL1 (extracellular loop 1) between TM1 and TM2 was created separately and was inserted into the final homology model.

For modelling the ECL1 the crystal structures PDB-ID: 3QF4 (ABC transporter TM287-TM288 from *Thermotoga maritima*, PDB-ID: 4AYT a *human ABCB10 transporter* and PDB-ID: 4F4C from *caenorhabditis elegans* were additionally taken for an alignment. Subsequently the the external loop was modelled in vacuum as the homologous parts of P-gp by identical settings and finally inserted into final homology model. The residue bonds between the protein

and the inserted loop were minimized. For the minimization as well as for the minimization of the generated intermediate structures the CHARMM27 forcefield was chosen. C. Globisch [53] used this forcefield for the NBD-closed homology model of P-gp, and since a forcefield defines the inner atomic potential, the ambition was to accomplish a homology model of a different conformation with equivalent atomic potential.

2.3 Refinement and validation

When a homology model is generated, first the backbone of the protein is built following by the side-chains. Building the backbone is the easiest step, in which the backbone atoms of the target are arranged identically to the backbone of the template. The side-chain generation is more challenging, here especially the differences of the sequences are taking place. Residues that are conserved can be obtained from the template, but residues that differ have to be predicted. Side-chains in a protein tend to exist in a limited number of low energy conformations (rotamers). These rotamers can be found in libraries which contain information of residues based on known crystal structures or generated from *ab initio* modelling. Considering this informations the side chains of the residues that are not conserved are added.

2.3.1 Ramachandran Plot

The plausibility of the final model was validated by means of *Ramachandran Plots*. The torsion angles *Phi* and *Psi* (shown in figure 2.2 show the flexibility of the protein backbone to adopt to a cer-

tain fold. The partial double-bound character of the peptide bond, restricts rotations around the N-C bond.

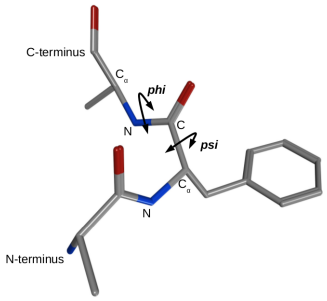


Figure 2.2: *phi* and *psi* rotations.

Therefore *Phi* and *Psi* are restricted to certain values, since other rotations will result in a steric clash between the backbone and the side-chain atoms.

The Plot has different regions representing the *most favoured* combinations of *Phi-Psi* values, followed by the *allowed* regions and finally the sterically *disallowed* regions. Residues, which are in *disallowed* regions are also called outliers Ideally over 90% of the residues are in the *most favourable* also called "*core*" regions.

Procheck [54]

	most favoured	**additionally allowed**	**generously allowed**	**disallowed allowed**
homology model	89.9%	8.2%	0.9%	0.9%
4M1M template	90.8%	7.9%	0.9%	0.3%

Table 2.2: Distribution of residues by means of ramachandran plots analysed by Procheck.

MOE [51]			
	most favoured	**allowed**	**disallowed**
homology model	91.1%	8.4%	0.5%
4M1M template	93.9%	5.7%	0.3%

Table 2.3: Distribution of residues by means of ramachandran plots analysed by MOE.

As shown in table 2.2 and 2.3 template and homology model have both approx. 90% of the residues in most favoured regions. Furthermore there are not more than 1% of the residues in disallowed regions.Therefore the homology model has comparable results as the crystal structure and is of good quality.

The outliers (residues that are in *disallowed* regions) are *ASP188*, *ASP689*, *TRP698*, *PHE971*, *GLU1039*, *VAL1040*.

ASP684 and *TRP698* terminate the internal linker, which was cut out, and are located on the membrane surface, so they can be neglected.

The crystal structures PDB ID: 4M2S and 4M2T, that correspond to the template, have one/two ligands co-crystallized. A closer look to these two crystal structures and the homology model shows, that *ASP118*, which is positioned within TM3, is far away from the drug-binding pocket of the co-crystallized Ligand QZ59. *PHE971* is on the connecting loop of TM11 and TM12 and close to the drug-binding pocket, but not directly involved in drug binding. Furthermore *ASP188* and *PHE971* were not identified to be located in putative the drug-binding sites (see section 2.4.3). Considering the outliers not to be involved in drug-binding, they can as well be neglected for the further work.

2.3.2 Superposition of the homology model and the crystal structure 4M1M

The degree of similarity of two protein structures is measured by the root-mean-square displacement (RMSD) between equivalent atom pairs *N*. The distance *d* between equivalent atoms *i* is computed after optimal superposition of the two structures and it depends on the number of the equivalent pairs.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N_{atoms}} d_i^2}{N_{atoms}}} \tag{2.1}$$

As shown in figure 2.3 the RMSD value of two proteins depends upon the extent of their sequence identity in their core regions [55]. The RMSD value is between 2.31Å (low sequence identity) and 0.26Å (high sequence identity). A protein structure will be a good template, if the sequence homology is over 50 %. With a sequence identity of 86 % a RMSD of approx. 0.5 Å is expected and satisfied by a RMSD value of 0.517 Å.

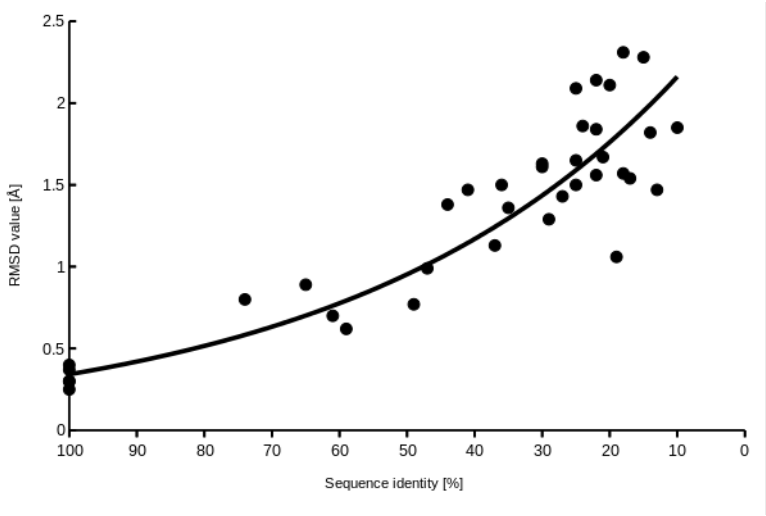


Figure 2.3: Relation of sequence identity and the backbone atom RMSD value (modified from C. Chothial and A.M. Lesk [55]).

2.3.3 Z-score

Potentials for protein folding derive from statistical analysis of known protein structures. The quality of such potentials is often assessed by so called Z-scores, which test how well the potentials differentiate the native fold of a protein from an ensemble of misfolded structures [56].

ProSA-web [56, 57] displays the z-value in a plot, which contains the z-scores of all experimentally determined proteins in current PDBs (X-ray, NMR). It can be used to check whether the z-score of the input structure is within the range of scores typically found for native proteins of similar size. The size of the native proteins corresponds not to the whole protein, but to the single protein chains. The plot shown in figure 2.4 and 2.5 show the local model quality score as a black dot by plotting energies as a function of amino acid sequence position. In general, positive values correspond to

problematic parts. If the dot is outside of the characteristic range of native proteins, the structure probably contains errors.

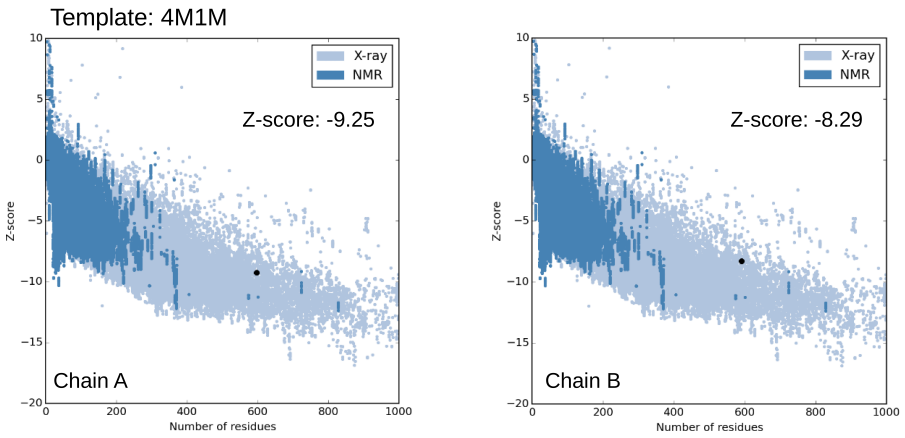


Figure 2.4: Overall model quality for the template.

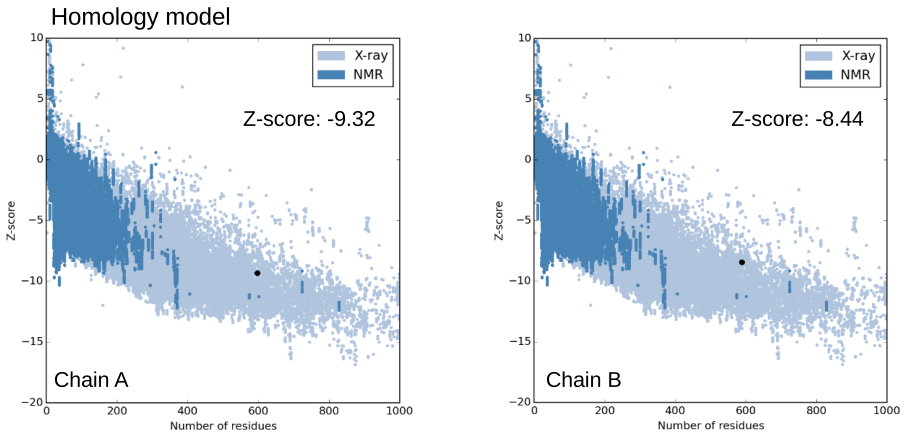


Figure 2.5: Overall model quality for the homology model.

Both chains of the homology, z-scores of -9.32 (chain A) and -8.44 (chain B), have similar results as the template chains (z-score of -9.25 and -8.29). These values indicate that the quality control values are within the range of native proteins. Therefore the homology model did pass the formal evaluation and needs to be considered as suitable for further investigation.

2.4 Further Analysis

2.4.1 Location of the transmembrane sequences

As mentioned in the fourth paragraph of section 2.2.1 the UniProt database [52] contains structural and functional information of proteins. To investigate whether the transmembrane sequences of the homology model match their predicted location, the in a membrane embedded protein was compared to the listings of the UniProt database.

	TM1	**TM2**	**TM3**	**TM4**	**TM5**	**TM6**
UniProt	45-67	117-137	187-208	216-236	293-316	331-352
Model	45-73	113-137	189-210	214-235	292-317	329-352
	TM7	**TM8**	**TM9**	**TM10**	**TM11**	**TM12**
UniProt	712-732	757-777	833-853	855-874	935-957	974-995
Model	710-734	754-777	832-853	856-875	935-960	972-996

Table 2.4: Comparison of the transmembrane sequences of the homology model and the UniProt listings.

As shown in table 2.4 the location of the transmembrane helices correspond very well to the listings. For example, while TM3 has only a two residue mismatch, TM9 is located almost identically (mismatch of only one residue).

A discrepancy with an eight residue mismatch (UniProt TM1: 45-67 to homology model TM1: 45-75) has only transmembrane helix 1. This might be caused by the separate modelling of ECL1 (extracellular loop, section 2.2.1). ECL1 of the homology model may have a larger width, which results to a lightly different angle of the TM1 and more residues spanning the membrane. With 3.6

residues per helix turn (pitch) the TM1 is a little more than one pitch inside the membrane, which should not lead to any problems during the following molecular dynamics runs (chapter 4) due to mislocations. Furthermore the residues *His61*, *Gly62*, *Leu65* and *Met69*, which are predicted to be involved in drug-binding lie all within the membrane, so there are no inconsistencies regarding location of drug-binding residues. .

2.4.2 Comparison with mutagenesis studies

Loo and *Clarke* are known for several cross-linking and mutagenesis studies. In some studies they used arginine-scanning mutagenesis for the TM segments of a P-gp processing mutant (G251V) [58, 59]. This mutant shows a partial maturation efficiency of approx 15 % and can be used to detect folding of P-gp in whole cells by monitoring the conversion of P-gp from a core-glycosylated P-gp (molecular of weight 150 kDa) to a mature P-gp (molecular weight of 170 kDa) that contains complex carbohydrates. An introduction of an arginine facing the lipid bilayer would likely inhibit maturation, whereas an arginine introduced into the aqueous face of the drug translocation pathway would not inhibit maturation of the mutant P-gp.

By plotting their arginine-scanning results onto heilcal wheels, Loo and Clarke observed, that most TM segments have distinct hydrophobic and hydrophilic faces.

In figure 2.6 residues of the transmembrane domain tested by Loo and Clarke are coloured in *magenta* (suggested to face the drug-binding cavity) or *yellow* (suggested to face the membrane bilayer).

The orientation of the residues within the model correlate with Loo and Clarkes observations [58]. The TMs, that are predicted to have distinct hydrophobic and hydrophilic faces, have their residues facing the lipid bilayer or lying within the inner drug-binding cavity, respectively.

The homology model shows TM9 not to be located directly into the inner drug-binding cavity, this displays the results, that TM9

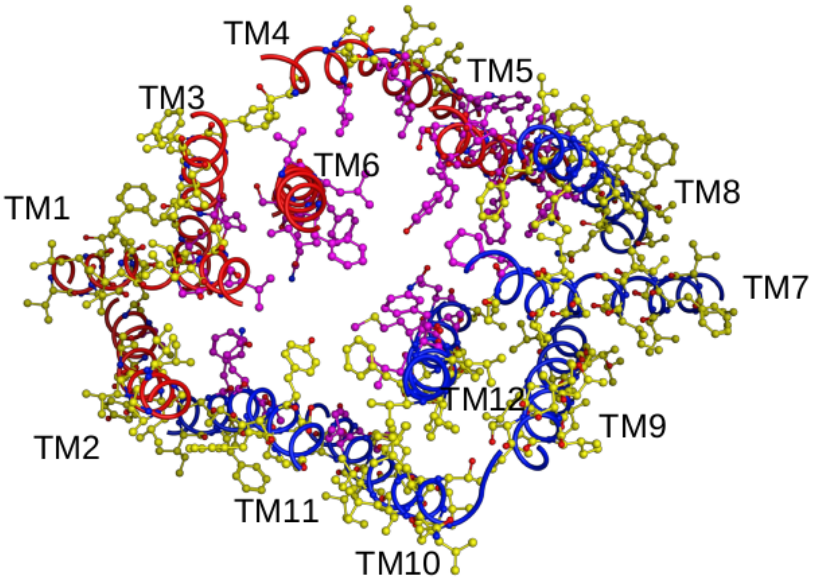


Figure 2.6: Residues of the homology model facing the drug-binding site. Residues suggested to face the bilayer are coloured *yellow*, residues suggested to face the drug-binding site are coloured in *magenta*. TMD1 is shown in *red* and TMD2 is shown in *blue*.

had the only transmembrane section without enhancer mutations.

Notability in their studies TM6 did not appear to contain a hydrophobic face and and was suggested not to be attached to the lipid bilayer, but rather to lie within the inner cavity. As shown in figure 2.6 the TM6 is not in contact with the lipid bilayer, because it is shield from the membrane by TM3 and TM4.

2.4.3 Drug-binding site

The MOE [51] *Site finder* is a geometric method unlike interaction energy methods. Regions of tight atomic packing are identified. Of these regions the ones that are "too exposed" to solvent are ignored. The remaining regions are then classified as hydrophobic or hydrophilic. If there are only hydrophilic sites in their neighbourhood, they likely are water sites and are ignored too. *Spheres* are calculated for each site by triangulating collected 3D points using a modified *Delaunay* triangulation. For each resulting simplex (collection of four points) there is an associated sphere called *alpha sphere*. This *alpha spheres* have differing radii corresponding to the planes of the convex hull of the point set. Of the resulting *alpha spheres*, those that correspond to inaccessible regions of the protein or are too exposed to the solvent are deleted. Finally the remaining *alpha spheres* are classified as hydrophilic or hydrophobic depending on whether they are or are not in a good hydrogen binding spot within the protein.

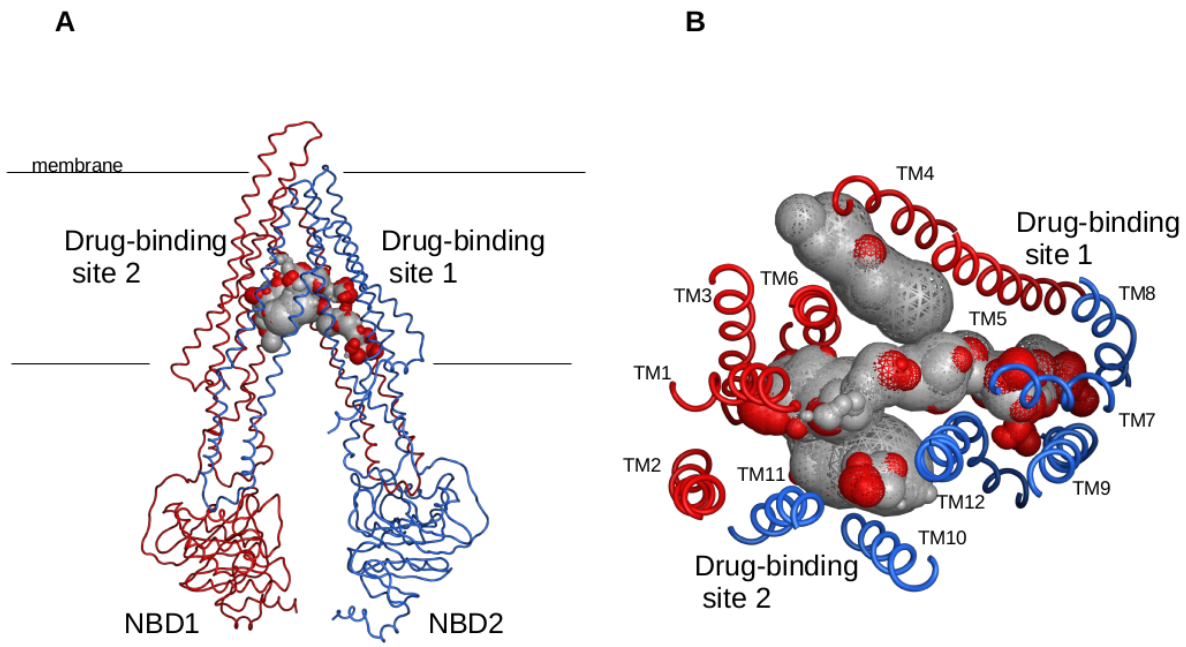


Figure 2.7: Location of the drug-binding sites 1 and 2. **A:** Side view from P-gp. **B:** Extracellularview from P-gp. Only TMs which are involved in drug-binding are shown. *Red* coloured are TMD1 and NBD1 and *blue* coloured are TMD2 and NBD2.

Site finder identified two possible drug-binding sites, which are divided in several smaller drug binding pockets. The sites are opposing each other and are located in the transmembrane regions. Both are connected to their respective suggested TM gate (TM4/TM6 and TM10/TM12) by a pocket which leads into the membrane.

The first drug-binding site is elongated and includes three drug-binding pockets. Two of the pockets have a hydrophobic core,

which is surrounded by smaller hydrohilic regions and which are located within TM1, TM5 and TM6 of TMD1 and TM7, TM8, TM9 and TM12 of TMD2. The third pocket leads toward the bilayer through TM3, TM4 and TM6, it has no hydrophilic spheres and may function as gate to the membrane, through which the substrates enter the transporter.

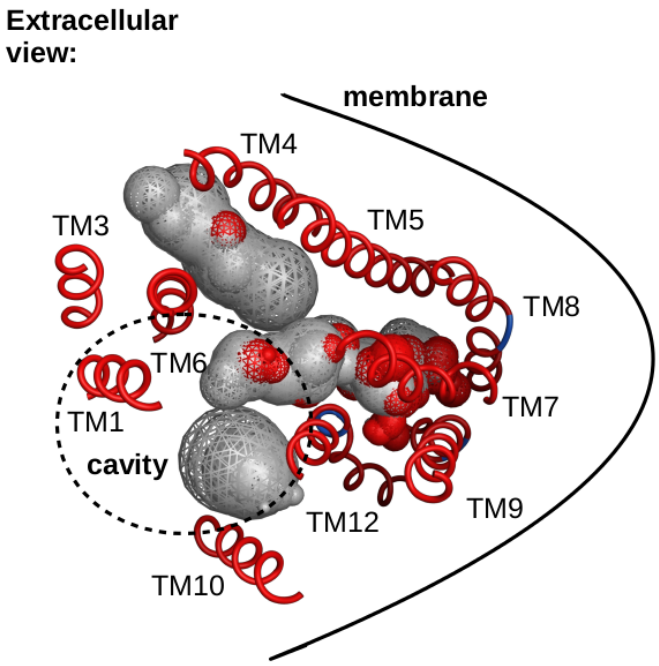


Figure 2.8: Drug-binding site 1 view from outside of cell. Only the involved transmembrane segments are shown. *Gray* spheres are hydrophobic and *red* spheres are hydrophilic. On top, the gate to the membrane, formed by TM4 and TM6, is visible. The inner cavity of P-gp is marked by a dotted circle.

I306, *Y307* (TM5), *L339* and *A342* (TM6) are predicted to be involved in vinblastine, cyclosporin A, colchicine and verapamil binding [60]. In contrast none of the identified residues was reported for rhodamine binding. Investigations on drug transport pointed to the existence of multiple drug-binding sites [61], where this site may represent the H-site (Hoechst33342 drug-binding site), since all listed residues are not identified to be involved in rhodamine binding.

This drug-binding site 1 was chosen for further investigations. Out of the listed 53 residues 15 (approx. 28 %) are reported to have an effect in drug-binding and twelve are listed to face the pocket [58, 62, 63]. Furthermore the co-crystallized ligands within this site can be used to define easily the binding area in docking approachesfor ligands as colchicine, which interacts with the H-site of P-gp.

In the following tables 2.5 and 2.6 all residues, which are identified to be located in the drug-binding site 1, are listed. Investigated residues are in bold and their suggested function with the appropriate literature is listed behind:

	Residue	Suggested function	Literature
TM1	**M69**	predicted to lie close to tariquidar bindinge-site	[64]
TM3	F200 F204	no information about drug-binding	
TM4	L219	no information about drug-binding	
	S222	predicted to lie close to verapamil binding-site predicted to lie in vinblastine and ciclosporin A binding site	[60] [65]
	L225	predicted to face drug-binding area	[58]
	G226	no information about drug-binding	
TM5	A292	no information about drug-binding	
	N296	predicted to face drug-binding area	[58]
	I299	predicted to face drug-binding area	[58]
	F303	predicted to face drug-binding area	[58]
	I306	predicted to lie close to the tariquidar binding-site may be part of signal to switch on ATP-hydrolysis when drug-binding site is occupied substitution reduced apparent affinity for vinblastine and cyclosporin A	[64] [66] [67]
	Y307	substitution leads to loss of substrate ability to inhibit labelling with IAAP predicted to lie close to the tariquidar binding site	 [68] [64]
	Y310	predicted to lie close to the tariquidar binding site	[64]
TM6	**F336**	predicted to lie close to the tariquidar binding site	[64]
	L339	predicted to lie close to verapamil, colchicine, vinblastine and cyclosporin A binding site substitution leads cyclosporin A to inhibit IAAP labelling colchicine, verapamil and vinblastine protected from dBBn inhibition	 [60, 69] [70] [71]
	I340	predicted to lie close to rhodamine binding site	[72]
	A342	predicted to lie close to verapamil binding site substitution confers resitance to colchicine, vinblastine and doxorubicin colchicine, verapamil and vinblastine protected from dBBn inhibition	[60] [73] [71]
	F343	predicted to lie close to the tariqidar and rhodamine binding site substitution leads to loss of substrate ability to inhibit labelling with IAAP	 [64, 72] [68]

Table 2.5: Table of residues (TMD1) located in drug-binding site 1.

	Residue	Suggested function	Literature
TM7	N721 G722 L724	no information about drug-binding	
	Q725	substitution leads to loss of substrate ability to inhibit labelling with IAAP	[68]
	F728	substitution leads to loss of substrate ability to inhibit labelling with IAAP predicted to lie close to the tariquidar binding site	[68] [64]
TM8	**F770**	predicted to face drug-binding area	[58]
	Q773 F777 Q778 G781 E782	no information about drug-binding	
TM9	K826 G827 I829 G830 S831 R832 A834 V835 Q838 N839 N842	no information about drug-binding	
TM12	**F983**	substitution affects inhibition of drug transport by cis-\trans-flupentixol	[74]
	G984	predicted to lie close to verapamil binding site	[60]
	M986 A987	no information about drug-binding	
	V988	caused reduction in the efficency of cyclosporin analogon	[75]
	Q990	predicted to face drug-binding area	[58]
	V991 S993 F994 A995 D997 Y998	no information about drug-binding	

Table 2.6: Table of residues (TMD2) located in drug-binding site 1.

The second drug-binding site is spherical, compact and includes six binding pockets. It is formed by TM1, TM3 and TM6 of TMD1 and TM10, TM11 and TM12 of the TMD2. The binding site is well exposed to the inner pore of the protein and has a hydrophobic core region. Hydrophilic regions are located towards TM1, TM3 and TM6.

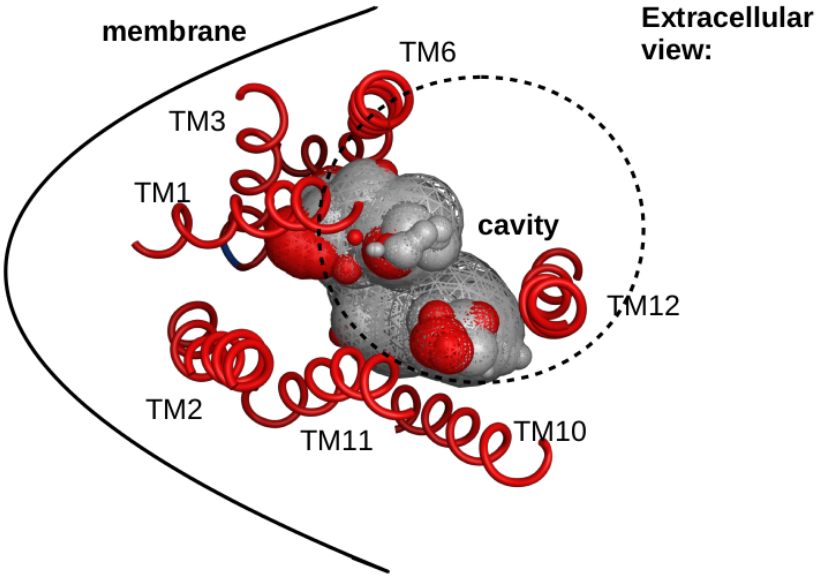


Figure 2.9: Drug-binding site 2 view from outside of cell. Only the involved transmembrane segments are shown. *Gray* spheres are hydrophobic and *red* spheres are hydrophilic. The inner cavity of P-gp is marked by a dotted circle.

The residue *T953* (TM11) located in this drug-binding site is proposed to form hydrogen bonds not only with propafenone analogs, but also with rhodamine123 [76]. Pretreatment with rhodamine B

protected the residues *G62*, *L65* (TM1), *I340*, *F343* (TM6) and *V982* (TM12) from inhibition with MTS-rhodamine.

These residues are reported to lie within or close to the rhodamine binding-site, respectively [72]. The named residues are not listed in the first binding-site, so that this drug-binding site may represent the rhodamine binding site of P-gp.

Interestingly, the identified residues of TM2, TM3, TM8, TM9 and TM10 are not explicitly mentioned to be involved in drug-binding, but are mostly close to listed residues. For example *S196* (TM3) is very close to *T199*, which is suggested to be involved in propafenone analog binding. All residues found in TM8 are within the listed residues 755-784 sequence, that is as well suggested for propafenone binding [76].

P-gp contains multiple sites for drug recognition or binding, so that the differentiation of single drug-binding sites is quite difficult (following chapter Morphing, section 3.4.2).

In the following tables 2.7 and 2.8 all residues, which are identified to be located in the drug binding site 2, are listed. Investigated residues are in bold and their suggested function with the appropriate literature is listed behind.

	Residue	Suggested function	Literature
TM1	**H61**	predicted to lie close to the tariqudar binding site	[64]
	G62	predicted to lie close to the rhodamine binding site	[72]
	L65	predicted to lie close to the tariquidar and rhodamine binding site	[64, 72]
	M68	no information about drug-binding	
	M69	predicted to lie close to the tariquidar binding site	[64]
	F72	no information about drug-binding	
TM2	V133	no information about drug-binding	
TM3	K189 M192 G193	no information about drug-binding	
	Q195	predicted to face drug-binding area	[58]
	S196	predicted to face drug-binding area	[58]
TM6	**F336**	predicted to face drug-binding area	[58]
	I340	predicted to lie close to the rhodamine binding site	[72]
	F343	predicted to lie close to the tariquidar and rhodamine binding site	[64, 72]
	S344	predicted to face drug-binding area	[58]
	Q347	predicted to face drug-binding area	[58]
	A348	no information about drug-binding	
	N350	caused reduction in the efficency of cyclosporin analogon	[75]
	S351	no information about drug-binding	

Table 2.7: Table of residues (TMD1) located in the drug-binding site 2.

	Residue	Suggested function	Literature
TM10	V865		
	I868	predicted to face drug-binding area predicted to lie in vinblastine and ciclosporin A binding site	[58] [65]
	A869	no information about drug-binding	
	G872	predicted to face drug-binding area	[58]
	E875 L879	no information about drug-binding	
TM11	F938	no information about drug-binding	
	T941	predicted to lie in propafenone analogs binding site	[77]
	F942	predicted to lie close to the tariquidar binding site colchicine, verapamil and vinblastine protected from dBBn inhibition	[64] [78]
	T945	colchicine, verapamil and vinblastine protected from dBBn inhibition	[78]
	Q946	predicted to lie close to the tariquidar binding site	[64]
	M949	no information about drug-binding	
	Y950	predicted to lie close to the tariquidar binding site	[64]
	S952	no information about drug-binding	
	Y953	predicted to lie close to the tariquidar binding site suggested to form H-bonds with propafenone analogs and rhodamine123	[64] [76]
TM12	**F978**	dramatically altered drug resistance profiles	[79]
	V982	predicted to lie close to colchicine, vinblastine, cyclosporin A and rhodamine binding site substitution leads to loss of substrate ability to inhibit labelling with IAAP colchicine, verapamil and vinblastine protected from dBBn inhibition	[69, 72] [68] [71]
	F983	substitution affects inhibition of drug transport by cis-\trans-flupentixol	[74]
	A985	colchicine, verapamil and vinblastine protected from dBBn inhibition	[71]
	M986	no information about drug-binding	
	V988	caused reduction in the efficency of cyclosporin analogon	[75]

Table 2.8: Table of residues (TMD2) located in the drug-binding site 2.

Chapter 3

Morphing

3.1 Introduction

P-gp undergoes large conformational changes powered by binding and hydrolysis of ATP. These conformational changes alter the transmembrane domains from open to the cytoplasm (NBDs are fullly disengaged, NBD-open structure) to open to the extracellular space (NBDs are fully engaged, NBD-closed structure). The movement of substances across the membrane from the cytoplasmic to the extracellular space may be coupled to these conformational changes.

As mentioned in section 2.1 crystal structures of mainly two conformations are available. To understand the molecular process of the large conformational changes and drug transport of the catalytic cycle of P-gp, transition structures are needed. Therefore different approaches are possible.

3.1.0.1 Targeted molecular dynamics

Frequently used is a molecular dynamics approach named targeted molecular dynamics. As molecular dynamics was performed in the further work, it will be discussed in Chapter 4 so only a small overlook is given now. By molecular dynamics, interactions and movements of proteins can be studied, but normally sampling of big molecules and conformational changes is unlikely due to large energy barriers and computational time. In targeted molecular dynamics the current positions of each atom of the protein is steered toward target coordinates derived from crystal structures of different conformations. A force vector is calculated to decrease the distance between the present and the target positions [80]. Due to lack of crystal structures of different conformations, this method was not implemented.

3.1.0.2 Linear interpolation

Another approach is to generate transition structures by a simple linear interpolation. In a linear interpolation unknown points as conformations of a protein can be generated if they are within the range of known points (known conformations).

This interpolation needs to be well analysed, as the generated structures may be unphysical.

3.1.0.3 Normal mode analysis

The third possibility is a normal mode approach. Normal Mode Analysis (NMA) is a less time-consuming alternative to molecular dynamics to simulate large conformational changes of proteins. Based on an analytical solution of Newton's equations of motion,

by approximating the complex dynamical behaviour of a protein as a simple set of harmonic oscillators vibrating around a given minimum energy conformation, NMA can predict the most plausible motions of a protein [81].

The minimization of the conformational potential energy as a function of the atomic Cartesian coordinates is followed by a calculation of the second derivatives of the potential energy with respect to the mass-weighted atomic coordinates. This, so called Hessian matrix is a 3N x 3N matrix, where N is the number of atoms in the molecule. Diagonalizing the Hessian matrix leads to the eigenvalues and eigenvectors. The eigenvalues and eigenvectors encode all possible motions around the initial conformation. High frequency modes represent high energy localised displacements, while low frequency modes correspond to low-energy displacements [82]. It is assumed that the low frequency modes are functionally relevant, since they may exist by evolution rather than by chance.

Normal modes can be predicted by coarse-grained normal mode or by rigid cluster normal-mode approaches (see section 3.5).

3.2 Morphing by GROMACS

In a first attempt *gmx morph* of GROMACS [83] was used to calculate a linear interpolation between the two given conformations (NBD-closed and NBD-open), where the coordinates of the intermediate x_i out of N total intermediates correspond to:

$$x_i = x_1 + \left(first + \frac{i}{N-1} \right) (last - first)(x_2 - x_1) \tag{3.1}$$

Since the intermediates are linearly interpolated structures, there are possible atom clashes, outliers and probably big tensions in the molecule, hence the morphing was performed in several smaller steps (figure 3.1).

In the first step, the two homology models of Sav1866 and mPgp were taken as starting structures to generate a certain number (e.g. 20) of intermediates, followed by a minimization of the two closest intermediates respective to the starting structures (intermediate No. 1 and No. 20).

The minimization was performed in vacuum with the steepest descent algorithm of GROMACS (Chapter 4.2.1.4) using the GROMOS53a6 forcefield. The backbone of the protein was hold by position restraints of $1000 kJ * mol^{-1}$. The final step of this first run defined the two minimized structures (intermediate No. 1 and No. 20) as new starting structures.

The next morphing run was performed with these new starting structures. In this run the number of intermediates was reduced by two. Again, the two closest structures respective to the starting structures (intermediate No. 1 and No. 18) were minimized with the steepest decent algorithm and backbone position restraints of

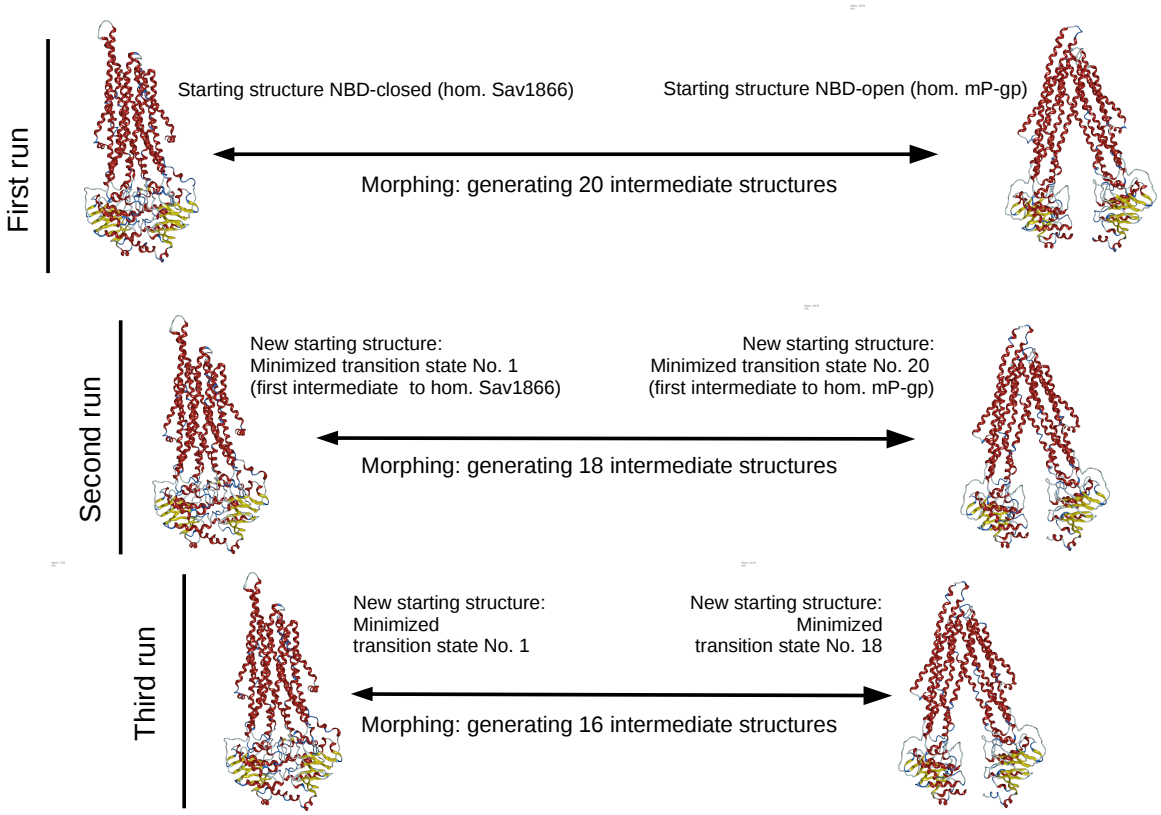


Figure 3.1: GROMACS morphing procedure

$1000kJ * mol^{-1}$ and subsequently defined as new starting structures.

This stepwise approach leaded to 18 transition states, that cover the whole movement from the NBD-open conformation to the NBD-closed conformation. The structures have a RMSD of approx. 0.5-0.6 Å to each other.

3.3 Morphing by UCSF Chimera

In an other attempt, the transitions states of P-gp were generated with *morph conformations* of UCSF Chimera [84]. In contrast to GROMACS [83] Chimera generates intermediates of two given conformations by a *corkscrew* interpolation method. The intermediates are related by as much movement as possible and by a translation along the axis of rotation.

For a run a starting and a target structure need to be defined. As the starting structure the NBD-open and as target structure the NBD-closed conformation were chosen. A single run could induce intermediate structures, but in the following, the structures could not be minimized to eliminate protein tensions and atom clashes.

Likewise to the GROMACS approach, the transition states were then generated stepwise (see figure 3.1) to reduce the mentioned tensions within the protein. Firstly 20 intermediates were generated and then the two closest respective to the starting structures were taken for minimization (Chapter 4.2.1.4). However, UCSF Chimera deleted during the morphing procedure side chains of certain residues, renamed residues and split the backbone. The deleted residue side chains needed to be added before minimizing the molecule. The *structure preparation* option of MOE [51] analysed the protein and added missing side chains. If MOE was incapable in adding the missing side chains, they were added by hand and minimized. The separated chains were connected and the bonds including the following two residues were minimized, as well. The minimization of the structure with position restrains of $1000 kJ * mol^{-1}$ on the backbone was done by GROMACS as in the GROMACS approach (see section 4.2.1.4). Atom clashes, which resulted during beginning of the morphing were completely

deleted and the clashes of the “middle” transition states were almost deleted.

Hence 18 transition states, that have a RMSD value of approx. 0.51-0.52 Å to each other, were induced.

3.3.0.4 Chimera cluster

Chimera clusters the intermediates [85] during the morphing procedure. The generated intermediates cover evenly the movement of P-gp from the NBD-open to the NBD-closed conformations, so the size of the clusters of each run should show an equal size (number of structures in a cluster).

Analysing the clusters showed, that in five of the nine morphing steps the biggest cluster included the NBD-open structures. To ensure that these clusters do not result from generation of the intermediates by the defined starting (NBD-open) to the target-structure (NBD-closed) the morphing was repeated with the NBD-closed structure as the starting structure and the NBD-open conformation as the target structure. The morphing was performed as in section 3.3. The cluster sizes were distributed as in the former approach.

At the beginning of the morphing, the biggest cluster includes then NBD-open structure, while the NBD-closed structure is within the smallest cluster. The mid-morphing structures have the same cluster size and at the end of the morphing procedure the bigger cluster is again on the side of the NBD-open structures.

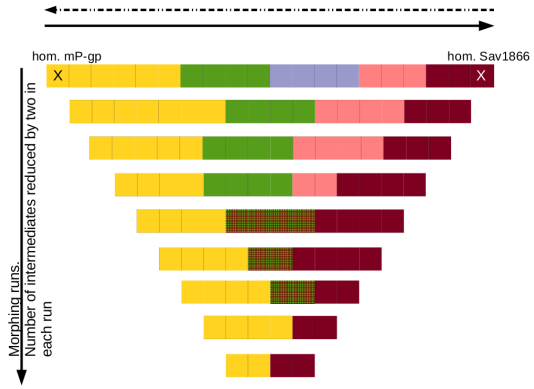


Figure 3.2: Cluster results of morphing by UCSF Chimera. The bars represent the morphing runs. Both attempts, with the NBD-open as starting and the the NBD-closed as target and vice versa, are shown. *Yellow* represents the cluster of the NBD-open conformation and *bordeaux* represents the cluster of NBD-closed conformation

.

Since the biggest cluster involve mostly these structures, the NBD-open structures seem to resemble more to each other than the NBD-closed structure intermediates. This could mean that during the transition from the NBD-open to the NBD-closed state of P-gp, the NBD-open conformational state is more stable than the NBD-closed structures, otherwise the cluster size should have been similar.

Similar observations were made in the molecular dynamics analysis, section 5.1.

3.4 Analysis of the morphing structures

Out of 18 generated intermediate structures only three intermediates from each morphing attempt (GROMACS and UCSF Chimera) were taken for further analysis. The structures were chosen as follows:

NBD-1/2-closed: The mid conformation of the movement from NBD-closed to NBD-open. This structure represents the "conversion" from the NBD-open to the NBD-closed state.

NBD-1/4-closed: This structure was chosen to be in between the NBD-closed and the mid-transition state structure.

NBD-3/4-closed: This structure was chosen to be in between the NBD-open and the mid-transition structure.

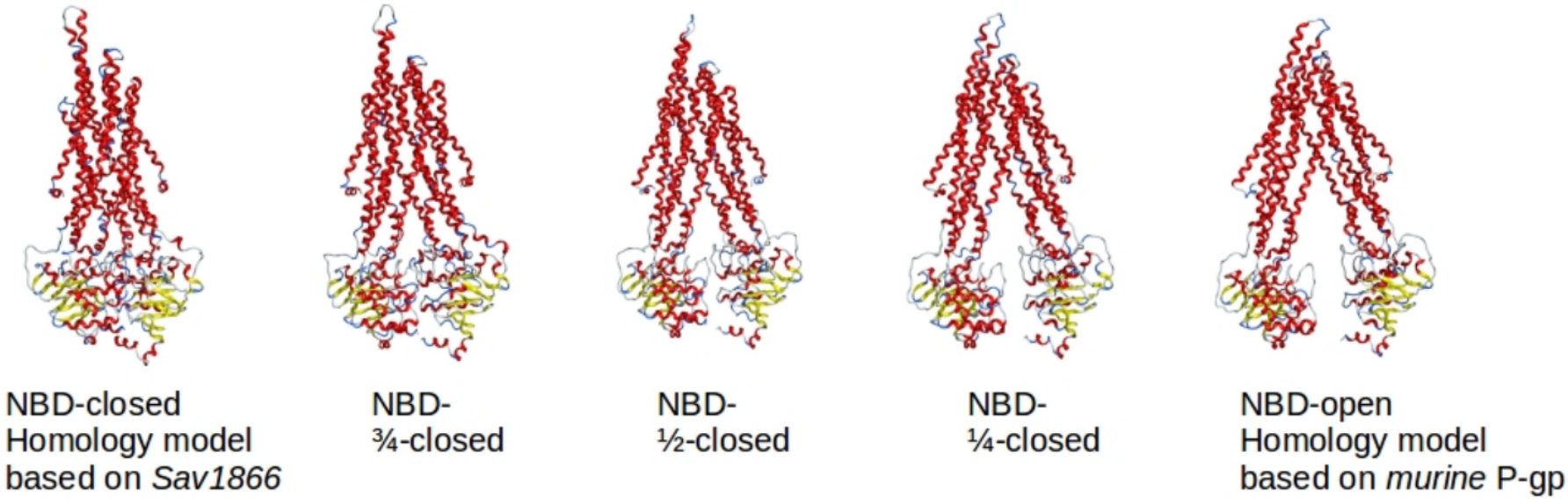


Figure 3.3: Morphing structures

As shown in table 3.1 the structures have an almost uniform distance from one another. By choosing these structures, the intention is to see in the following molecular dynamics, if they have any conformational preferences. Especially the NBD-1/2-closed structure should show a preference in movement by closing or opening the NBDs. Furthermore the structures may show a different flexibility in certain regions.

	NBD-closed	3/4-closed	1/2-closed	1/4-closed	NBD-open
NBD-closed	-	2.96	5.13	7.20	9.69
3/4-closed	2.96	-	2.33	4.37	6.84
1/2-closed	5.13	2.33	-	2.15	4.66
1/4-closed	7.20	4.37	2.15	-	2.55
NBD-open	9.69	6.84	4.66	2.55	-

Table 3.1: RMSD [Å] of the five conformations

3.4.1 Ramachandran plot

As described in Chapter 2.3.1 the *Ramachandran plot* represents the quality of the 3D structure of the protein. The results of the morphing intermediate structures are listed in the tables above:

At first glance the structures generated by UCSF Chimera seem to be a little better than the GROMACS structures, but at the second glance both methods generated transition states with the same quality. The percentage of residues in disallowed regions is slightly different (0.1%).

Although UCSF Chimera used an algorithm in which the movement of the residues is related to each other, the structures needed more effort to eliminate the atom clashes and to minimize and reduce the outliers, while the GROMACS structures were almost clash free.

GROMACS morphing transition states

		NBD-$^1/_4$-closed	**NBD-$^1/_2$-closed**	**NBD-$^3/_4$-closed**
Procheck	most favoured	88.1%	88.2%	89.4%
	additionally allowed	9.7%	9.4%	9.8%
	generously allowed	0.9%	1.0%	0.8%
	disallowed	1.3%	1.3%	0.0%
MOE	most favoured	91.5%	90.6%	91.9%
	allowed	8.0%	8.5%	7.6%
	disallowed	0.5%	0.9%	0.5%

Table 3.2: Ramachandran plot results of the morphing transition states generated by GROMACS

UCSF Chimera morphing transition states

		NBD-$^1/_4$-closed	**NBD-$^1/_2$-closed**	**NBD-$^3/_4$-closed**
Procheck	most favoured	90.2%	89.0%	90.2%
	additionally allowed	7.6%	8.7%	9.1%
	generously allowed	1.0%	0.9%	0.6%
	disallowed	1.2%	1.4%	0.1%
MOE	most favoured	93.2%	92.3%	93.2%
	allowed	6.4%	6.7%	6.1%
	disallowed	0.4%	0.9%	0.7%

Table 3.3: Ramachandran plot results of the morphing transition states generated by UCSF Chimera

For the elimination of the atom clashes, the clashing atoms were selected and minimized. This step usually leaded to a success. To reduce the number of outliers, they were also selected and minimized. In a second step the remaining outliers and the three closest residues to both sides of the peptide chain were selected and minimized, as well. By this procedure the number of outliers could be reduced drastically, e.g. the 16 outliers of NBD $^1/_4$-closed structure

generated with UCSF Chimera were reduced to five.

For further investigations the GROMACS generated transition state structures were used. In contrast to UCSF Chimera the generated 18 intermediate structures were clash free, besides the three mid-transition structures. Considering that they are "just" linear interpolated structures, their quality is comparable to the UCSF Chimera structures and in regard of possible new and better crystal structures in the future, they are by far means easier to generate, without any elimination problems.

3.4.2 Drug-binding sites

In the following section the drug-binding sites of the P-gp transition states are discussed. To built on the results of the mPgp based homology model, this analysis starts with the NBD-$^1/_4$-closed structure. All named residues were additionally identified in the particular conformational state.

3.4.2.1 Drug-binding site NBD-$^1/_4$-closed structure

The two putative drug-binding sites identified in the NBD-open structure, are in this NBD-$^1/_4$-closed structure not clearly separated.

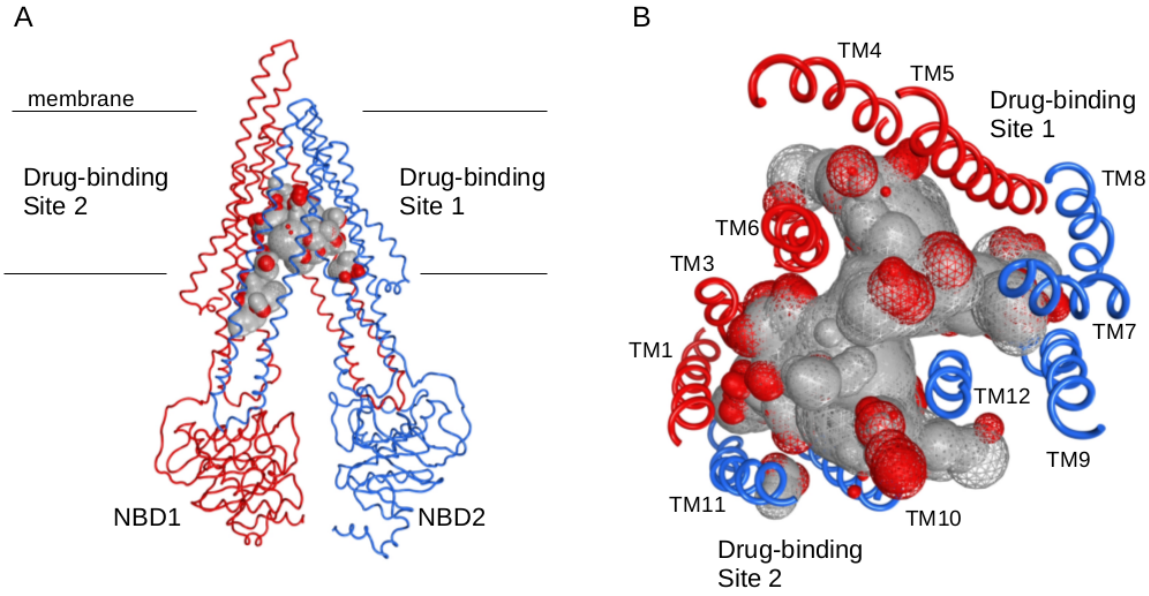


Figure 3.4: Morphing structure NBD-$^1/_4$-closed

The suggested H-site is not elongated anymore, but more spherical shaped. It is formed by six pockets and has a hydrophobic core, which is well exposed to the inner cavity of the protein, and hydrophilic smaller sections adjacent to the TMs. The binding site is formed by TM1, TM3, TM4, TM5 and TM6 of the TMD1 and TM7, TM8, TM9 and TM12 of TMD2. To the named residues of TM3 in table 2.5 chapter 2.4.3 the residues *M192, F193, Q195, S196* and *S196* which were suggested to form the R-binding site in the NBD-closed structure were additionally identified. *Q195* and *S196* were predicted in mutagenesis studies to face the translocation pathway. The overlap of the two binding sites results from closure of the NBDs, the movement of the TMs and the resulting decreasing inner cavity.

The second drug-binding site is formed by TM1, TM2, TM3 and TM6 of TMD1 and TM10, TM11 and TM12 of TMD2. This site is composed by six binding pockets, where the suggested gate located between TM10/TM12 increased, while the binding site formed by the remaining five pockets is spheric with a hydrophobic core as in binding site 1.

Notability the two binding sites look very symmetric to each other.

3.4.2.2 Drug-binding site NBD-1/2-closed structure

The biggest identified pocket in this conformational state includes residues from both above mentioned drug-binding sites.

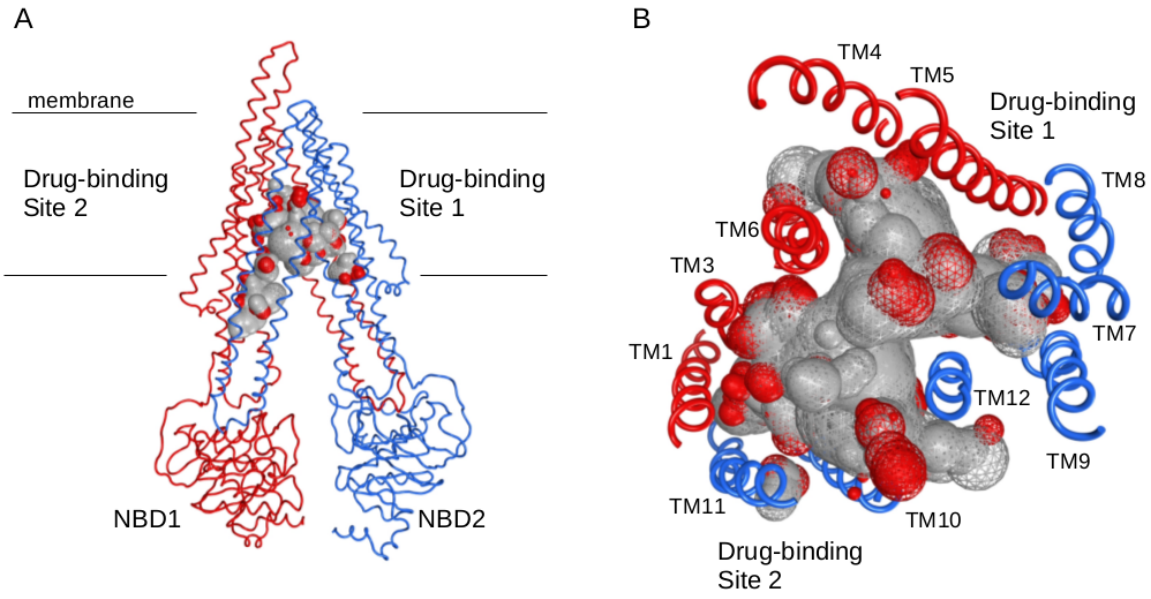


Figure 3.5: morphing structure NBD-1/2-closed

Despite the bigger hydrophobic core, which is well exposed to the inner cavity, the putative H-site is still clearly formed by TM1, TM3, TM4, TM5 and TM6 of the TMD1 and TM7, TM8, TM9 and TM12 of TMD2. New residues identified to lie within the drug-binding site are located in the TM4, of which *E243, L244* and *Y247* are also in the NBD-3/4-closed and in the NBD-closed structure [53] binding-site.

The second drug-binding site did not changed much. The bind-

ing pocket forming the gate of TM10 and TM12 has a bigger size, while the remaining pockets are exposed to the inner cavity opposing the H-site. The hydrophilic regions towards the transmembrane segments are still remaining. A smaller drug-binding pocket protrudes the membrane plane towards the NBD1 to the inner of the cell.

3.4.2.3 Drug-binding site NBD-3/4-closed structure

The NBD-3/4-closed structure represents a state of P-gp in the catalytic cycle in which the drug release may take part. The biggest pocket is exposed to the extracellular space. It involves residues, which were identified to compose the putative H- or R-site.

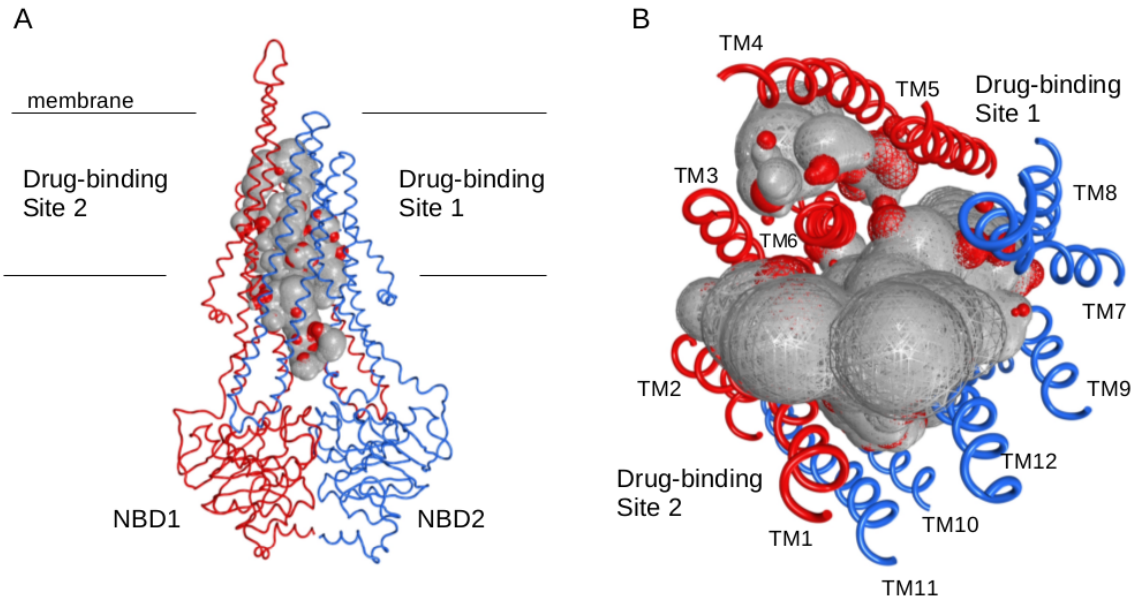


Figure 3.6: Morphing structure NBD-3/4-closed

As a result of the conformational changes of P-gp and resulting changes in the TMs, the drug-binding site 1 is not only exposed to the extracellular space, but also separated into two parts. Notability the residues of the smaller pocket formed by TM3, TM4, TM5 and TM6 overlap with the amino acids reported to be located close to the tariquidar binding site [64]. The pocket is hydrophobic with hydrophilic section towards the second pocket.

The second pocket is defined by TM2, TM3, TM5, TM6 of TMD1 and TM7, TM8, TM9, TM11 and TM12 of TMD2. The residues *C137, L138, C141* (all TM2) and *S180, K181, N183* and *E184* (TM3) were also identified in the NBD-closed structure [53] where *C137* could be protected from cross-linking by hoechst33342 and vinblastine [86].

The binding pocket of drug-binding site 2 forming the gate of TM10 and TM12 has disappeared, while the remaining pockets are building a cavity exposing the extracellular space. The hydrophilic regions towards TM6 are still remaining, while the cavity is hydrophobic. The small drug-binding pocket, which protruded the membrane plane towards the NBD1 to the inner of the cell is gone.

In all three morphing structures the putative H-site, which corresponds to the binding-site of the co-crystallized ligand QZ59-RRR is detectable. This confirms the binding site 1 to be suitable for docking with different ligands. As this binding site changes its size and orientation during the conformational changes of P-gp, the molecular dynamics will show different behaviour of the ligand to the protein and vice versa.

3.5 Normal mode by NMSim Server

The normal mode approach was done by the NMSim Server [87]. Normal modes can be predicted by coarse-grained approaches as the elastic network model (ENM), where a protein is represented as a 3D network. Particles e.g. C_alpha atoms have simplified potentials in terms of Hookean elastic springs. The elastic network model can be combined with concepts of rigidity which lead to a multi scale approach of rigid cluster normal mode analysis (RCNMA). In RCNMA the structure is considered to be constructed furthermore of rigid areas (rigid bodies). The protein is checked for structural rigidity and is partitioned into rigid regions and flexible joints between them. The rigid regions can vary in size and are allowed only translational and rotational degrees of freedom (no relative motions within a rigid region).

The normal modes calculated for the structure are followed by movement along directions of low-frequency normal modes. Out of this movement a conformation is generated and a new set of normal modes is then calculated using the previously generated structure.

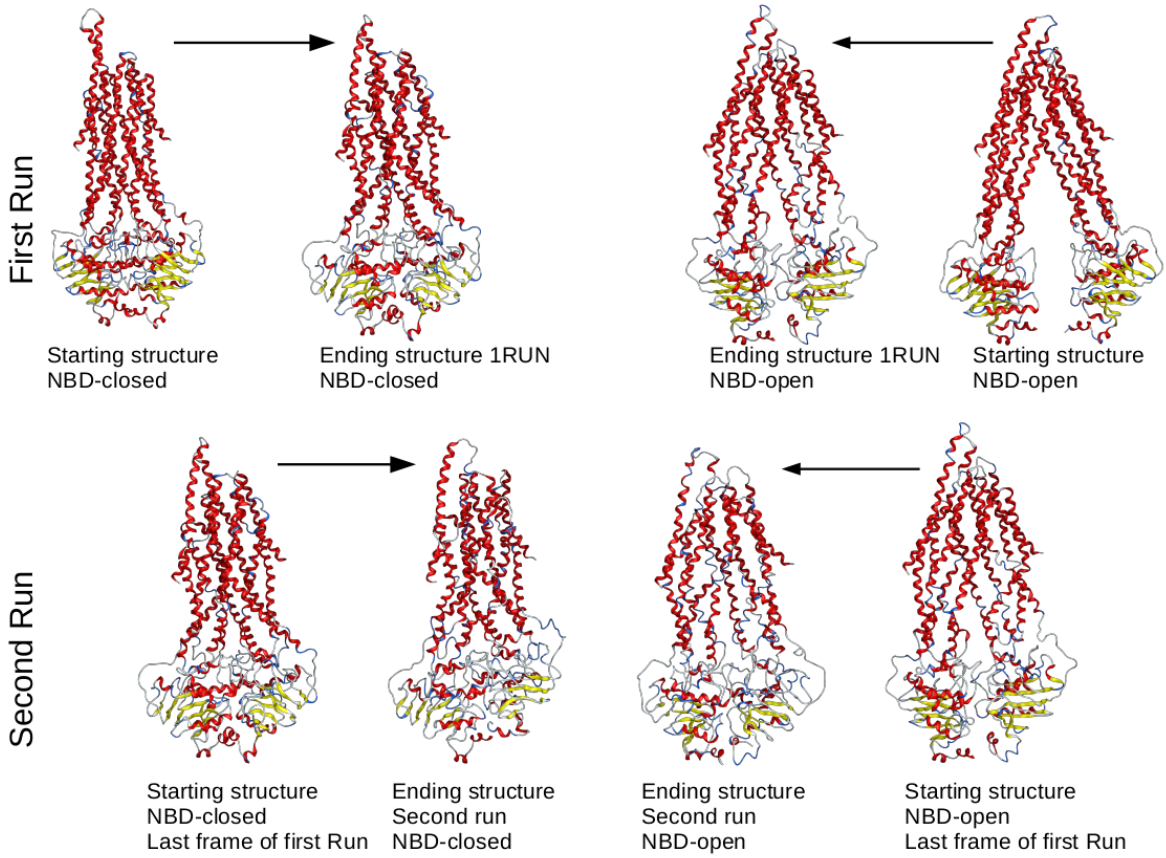


Figure 3.7: Normal mode approach

NMSim server accepts only proteins with a continuous amino acid sequence, however both homology models are missing the intermediate linker sequence (residues 631-687). This lacking 56 amino acids were inserted as a connecting chain outside the protein (see figure 3.8).

All simulations were performed as *targeted simulation*. As a result of an unsuccessful run generating a continuous conformational change from the staring to the target structure, the simula-

tions were performed twice. Once with the NBD-closed structure as starting point and the NBD-open structure as ending point and vice versa.

In the first implementation, the targeted simulation was chosen as a rigid cluster normal mode analysis. The simulation parameters were kept by default, besides the number of simulation cycles and the step size was varied. By the step size, the distortion of the structure along the normal mode directions can be chosen. A number of intermediates could be generated by this method, however not continuously to the target structure. Furthermore by increasing the step size to its maximum of 2Å or by increasing/decreasing simulation cycles, there was no success. Therefore the run was repeated with the two intermediates closest to each other (see figure 3.7). In this second run, the gap remained the same between the two conformations closest to each other. The calculated normal modes seem not to overcome the big conformational changes of P-gp.

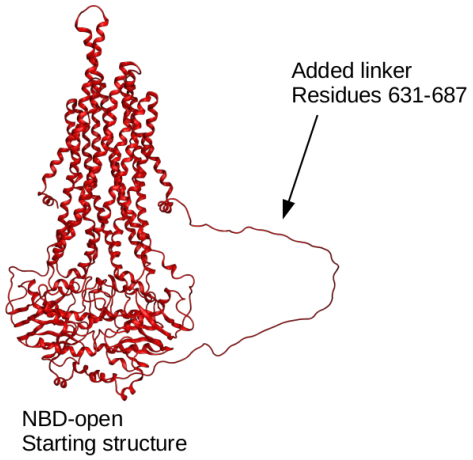


Figure 3.8: Inserted linker for the normal mode approach. The missing 56 residues were added to connect the two homolgous parts

In an additional running the elastic network model was chosen for the *targeted simulation*. The parameters were kept by default, besides the number of simulation cycles and the step size. Here too, a variation of step size and simulation cycles could not generate continuous intermediates.

The normal modes are calculated from a local minim state of the protein, therefore the local minimum may not represent the “real” conformation of P-gp.

Since P-gp hydrolyses ATP, an energy barrier may separate the NBD-open conformation from the NBD-closed state. The normal modes of the whole protein may not capture the relationship of the flexibility of the single helices of the two homologous pars of P-gp.

3.5.1 Analysis of the generated structures

3.5.1.1 Rigid cluster decomposition

All atoms of the starting structure (for the NBD-open and the NBD-closed structure, respectively), which are part of a rigid cluster are listed in the following table. By the rigid cluster the flexible and inflexible parts of P-gp can bee seen.

As shown in tables 3.4 and 3.5 there are different rigid cluster decompositions, not only between the two conformations, but also between the two subsequently runs.

The NBD-closed structure did not show as many rigid regions as the NBD-open conformation.

The first run showed that the NBD-open structure starting structure has its biggest rigid region in the TMD2 part (TM9/TM10), while the NBD-closed starting structure showed the rigid region in TMD1 (TM4/TM5).

In the second run the rigid cluster decomposition changed drastically. The NBD-open structure reduced the biggest rigid region from 73 involved residues to 60 residues (TM2) of NBD1. The protein was divided in 25 rigid region adding the intra-cellular-loop 1 (ICL1) and splitting the former biggest rigid region TM9/TM10 into four smaller parts.

NBD-open structure (homology of mP-gp)			NBD-closed structure (homomology of Sav1866)		
sequence	**residue No.**	**clustersize**	**sequence**	**residue No.**	**clustersize**
TM1	44-78	35	TM1	48-89	42
TM2	104-158	55	TM2	94-125 126-157	32 32
			ICL1	159-161	3
TM3	167-184 187-211	18 25	TM3	167-183 184-211	17 28
TM4	213-259	47	TM4/TM5	212-325	114
ICL2	260-267	8			
TM5	269-282 287-325	14 39			
TM6	327-371	45	TM6	327-375	49
TM7	709-741	33	TM7	696-737	42
TM8	747-799	53	TM8	743-799	57
ICL3	800-805	6	ICL3	800-805	6
TM9	806-810 811-829	5 19	TM9	806-851	46
TM9/TM10	830-902	73	TM10	854-902	49
ICL4	903-910	8			
TM11	912-967	56	TM11	912-968	57
TM12	971-992 994-1013	12 20	TM12	970-1014	45

Table 3.4: Rigid cluster decomposition of the **first RCNMA run** by order. The clustersize describes the number of residues in one cluster. The two conformations differ in their rigid cluster decomposition. The NBD-open conformation has the biggest cluster as the TM9/TM10 section, while the NBD-open conformation has the biggest rigid cluster as the TM4/TM5 area. (NBDs are excluded)

NBD-open structure (homology of mP-gp)			NBD-closed structure (homomology of Sav1866)		
sequence	**residue No.**	**clustersize**	**sequence**	**residue No.**	**clustersize**
TM1	39-79	41	TM1	51-65 66-89	15 24
TM2	99-158	60	TM2	94-124 125-153	6 29
ICL1	159-163	5			
TM3	170-179 185-207	10 23	TM3	168-181 183-206	14 24
TM4	213-221	9	TM4/TM5/ TM6	212-374	163
ICL4	261-266	5			
TM5	269-285 288-324	17 37			
TM6	327-371	45			
TM7	708-741	34	TM7	709-730	22
TM8	747-778 779-798	32 20	TM8/TM9	743-848	104
TM9	820-851	32			
TM10	855-875 876-889 890-902	21 14 13	TM10	854-873 886-901	20 16
ICL4	903-911	9			
TM11	912-930 931-966	19 36	TM11	903-923 925-952	21 28
TM12	971-982 986-992 995-1013	12 7 19	TM12	971-994 995-1013	24 19

Table 3.5: Rigid cluster decomposition of the **second RCNMA run** by order. The clustersize describes the number of residues in one cluster. The NBD-open conformation has the biggest rigid cluster in the TM2 section (60 residues), while of the NBD-closed structure shows the biggest region spanning the TM4/TM5/TM6 area (163 residues). (The NBDs are excluded)

In contrast the NBD-closed structure was not additionally divided, but rather changed its rigidity. The biggest rigid region increased size by adding TM6 involving 163 residues (former size was 114). The ICL1 was excluded of the rigid cluster decomposition, so that it is considered as a flexible region. While the TM8 and TM9 were merged the remaining TMs were separated into two parts.

This contrasting rigid cluster decomposition of the two structures NBD-open and NBD-closed may indicate P-gp to have changing rigid regions and flexibility due to the conformational state of the protein.

On the other hand, this may be a reason why the conformational change of P-gp was not successful by the NMSim Server approach. The protein can not reach the target conformation without changing the rigidity and flexible joints.

3.5.1.2 Superposition of the generated structures

Superposing the two conformations of P-gp, the RMSD (root mean square deviation, see section 2.3.2) is 9.69 Å. By generating intermediates and a dimerization of the NBDs the RMSD value should decrease from the NBD-open to the NBD-closed structure as shown by the morphing structures in section 3.4. Interestingly the RMSD values of the generated structures are not decreasing in the second run. The second run was performed to close the gap between the two starting conformations, so the generated structures should show a smaller RMSD than in the first run.

		NBD-open structure			**NBD-closed structure**		
		starting structure	last structure run 1	last structure run 2	last structure run 2	last structure run 1	staring structure
	starting	-	7.23	7.68	10.55	9.33	9.69
NBD-open	last run1	7.23	-	4.75	8.20	7.47	7.16
	last run2	7.68	4.75	-	7.74	5.97	7.36
	last run2	10.55	8.20	7.74	-	5.19	6.69
NBD-closed	last run1	9.33	7.47	5.97	5.19	-	5.22
	starting	9.69	7.16	7.36	6.69	5.22	-

Table 3.6: RMSD [Å] of the generated structures

For example the ending NBD-closed structure compared to the NBD-open structure has an RMSD value of 10.55 Å, while in the first run it was 9.33 Å. This value is even bigger than the difference of the starting conformations.

Generally this RMSD values do not give an accurate picture. The inconsistency is not only because the NBDs approach each other only to a certain point, but also, because there is a loss of helical structure in the protein, not only at the points, where the rigid cluster decomposition split the helices, but also within the

rigid parts. This loss has an effect on the RMSD.

As visible in figure 3.9 the broken helices are planar to each other and span from the membrane surface to the mid-membrane level.

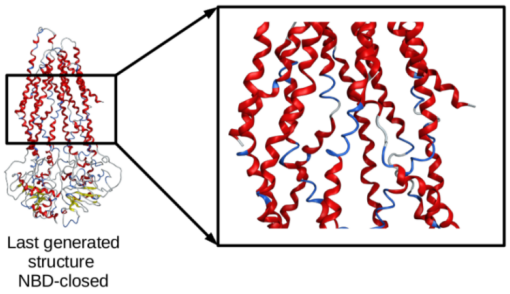


Figure 3.9: Although the TM regions were identified as rigid bodies, they show at certain areas a loss of helical structure.

Noticeable, also in the molecular dynamic simulations 5.2.1 loss of helical structure appeared, especially within TM3 (*Lys181-Met192*) and TM9 (*Asp821-Ile829*).

3.5.1.3 Joints in the structures

P-gp needs for its big conformational change flexible regions within the protein. Therefore firstly the joints of the rigid cluster decomposition were examined. Joints are parts in the protein between the rigid regions.

All four rigid cluster decompositions showed inconsistent joints (shown in table 3.4 and 3.5), only TM3 has a joint in all four rigid cluster decompositions created by the residues *Asn183* and *Glu184*. This two residues appear also in the loss of helical structure of TM3 in section 5.2.1. Therefore the joints were not examined separately, but all marked in one structure to clarify the regions and levels as a whole.

There were three different levels in which joints appeared. One level in the membrane section (TM1, TM2 and TM5), one on the membrane-cell border (TM10 and TM12) and the third and biggest one in the intracellular space (TM3, TM5, TM10 and TM11).

The joints were better visible in the NBD-open structure rather than in the NBD-closed structure.

Chapter 4

Molecular Dynamics

4.1 Introduction

4.1.1 Theory

Molecular dynamics (MD) is a method for studying the time dependent behavior of a molecular assembly by taking their interactions into account. These particles can correspond to atoms, but they may also represent distinct entities (e.g. chemical groups). Newton's equation of motion is used to determine the net force and acceleration experienced by each particle:

$$F_i = m_i a_i \tag{4.1}$$

F is the force exerted on the particle i, m is the mass and a is the acceleration of the particle i.

The Force F can also be expressed as the negative gradient of a potential energy:

$$F_i = -\nabla_i V \tag{4.2}$$

the potential energy V, is the energy of the system (described by a

force field) and is related to the changes of positions of the particle i as a function of time. This movement of particles during a MD is the trajectory. The equation is solved numerically at discrete time steps (e.g. Verlet algorithm)

4.1.2 Force Fields

A force field describes the geometric properties of interactions of molecules. A molecule is described as a series of charged points (atoms) linked by springs (bonds). These molecular properties can be classified in two groups: bonded and non-bonded interactions:

$$E_{total} = E_{bonded} + E_{non-bonded} \tag{4.3}$$

The sum of all bonded pairs results from following terms:

$$E_{bonded} = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi[1 + \cos(n\chi - \sigma)] \tag{4.4}$$

The first term is modelled as a harmonic potential according to Hooke's law, where K_b describes the stiffness and b_0 the equilibrium state. The second term describes the deviation of an ideal bond angle geometry. K_θ represents the angle bending constant and θ_0 the equilibrium state. The last term shows the steric barrier between atoms separated by three covalent bonds. The periodic (*cos*) rotation arround the middle bond is the dihedral angle χ. K_χ is the rotational barrier, n is the periodicity and σ the phase of the rotation.

Non-bonded interactions are interactions of atoms in the same molecule that are separated by three or more bonds, or interactions of atoms of different molecules.

$$E_{non-bonded} = \sum_{ij} \left(\epsilon_{ij} \left[\left(\frac{R_{min,i}}{r_{ij}} \right)^{12} - 2 \cdot \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right) \tag{4.5}$$

The first part is the Lennard-Jones equation. The atoms feel an increasing attraction with decreasing distance, until an energy minimum is reached. At closer distances the increasing energy leads to a repulsion. The pre-factor ϵ is a parameter based on the atoms *i* and *j*. R_{min} defines the energy minimum.

The second part is Coulomb's law for electrostatic interaction between non-bonded pairs of atoms. The interactions of two atoms *i* and *j* with charges *q* decay slowly with increasing distance *r*. Calculating these distances is the most time consuming part of the MD. To speed up the calculation cut-off distances are used, which define greater distances to be ignored.

The Force field specifies not only the parameters of this equations, but also the atom types. An atom type is defined by its number of atoms, hybridisation state and environment (e.g. bare sp^3 carbon bound with four heavy atoms or in an aliphatic CH_3 group).

There are three different force field types:

1. **all atom force fields:** every single atom in the system is specified (e.g. OPLS-AA, CHARMM22 [88, 89])

2. **united atom force fields:** all atoms are specified except the non-polar hydrogens (e.g. OPLS-UA, CHARMM19 [88, 89], GROMOS [90])

3. **coarse grained force fields:** multiple atoms are combined

into one interaction site (e.g. MARTINI [91]). Coarse grained force fields allow a simulation of larger systems for a longer time, however, this is achieved at the expense of atomic resolution.

In this thesis the united atom forcefield GROMOS53A6 was used.

4.2 Moclecuar dynamincs run

The molecular dynamics simulations were performed using GROMACS 4.6 and 5.1 [83] with the gromos53a6 force field. The following procedure was done for all structures (the two homology models NBD-open and NBD-closed and the generated intermediate states NBD-$^1/_4$-closed, NBD-$^1/_2$-closed and NBD-$^3/_4$-closed).

4.2.1 Building the simulation system

The ambition of molecular dynamics is to simulate the physiological environment as much as possible.

4.2.1.1 Inserting P-gp into a membrane

A POPC (1-Palmitoyl-2-oleoylphosphatidyl-chinoline) bilayer including 512 molecules was selected as an appropriate membrane. The POPC membrane has a phase transition (liquid crystalline phase) at a simulation temperature of 310 K, which complies to the human body temperature.

Before inserting P-gp into the membrane, the optimal position of the protein in the bilayer had to be calculated. This was done by *lambada-align* [92], which aligns the protein and membrane based on an optimal rotation and translation by a hydrophilicity profile-derived scoring function.

InflateGRO [92] and g_membed (GROMACS implementation) [93] are two applications for inserting a protein into a membrane, which differ in their implementation. InflateGRO inflates the membrane, while the protein remains constant, and deletes in the shrinking process overlapping lipids based on the area occupied by the pro-

tein.

Differently, g_membed of GROMACS, which was chosen, scales the coordinates of the protein with respect to the geometrical center by 0.1. Subsequently every molecule of the equilibrated membrane overlapping the down-sized protein is deleted. Then the protein size increases stepwise to 1.0 again. Each membrane molecule that has contact to any protein atom during this procedure is deleted, too. The big advantage of g_membed in comparison to inflateGRO is, that the equilibrated membrane remains as it is, while InflateGRO needs to bring back the membrane to natural dimensions, which means that the equilibration is not preserved [92].

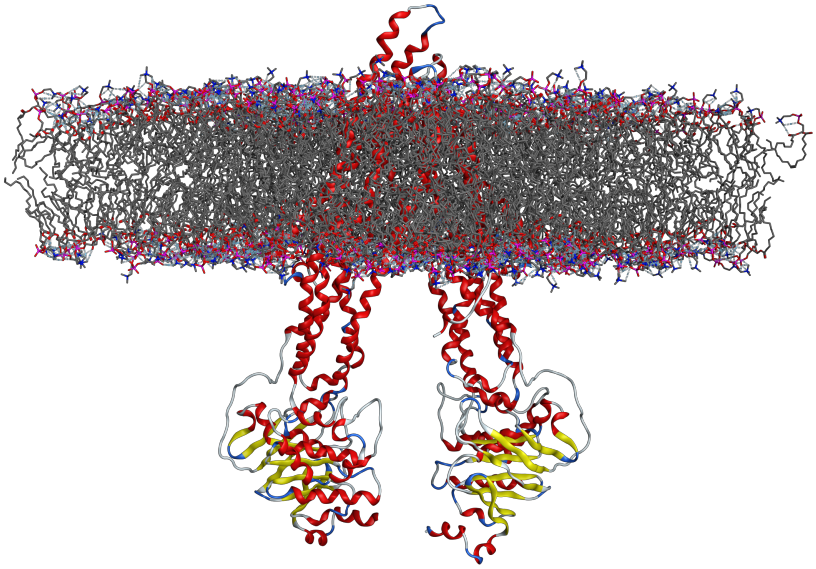


Figure 4.1: NBD-open P-gp conformation embedded in POPC bilayer membrane.

For each molecular dynamics simulation of the several conformations of P-gp (NBD-open, NBD-closed and the three transition states), the protein was inserted into the equilibrated POPC membrane (figure 4.1). Since the conformations fill the membrane space differently 27-31 POPC molecules were deleted (see table 4.1).

4.2.1.2 Periodic boundary conditions of the simulation box

The whole protein membrane system was in a cubic simulation box with periodic boundary conditions (PBC), which replicate the simulation box infinitely in all directions. This procedure enables the simulation to be performed by using a relatively small number of particles in such a way, that the particles experience forces as though they were in a bulk solution. Therefore there are no surface

effects like artifacts of a finite system. Particles crossing a boundary of the simulation box emerge back from the opposite side. The size of the membrane defined the X and Y axes of the cubic box. The Z axis size was defined by the protein. Here it is important to ensure, that the box vector is big enough for the protein not to "see" its mirror image, yet the box vector should be as small as possible to keep the computational cost manageable.

4.2.1.3 Solvation of P-gp embedded in the membrane

Natural environment of proteins features not only membranes, but also water and ions. There are various water models of which only three will be mentioned. SPC (simple point charge) water [94] and TIP3P (transferable intermolecular potential) [95] water have just three interaction sites centered on the nuclei. The geometry of the water molecule is planar, while the TIP5P model [96] is almost tetrahedral by adding two lone pair sites, which lead to two additional interaction sites and an increasing computational cost. Therefore, to keep the calculation time as small as possible the SPC water was chosen. A physiological environment should be isotonic, therefore sodium chloride had to be added. An isotonic sodium chloride solution has a concentration of 0.154 mol/L. Since water has a molarity of 55.56 mol/L, one sodium chloride molecule comes to 361 water molecules. The P-gp homology model had a negative charge, which was also balanced by adding NaCl.

The following table 4.1 lists the amount of molecules, which the different simulations of the P-gp conformations embedded in the POPC membrane and solvated in water had.

The P-gp NBD-$^1/_2$-closed conformation had the compactest structure of all conformations, with the NBDs in mid distance and the

Conformation of P-gp	amount of molecules POPC	water	sodium	chloride
NBD-closed	472	76512	206	218
$^3/_4$-closed	476	81504	221	233
$^1/_2$-closed	478	82112	222	234
$^1/_4$-closed	472	74874	204	216
NBD-open	474	74205	200	212

Table 4.1: Amount of molecules of each simulation.

TMs lying closely together. Therefore this MD system showed the smallest number of deleted POPC molecules and the highest amount of water molecules, followed by the NBD-$^3/_4$-closed conformation.

4.2.1.4 Minimization

By inserting P-gp into the POPC membrane and adding SPC water into the simulation box the initial geometry of the molecules did not necessarily corresponded to stable conformers, and atoms could lie too close together, so that the energy could be raised above natural energy levels, which leads to local distortions. Therefore, energy minimization was not only carried out to determine a stable conformer (lower energy conformation), but also to correct the system and reduce flaws, as steric clashes and distorted bond angles and lengths, which would result in an unstable simulation.

Energy minimization is a numerical procedure for finding a minimum on the potential energy surface starting from the higher energy initial structure. The *steepest descent* method, which was used in this thesis, determines the energy and the slope of the function at the starting point. If the slope is positive, it is an indication that the potential energy of the conformation is too large

and the conformation is then adjusted. The second minimization is carried out by determining again the energy and the slope starting from the last point and moving steepest to the remaining direction. This process continues until a minimum has been reached.

Since the simulation system contained almost over 280'000 atoms, the minimization was done in several smaller steps. Firstly, only the water molecules were minimized and the rest of the system was held by position restraints of 1000 $kJ * mol^{-1}$, doing so the particles remained at fixed positions, thus they could move, but only with a large energy penalty. Restraining one part of the simulation system, avoided drastic rearrangements of the system (motions of the protein), which can occur, if the added solvent which is not yet equilibrated subjects large forces to the protein and membrane. After minimizing the water molecules, the POPC membrane was included in a second minimization step, with P-gp held by position restraints of 1000 $kJ * mol^{-1}$.

The minimization ensured that the solvent and membrane had a reasonable structure, in terms of geometry and orientation, which results in a system with an energy near the lowest possible. Thus, this state corresponds to a temperature near 0 K, where no motion can be seen [97].

4.2.2 Equilibration of the system

Having a temperature of approx. 0 K, the simulation system needed to be brought up to the temperature of interest 310 K. Therefore a short 100 ps NVT ensemble (constant number of particles, volume and temperature) with position restrains on the protein of 1000 $kJ * mol^{-1}$ was performed, as calculation of pressure is inaccurate at low temperatures. After reaching the temperature, the pressure was conducted under a NPT ensemble (constant number of particles, constant pressure of 1 bar and constant temperature of 310 K).

For the non-bonded interactions a cut-off of 0.9 nm was used, while the Particle Mesh Ewald method [98] was employed to capture the non-bonded interactions at longer distances. Temperature and pressure coupling were controlled using the Nose-Hoover [99] thermostat. The SHAKE [100] algorithm constrained bonds and the highest frequency motions involving hydrogen atoms from force evaluation.

A time step of 2 fs was used during the equilibration and the production run time.
After the 10 ns NPT run, the protein position restraints were released stepwise, first holding only backbone atoms, following by only holding only C^{α} atoms. The whole equilibration procedure lasted for 11 ns.

4.2.3 Production run

The production run lasted for 35 ns. To verify the observations, all simulations of each P-gp conformation were repeated twice.

4.3 Inserting a ligand by molecular docking

Molecular docking includes the calculation or expectation of an interaction of a small molecule (the ligand) to a protein. In the process the ligand conformation, as well as the position and orientation within the binding site of the protein is predicted.

This prediction can be achieved through two interrelated steps: first conformations of the ligand in the binding site are sampled; then these conformations are ranked via a scoring function, which represents the quality of the docking result. The generated conformations ranked highest by the docking scores do not assure whether the ligand binds and what the affinity to the protein is. This needs still to be confirmed by experimental assays.

4.3.1 Docking preparation

Knowing the location of the binding-site before docking a ligand significantly increases the docking efficiency.

4.3.1.1 Validation

For reliable results, first the docking parameters and binding site had to be defined and verified. Thus, the co-crystallized ligand of the 4M2T crystal structure was “re-docked” into its binding site within the 4M2T structure.

As a docking program GOLD (Genetic Optimization for Ligand Docking) [101, 102] was used, which modifies and optimizes dihedrals of ligand rotatable bonds, ligand ring geometries, dihedrals of

protein OH- and NH_3 groups and the mapping of the fitting points, to fit the ligand into the binding-site.

To determine the correct poses from the incorrect ones and to rank the generated conformations the *goldscore* was used, which is a molecular mechanics-like (classical forcefield based) function, and which determines the binding energy by calculating the sum of non-bonded interactions, including protein-ligand hydrogen bonds, protein-ligand van der Waals interactions and intramolecular tensions of the ligand.

The ligand poses in the binding-site were refined using the *chemscore* [103, 104] as rescore function. *Chemscore* estimates the free energy of binding of a ligand to a protein, by considering hydrogen bonding, lipophilic interactions and entropy of the ligand upon binding to the protein.

100 solutions (docking poses suggested by GOLD) were generated using GOLD, which were analyzed by RMSD (degree of similarity to the extracted ligand) and docking positions in the drug-binding site. The best ten results were within the first 25 poses.

In this thesis docking was used to insert a ligand into the drug-binding site of the co-crystallized ligand QZ59-RRR for ongoing molecular dynamic simulations. Therefore re-docking the QZ59-RRR ligand verified the docking parameters (*goldscore* and *chemscore*) for the further proceedings.

4.3.1.2 Defining the drug-binding site of the P-gp homology-model

To transfer the docking protocol to the homology model, the crystal structure PDB-ID: 4M2T with the co-crystallized ligand QZ59-RRR

was aligned with the NBD-open conformation of P-gp. This alignment placed the ligand into the putative drug-binding site of the homology model. By saving the coordinates of ligand and protein, GOLD recognizes the drug-binding site, so that the ligand can be extracted and the active docking site can be defined. The docking was performed as described above in section 4.3.1.1, besides the *generate cavity file* was especially selected, to generate a drug-binding site file in which all involved residues are listed.

4.3.2 Inserting the ligand into the protein

Inserting a ligand (e.g. colchicine) into the P-gp conformation was done for each structure equally as follows:

4.3.2.1 Docking with GOLD

For docking colchicine into P-gp, the NBD-open structure was chosen first, and the drug-binding site was defined by all atoms within 6Å of the existing *cavity atoms* file (all drug-binding site atoms listed; section 4.3.1.1) generated in the former run with the QZ59-RRR ligand.

As a scoring function the *goldscore* and as rescore function the *chemscore* were selected. 25 solutions were generated and further analyzed. The colchicine position closest to the QZ59-RRR position was chosen for the molecular dynamics run.

4.3.2.2 Preparing the ligand for molecular dynamics

GROMACS needs for molecular dynamics simulations two informations: First, the coordinates (in GROMACS file format), in which the coordinates of all atoms within the system are listed and to which the velocities are assigned. Second, the topology, which specifies each atom type, mass, charge and most importantly the connected atoms of a molecule. GROMACS can only generate coordinates and topologies for proteins, and not for small molecules. Therefore ProDRG [105] was used, which translates the coordinates and the topology to a GROMACS readable format.

To extract for the ProDRG server the exact position of the ligand within the drug-binding site, the P-gp structure used for GOLD docking and the P-gp structure embedded into the membrane (in which the ligand is inserted to) were superposed. This step was necessary, because GOLD docking generated new coordinates of the ligand and the protein and if no superposition was performed the ligand with the GOLD coordinates would be placed not into the correct position or even into plain space.

The conformation of the ligand should stay consistent, thus the energy minimization option was turned off. The ligand was then included into the P-gp conformation embedded in the POPC membrane.

4.3.3 Equilibration

The procedure of the following equilibration was performed as in sections 4.2.1.4 and 4.2.2.

After positioning the ligand into the drug-binding site of P-gp, the simulation box was filled with SPC water. Then sodium and chloride ions were added until the solution was isotonic. The steric clashes and distorted bond angles and lengths were corrected by a minimization with positions restraints of 1000 $kJ * mol^{-1}$ on P-gp, colchicine and POPC. A second minimization was performed with restraints on the protein and ligand. This restrains were kept during the short 100 ps NVT and the following 10 ns NPT equilibration phase. Afterwards a small NPT simulations followed with decreasing position restrain forces on the protein. During the 11 ns equilibration time the ligand was kept at the docking position.

4.3.4 Production run

The production run lasted for 35 ns. To verify the observations, the simulations were repeated, as the ligand-free conformations, twice.

4.3.5 Analysis

Besides the above mentioned RMSD value (section 2.3.2) to analyze the degree of similarity, the conformation trajectories were analyzed by several other measurements.

4.3.5.1 RMSF

The root mean square displacement fluctuation is a measure of the displacement of a particular atom or group of atoms, relative to the reference structure, averaged over the number of atoms. It gives, in contrast to the RMSD (measure for overall displacement), insights into the flexibility of particular regions of the protein (e.g. local chain flexibility).

4.3.5.2 DSSP

DSSP (*define secondary structure* in *protein*) is an analysis program [106, 107], which reads the positions of all atoms in a protein, followed by a calculation of the hydrogen-bond energy between all atoms. As a pattern-recognition process of the H-bonded and geometrical features extracted from x-ray coordinates [106], different types of secondary structure are assigned to the protein. Out of the eight secondary structures, the three relevant for the analysis in this thesis (section 5.2.1) will be elucidated:

1. *bending-helix*: region with high curvature formed by at least five residues.

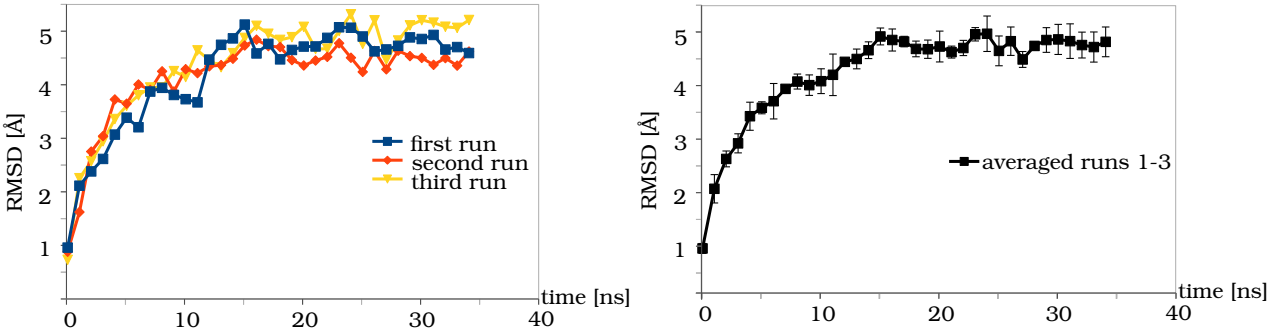
2. *turning-helix*: hydrogen bonded turn in helix.

3. *coil*: an unwinding, or a helical structure, in which dihedral angles lie not within the helix definition.

4. π-*helix*: region, which shows instead of H-bonds between the nitrogen group of residue i and the carbonyl group of residue $i+4$ (α-helix), interactions between i and the $i+5$ residue.

4.3.5.3 Data merging

All molecular dynamics simulations of the generated P-gp conformation were performed three times. To compare the resulting data for distance measurement, RMSD, RMSF, H-bonds etc. of the different P-gp conformations with each other, the results of each conducted MD run per conformations was averaged (figure 4.2).



(a) All values. **(b)** merged data

Figure 4.2: Example of data merging: the RMSD values of the three performed NBD-open conformation simulations (**a**), were merged into one graph (**b**).

Chapter 5

Analysis of the P-gp conformations

The conformation and structural flexibility of P-gp captured in the simulations was analyzed by monitoring the center-of-mass (COM) distance between the two NBDs (section 5.1), by the H-bonds between residues i and i+4 in the transmembrane helices (section 5.2.2) and by the helices themselves (section 5.2.1).

The two NBDs dimerise upon binding ATP and separate after ATP hydrolysis and because of their tight coupling to the TMDs, this results in the TMDs to alternate between the NBD-closed and NBD-open conformation [108].

In addition, cross-linking data was compared with each conformation during the molecular dynamic simulations, section 5.5.

Figure 5.1: Superposition of the NBD-open conformation before the dynamics (**yellow**) and after 35ns. **Red** shows the first, **blue** the second and **green** the third run. The differences of the NBD movement of the three simulations are clearly visible between the second and the third run.

5.1 NBD movement of the different P-gp conformations

Overall, the NBD center of mass (COM) distances of the P-gp conformations obtained from the simulations, appeared to approximate each other, or stay consistent with distances observed from 26-38 Å.

A large amplitude of motions of the NBDs showed only the ligand-free NBD-open (figure 5.1) and NBD-1/4-closed conformations. The width of the fluctuations was from 40 Å to 48 Å, with a starting value of 43 Å (see figure 5.1). The NBD-1/4-closed conformation showed a span of 6 Å (40-46 Å). These distances result the NBDs being more apart than in the murine P-gp crystal structure (PDB.ID: 4M1M) and suggesting a highly mobile nature of P-gp in the NBD-open and NBD-1/4-closed conformation.

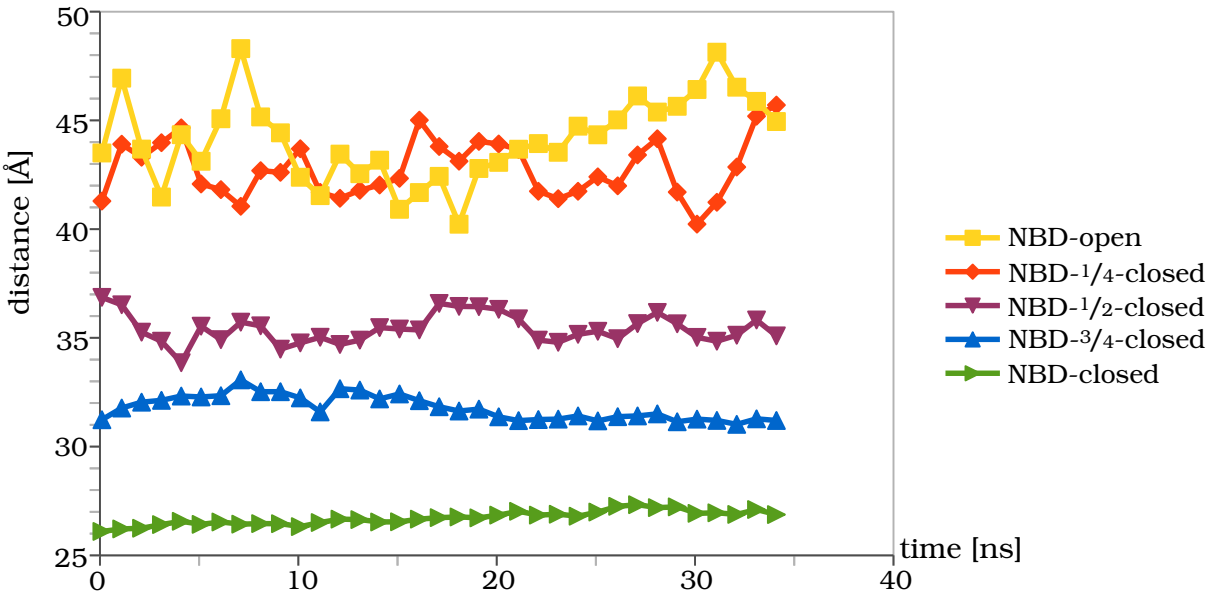


Figure 5.2: COM calculations of the NBD distances of ligand-free P-gp conformations.

The NBD movement of the remaining ligand-free P-gp conformations seemed marginally in comparison to the fluctuations of the NBD-open and NBD-$^1/_4$-closed structures. The NBD-$^1/_2$-closed conformation showed an average distance of $\sim$35 Å, while the NBD-$^3/_4$-closed reached 31.7 Å and the NBD-closed conformation stayed at $\sim$26 Å (figure 5.2).

The differences of the COM calculation starting values of the ligand-free and the colchicine docked conformations, occur from the equilibration phase of both, respectively. As mentioned in section 4.2.2 and 4.3.3 the last step of each equilibration was a small 100 ps simulation with no position restrains holding the Cα atoms of P-gp, so that a relaxation and adaption is possible.

Colchicine facilitated a closure of the NBDs of NBD-open and NBD-$^1/_4$-closed conformation with both reaching a NBD-NBD distance of 38 Å. This result was achieved in the first 10 ns of the simulation time.

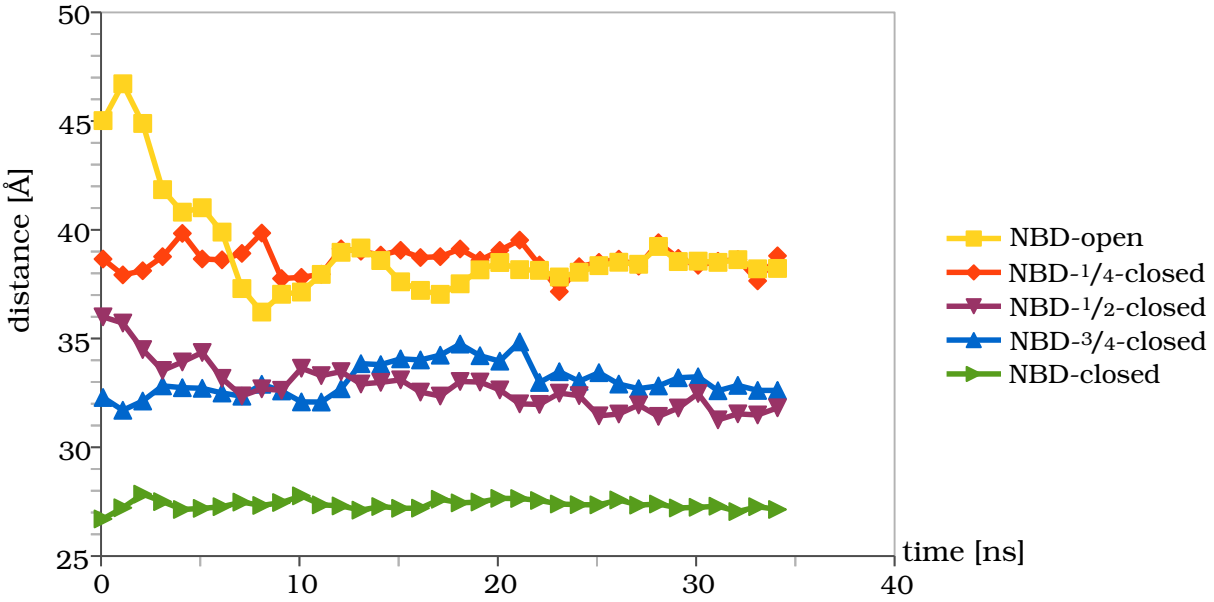


Figure 5.3: COM calculations of the NBD distances of P-gp conformations including colchicine.

Interestingly, colchicine seemed not to have this effect on the other three conformations. Only the NBDs of NBD-1/2-closed conformation slightly approximated each other by ~3 Å, reaching the distance of the NBD-3/4-closed conformation, while latter stayed, as the NBD-closed conformation, consistent throughout the simulation of 35 ns.

These results indicate that, P-gp may have three phases in the NBD closure. Although the sampled conformations had a comparable distance to each other (RMSD in section 3.4), they were not completely even at the "starting" distances of the NBDs (see table 5.1).

But even so, colchicine influenced the NBD-open and NBD-1/4-closed conformation in a way, that both reached the same distance of 38 Å. These two conformations showed the highest separation of

Conformation	COM distance NBDs [Å]
NBD-closed	45
$^3/_4$-closed	39
$^1/_2$-closed	36
$^1/_4$-closed	31
NBD-open	26

Table 5.1: center of mass calculation of the NBD distances before the molecular dynamics simulation of the colchicine docked conformations.

6 Å at the starting distances of the NBDs, while the NBD-$^1/_2$-closed conformation, where the NBD starting distance was only 3 Å apart to the NBD-$^1/_4$-closed conformation, preferred to change its conformation by approximating the NBDs even more to each other, rather than to separate them, concluding that colchicine promotes the closure of the NBDs.

A separation of the NBDs may only be achievable after ATP hydrolysis, which could not be seen in the simulations, since no ATP was docked into P-gp [108].

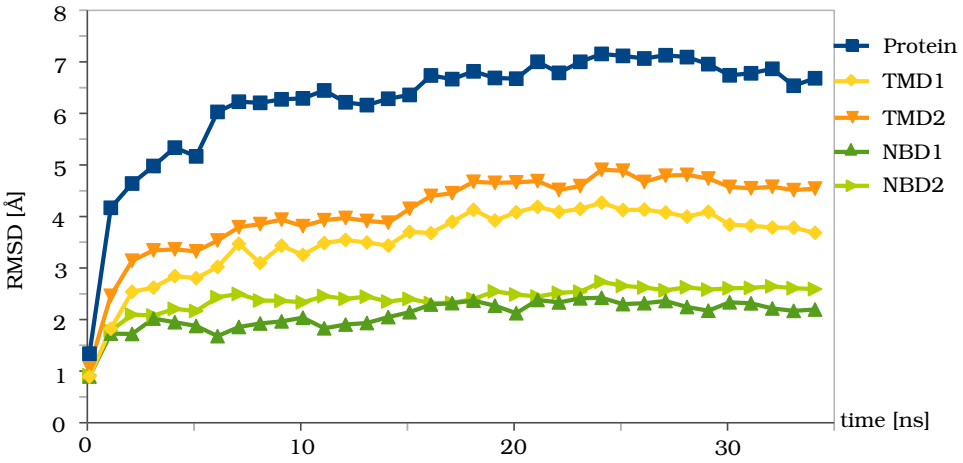
5.2 Origin of structure flexibility

5.2.1 Unwinding of helices

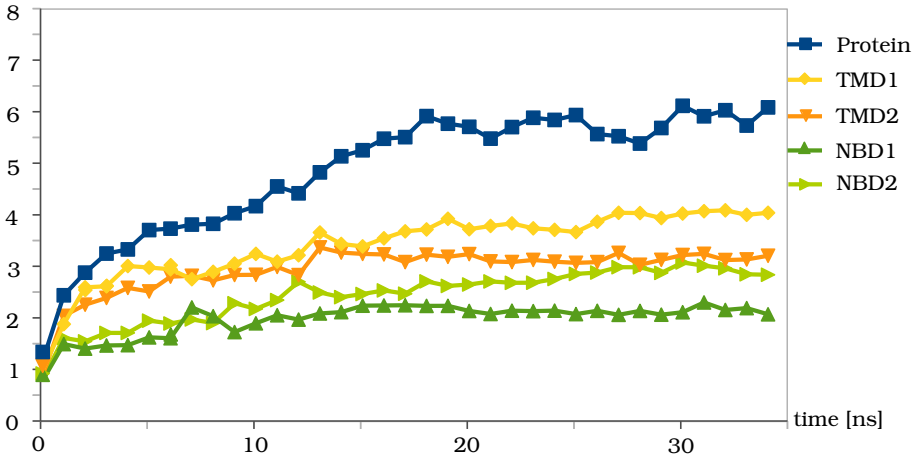
Although the NBDs appeared to exhibit greater movement (clearly visible in the NBD-open and NBD-1/4-closed conformation, see section 5.1), they behaved like a rigid body with an internal RMSD even lower than the TMDs (figure 5.4).

As an example, the RMSD of of NBD-open with docked colchicine, is plotted next to the NBD-1/2-closed conformation (figure 5.4). Although the NBD-open conformations showed a drastic decrease of the NBD distance (figure 5.3) the RMSD of both NBDs is smaller than of the TMDs, and is similar to the RMSD values of the NBD-1/2-closed conformation, which stayed throughout the simulations constant.

This higher RMSD of TMD1 and TMD2 could be caused by hinging motions during the dynamics. Therefore the morphing structures were additionally analyzed for helical kinks or other structural evidence of possible hinge regions, which may be a loss of structure and should ideally be located at, or close to high RMSF sections.



(a) NBD-open including colchicine.



(b) NBD-1/2-closed.

Figure 5.4: RMSD of the whole Protein, and the separate domains.

All morphing structures showed after 35 ns simulation time, structural defects in following positions:

Domain	Area	Residues
TMD1	TM3 (lower POPC leaflet)	*Gly185 - Met192*
	TM6 (lower POPC leaflet)	*Phe343 - Ser351*
TMD2	helix 9 (cytoplasm)	*Asp821 - Ile829*
	helix 10 (transmembrane to cytoplasm)	*Ser880 - Gly894*
	helix 11 (cytoplasm)	*Leu924 - Arg929*
	TM11 (lower POPC leaflet)	*His936 - Ile940*
	TM12 (lower POPC leaflet)	*Ala987 - Pro996*

Table 5.2: Regions of helical unwinding in all conformational P-gp structures.

To clarify, when exactly the α-helix changed during the simulations, a DSSP (hydrogen bond estimation algorithm) analysis was performed.

The DSSP analysis differentiates, between helical *turning*, *bending*, *coil* and a π*-helix*. Latter produces not only significant bends and local distortions in the α-helix segments, but also leads to an appropriate orientation of the constituent residues, as to help proper protein folding [109].

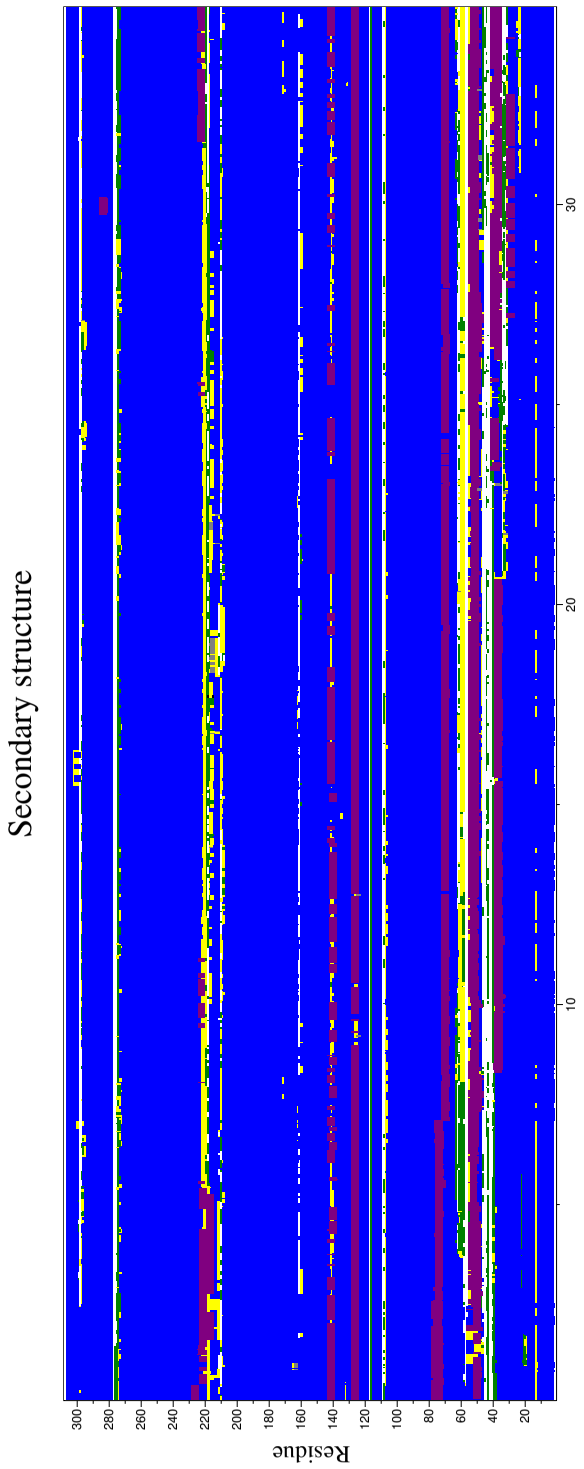


Figure 5.5: DSSP of the TMD1 of NBD-3/4-closed conformation of P-gp. The *coil, turning helix* of helix 3 (residues 130-175) and helix 6 (residues 292-305), appeared within the first nano seconds and remained until the end of simulation. *The residue numbers of the DSSP plot, are not exactly the P-gp residue numbers, since DSSP analysis numbered from the first residue of the generated conformations, which starts from residue Val36.*

Additionally to the unwinding helices in table 5.2, DSSP analysis revealed following structural types of secondary structure:

Domain		Residue	Structure
TMD1	TM2	*Gly102 - Met111*	π-helix
	helix 3*	*Leu175 - Ser179*	π-helix
	TM3	***Gly185 - Met192***	coil/turning helix
	TM5	*Gln270 - Lys272*	turning helix
	TM5	*Glu273 - Glu275*	bending helix
	TM6	***Ser349 - Ile352***	coil
	helix 6*	*Phe366 - Asp370*	coil
TMD2	TM8	*Gln750 - Leu757*	coil/turning helix
	helix 8*	*Thr785 - Tyr790*	π-helix
	helix 9*	*Thr810 - Thr815*	turning helix
	helix 9*	***Thr816 - Asp822***	bending helix
	TM9	***Ala828 - Arg832***/*Gln838 - Asn842*	π-helix
	TM10	*Val874 - Met876*	coil
	helix 10*	***Glu889 - Ala897***	π-helix
	helix 11*	***Tyr920 - Val926***	coil/turning helix
	TM11	*Lys934 - His936*	coil
	helix 12*	*Ala1005 - Ile1012*	turning helix

*intracellular parts of helices

Table 5.3: Different helical structures of P-gp conformations, identified by DSSP analysis. The **bold residues** were part of the unwinding helices at the end of the simulations.

If the helical unwinding/bending was not present from the beginning of the simulation, it occurred within the first 3 ns (figure 5.5) and lasted throughout the whole simulation time. This first nano seconds may still be seen as equilibration time, since the protein has no restrains and adapts freely to its environment. And the ligand colchicine readjusts itself within the drug-binding site (see section 6.2).

Figure 5.6: Unwinding of helix 9. Starting helix **blue**, after 35 ns NBD-1/4-closed (**purple**), NBD-1/4-closed including colchicine (**red**), NBD-1/2-closed (**cyan**), NBD-1/2 (**dark cyan**) and NBD-3/4-closed (**green**). *Although all conformations showed helical unwinding, due to clarity only a sample is shown.*

The locations of the helical deformations at the end of the simulations mostly corresponded to the unwinding/bending helix sections (table 5.3). TM2, TM5, TM8 and TM11 appeared additionally in the DSSP analysis, which plotted the helical structure more specific, and therefore helical *bendings* and π-*helices* could be identified. If the helical deformation sections were not identical to the DSSP analysis, they were located next to the sections. For example, all morphing conformations showed at the end of the simulation a deformation in TM9 from *Asp821-Ile829*(figure 5.6), which overlaps with two of three divided sections by DSSP: *helical turning*: *Thr811-Thr815*, *helical bending*: *Thr816-Ala822* and π-*helix*: *Gly826-Arg832* (figure 5.5).

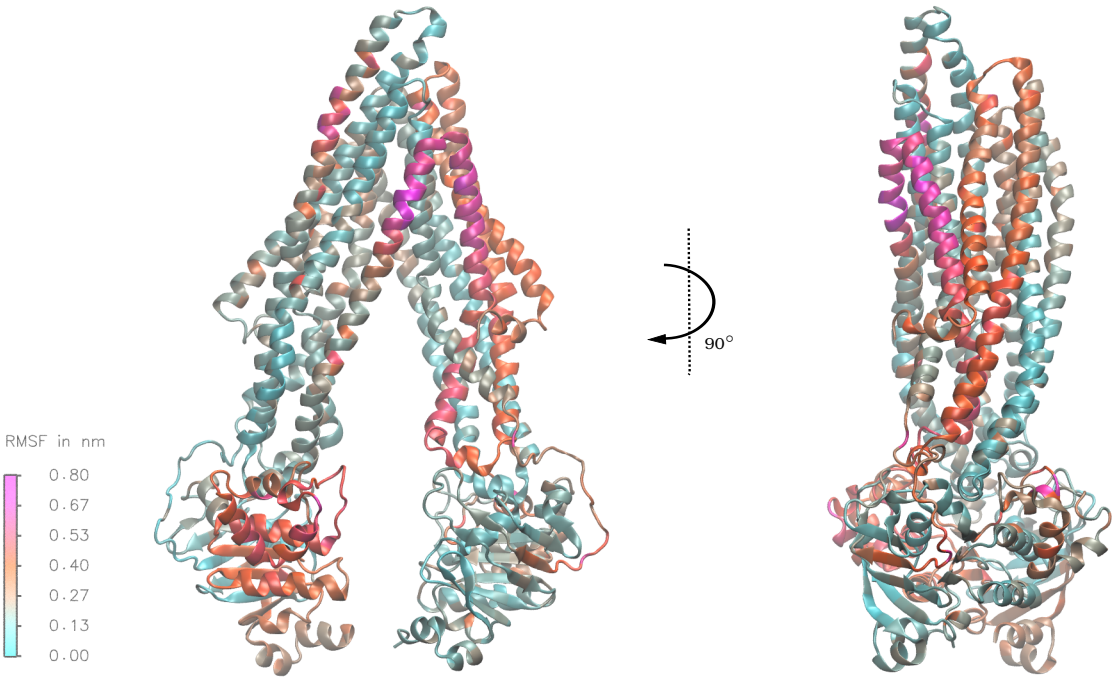


Figure 5.7: RMSF of NBD-closed conformation. Helices 7, 8, 9 and 10 showed next to NBD1, the highest fluctuations.

Visualizing the positions of this helical deformations, which correlated to the high RMSF regions of P-gp (see figure 5.7), showed, that despite the different sequences between TMD1 and TMD2, the regions were located at comparable heights, in respect to the POPC membrane (see figure 5.8).

The different P-gp conformations showed similar RMSF values. The differences correlated with the actual movement of the NBDs. High NBD fluctuations showed high RMSF values in helices 3 (TMD1), 8, 9 and 10 (TMD2) in comparison the the rest of the protein. Low NBD fluctuations had very low RMSF values.

There were three regions were the helices showed possible hinges:

1. **Cytoplasm level:** The identified helices close to ICL1 and ICL4, that are located in the inner cavity of P-gp (H3 and H10), were both π-*helices*, while H6 and H11 were identified as *coils* and *turning helices*. The helices leading to ICL2 and ICL3, had all different structure appearances, on the inner side the counterpart of H3, which is H9, was identified as a *bending helix* and the counterpart of H6, which is H12, was identified as *turning helix*. Only the π-*helix* of H8 is located at the "outside" of P-gp.

2. **Transmembrane, inner POPC leaflet, level:** The helices mentioned in the cytoplasm level, showed all but two (H8 and H12), also a *turning helix* section in the POPC membrane inner leaflet. TM9, being the only helix identified on the NBD2 side of P-gp, had also a π-*helix* part.

3. **Transmembrane, outer POPC leaflet, level:** the π-*helix* section of H2 and the *bending helix* part of H8 were located next to ICL1 and ICL4, respectively.

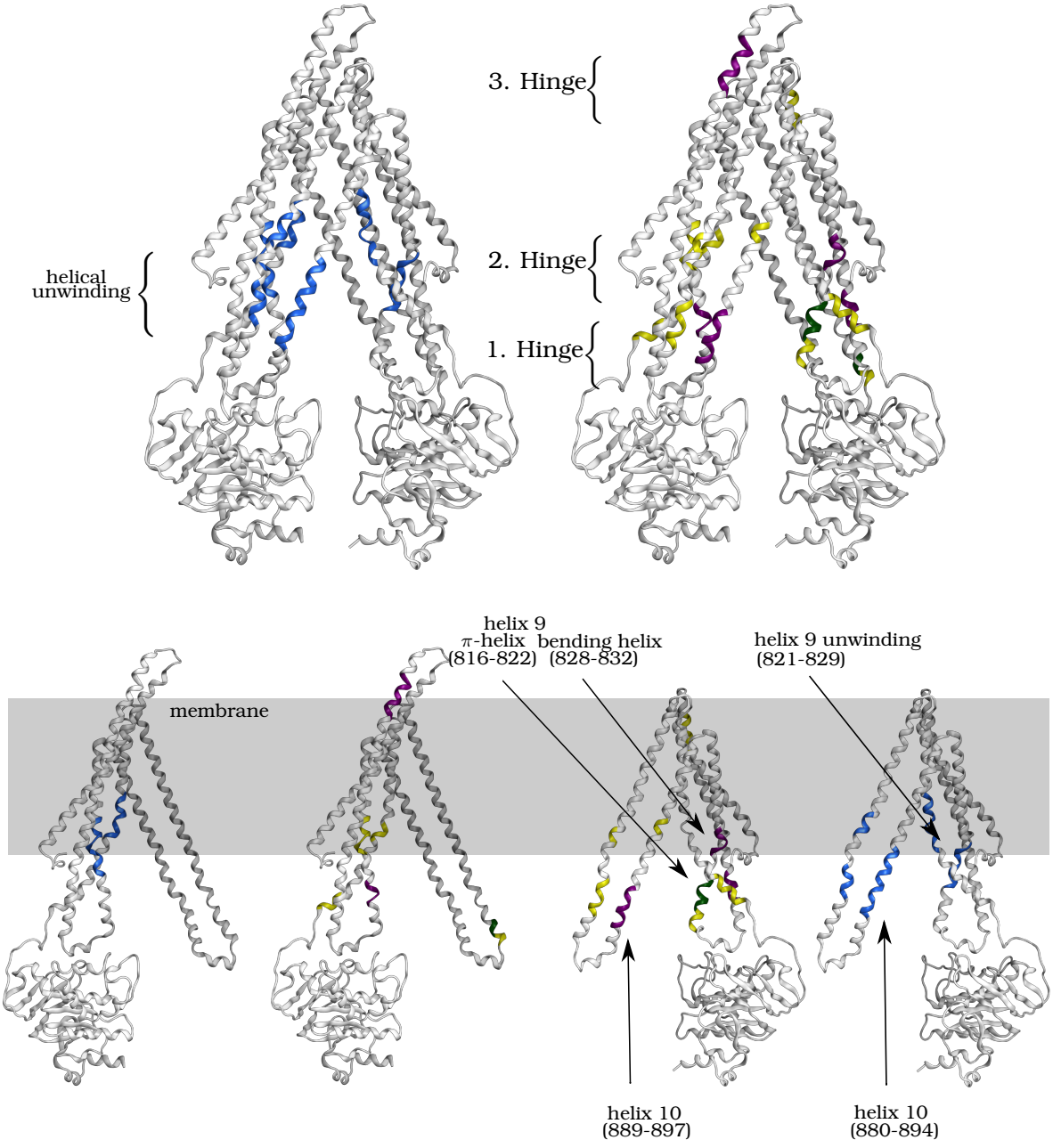


Figure 5.8: Suggested P-gp hinges. Marked areas of helical unwinding **blue** during the molecular dynamics; helical turning (**yellow**) and bending (**green**), and π-helices (**purple**) identified by DSSP analysis.

The helix 3 in TMD1, kinked/unwound at two points, corresponding approximately to the midpoint of the membrane, and close to the ICL1, and so did its counterpart in TMD2, namely, helix 9. Additionally in TMD2 helices 10 and 11 showed hinges at the same level as helices 3 and 9. Helix 12 showed an kink/unwinding only in the extracellular end, and similar unwinding was also present for the counter part region in TMD1, namely extracellular end of helix 6.

This two hinge levels suggest that the conformational changes take place on two sides. Firstly the extracellular hinge region, located mainly in the TMD2, which may induce the twisting motion of the NBDs. And secondly the closure of the intracellular facing cavity.

Noteworthy, the unwinding section of helix 6 located in the inner leaflet of the membrane, may also induce the opening of the TM4/TM6 gate to the membrane (see section 5.4) and therefore affect its neighbor helix 3.

Additionally the unwound section of helix 2 and its counterpart helix 8 (in addition to the above mentioned helices and their counterparts) suggest that there is a symmetry between the two TMDs, which is not apparent from their sequence, but their topological similarity.

Since all conformations, regardless if in complex with colchicine or not, were involved in the observations, three putative hinge levels could be localized, but a trigger for the NBD-closure invoked by colchicine could not be determined.

5.2.1.1 Comparison with rigid cluster results of Normal mode approach by NMSim Server

The helical unwinding/bending sections were compared with rigid cluster regions, or rather with the flexible parts of P-gp of the rigid cluster decomposition results of section 3.5.1.1. Only the rigid cluster decomposition of the second RCNMA run showed comparable results. A flexible region of TM3 listed in section 3.5.1.1 *Ser180-Glu184* lies in between the *bending* and π*-helix* of TM3 (see table 5.3). TMD2 showed more hinge region similarities, the *turning-* and *bending helix* section of H9, as well as a the *truning-* and a very small part of the π-helix of H10 were recognized as flexible regions. However, these similarities related only to the rigid cluster decomposition of the NBD-open conformation, while the NBD-closed conformation had no matches.

If P-gp has hinge regions on the three mentioned levels, this comparison of the flexible regions recognized by the rigid cluster decomposition of NMSim Server [87] may indicate, why the process was not successful. By dividing P-gp into to big rigid regions and identifying the flexible joints on the wrong helices and levels, the protein was not flexible enough for conformational changes.

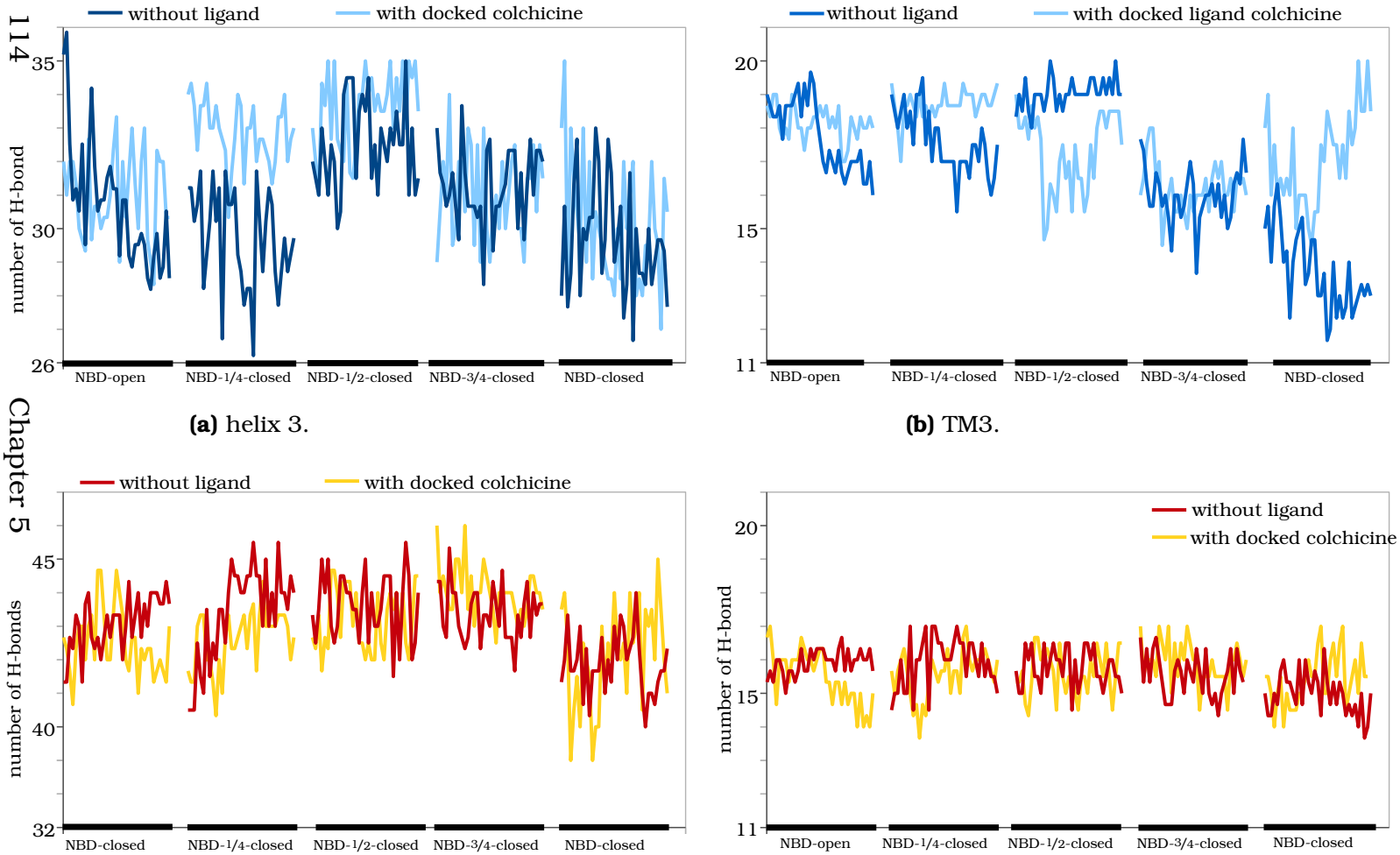
Noteworthy, the transmembrane hinge regions correlated very well with the rigid cluster mentioned in section 3.5.1.2, showing that the hinges were present but undetected by the rigid cluster decomposition.

5.2.2 H-bonds of helices 3 and 4 as example for indication of flexibility of P-gp

The helical unwinding/bending should also be visible by a reduction of the hydrogen bonds between the carbonyl and amine group of the helix backbone of residues i and $i + 4$, which stabilize the structure. Therefore the H-bonds of helices 3, 4, 8, and 9 as well as their TM parts were analyzed.

The analyzed H-bonds of the transmembrane sections of the helices had less H-bond alterations, than the whole helices, suggesting the helices being tighter packed and less flexible/bendable than the sections located in the cytoplasm.

The helices, which did not show helical unwinding at the end of each simulation and which are suggested not to have hinge regions, namely H1 and H4 of TMD1 and H7 of TMD2, had low H-bond alterations in comparison to the other helices, indicating that the helices indeed kept their tight packing and therefore may move only as a whole helix (figure 5.9).



(a) helix 3.

(b) TM3.

(c) helix 4.

(d) TM4.

Figure 5.9: Number of H-bonds of helix 3 of Apo P-gp (**blue**) and colchicine docked P-gp (**light blue**), as well as helix 4 (**red**) of Apo- and colchicine docked P-gp (**yellow**), respectively.

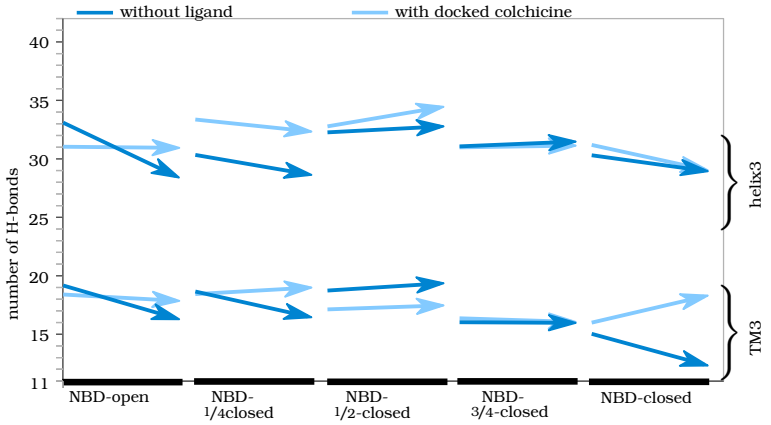


Figure 5.10: The trend of the H-bond number of helix 3 and its transmembrane part, respectively is shown of the ligand free conformations **blue** and the conformations with included colchicine **light blue**.

As shown in section 5.2.1, H3 and its counterpart H9 showed two possible hinges at two regions of P-gp. The number of H-bonds decreased during the simulations (fluctuating by 8-10 H-bonds), validating the suggestion of a hinging motion by reduced H-bonds. Interestingly comparing the trends of H-bond numbers of the whole helix and their transmembrane part in each simulation, showed that the reduced number of H-bonds in the whole helix, occurred mostly by the reduced H-bonds of the transmembrane region (figure 5.10 and 5.11), while the conformations with colchicine behaved differently (e.g. figure 5.10 and 5.11; NBD-1/4-closed and NBD-closed conformations).

Both, the whole H3, as well as the transmembrane section showed H-bond alterations. Interestingly, the NBD-1/2-closed conformation, showed a stable helix backbone H-bonding in its transmembrane part, although being in the mid transition state of P-gp, suggesting, that it should show the highest fluctuations.

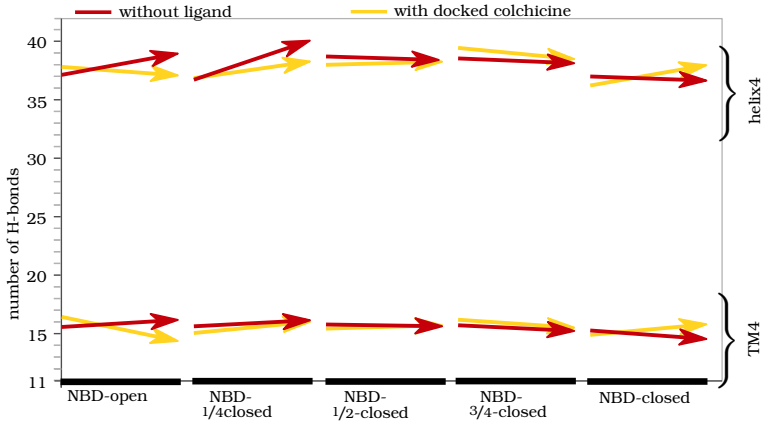


Figure 5.11: The trend of the H-bond number of the conducted simulations of ligand-free P-gp conformations **red** and the conformations with docked conformations **yellow**.

Helix 4, being one of the helices, which were suggested to have no hinge motions, showed an almost constant number of H-bonds (with a fluctuation of 3-5 H-bonds), indifferent if there was colchicine present or not.

This result correlates with the suggestion of the number of H-bonds between residues i and i+4 of the helices, to indicate flexible regions. Due to analyzing whole or transmembrane part of the helix, the exact point of the flexibility of the helix cannot be captured.

5.3 Rearrangements of helices during the simulation

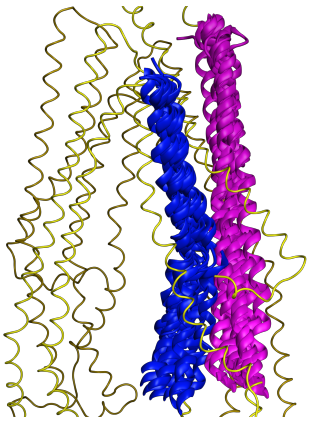
The distances between adjacent helices facing each other and the inner cavity, were monitored during the course of simulation, to investigate their movement.

The residues for the measurement, were picked as top residues, which are the residues towards the extracellular space and mid residues, which are located close the edge of the inner membrane leaflet to the intra-cellular space.

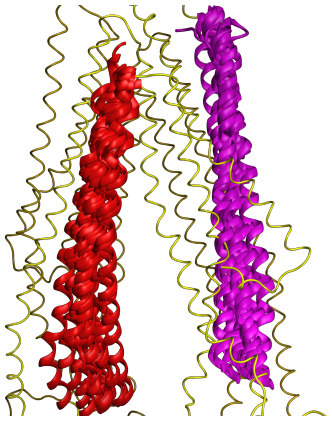
Residues at the bottom of the helices were also picked, which were located close to the ICLs and NBDs, respectively. As in section 5.5 the C_α-atom distances of these residues were measured.

Additionally the distance of the top and bottom residues was measured to investigate the length of helices. Also the selected residues were not in the unwinding helix areas of section 5.2.1, except for helix 11.

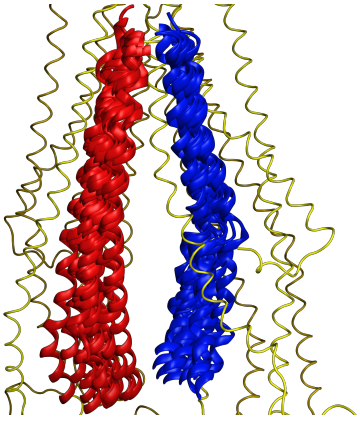
Due to calculation time, only helices 2, 3, 4 and 11 were picked.



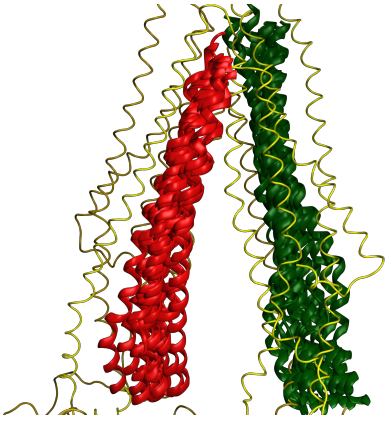
(a) Rearrangement of helix 2 (*purple*)to helix 3 (*blue*).



(b) Rearrangement of helix 2 (*purple*) to helix 4 (*red*).



(c) Rearrangement of helix 3 (*blue*) to helix 4 (*red*).



(d) Rearrangement of helix 4 (*red*) to helix 11 (*green*).

Figure 5.12: Location and movement of the investigated helices. The helices of the NBD-open homology model are shown in *yellow*. To visualize the movement, the investigated helix of all structures is plotted at the end of the simulation.

5.3.1 Helix 2 and Helix 3

Helix 2 and helix 3 are not only located next to each other, but are also connected to ICL1, which forms non-bonded contacts with NBD2. H2 goes straight toward the ICL1, while H3 shows a small kink, located close to the mid residue *Met192* (figure 5.12). This kink is next to the putative second hinge level of section 5.2.1.

Two different motions of the top residues of helices 2 and 3 could be seen (figure 5.13):

Met111 and *Arg210* decreased their distances during each simulation, only the NBD-open and NBD-1/4-closed conformation fluctuated around their initial morphing distances.

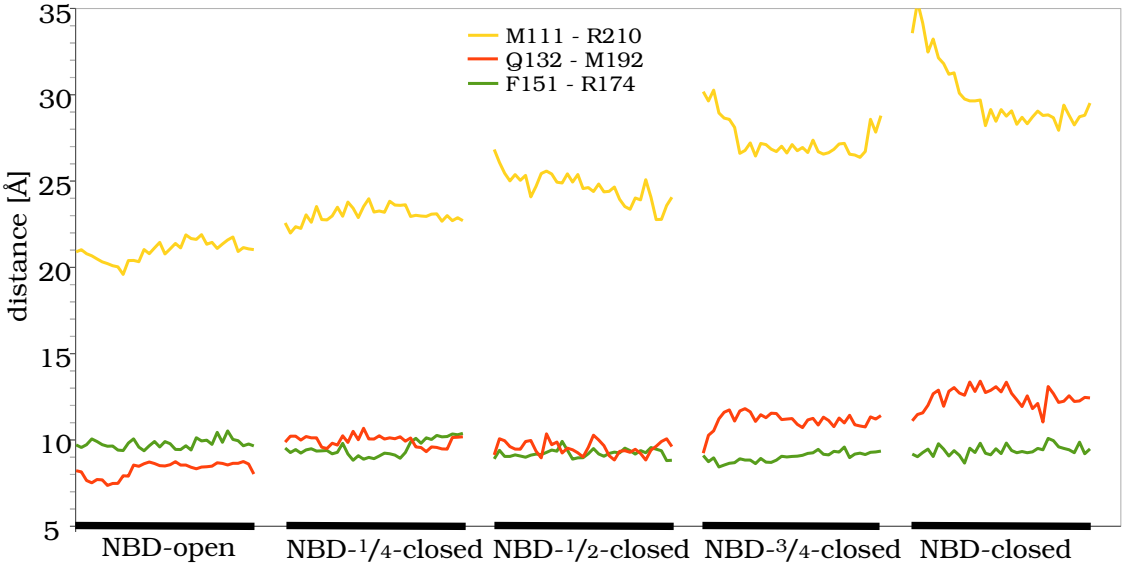
The structures from NBD-1/2-closed towards NBD-closed, in which colchicine was docked in, reduced their values by ∼3 Å, reaching a distance of 29.5 Å. This distance was the initial distance of the ligand-free NBD-closed conformation, which showed at the end of the simulation an constant distance of 24 Å, and showed with the other colchicine-free structures a reduction of the distance by ∼5 Å (figure 5.13a and 5.13b).

On the other hand, both helices moved apart from ∼21 Å to 29 Å (clearly visible by the colchicine including structures; figure 5.13b).

The mid residues *Gln132* and *Met192* fluctuated during each simulation around their initial value, being 8 Å for the NBD-open and NBD-1/4-closed conformation and increasing the distance from 10 Å for NBD-1/2-closed and NBD-3/4-closed to ∼12 Å for the NBD-closed including colchicine conformation (figure 5.13a and 5.13b).

The bottom residues *Phe151* and *Arg174* stayed at a constant distance of ∼9 Å to each other, indicating, that helix 2 moved from the mid residue to the ICL1 along with helix 3 in the same direction, while the upper part of these helices moved apart. By this movement P-gp may be capable to dislocate ECL1 and open the outer P-gp pore.

(a) Conformations without colchicine.



(b) Conformations with docked colchicine.

Figure 5.13: Rearrangement of helix 3 and helix 3.

5.3.2 Helix 3 and Helix 4

The two helices 3 and 4, which are as an inverted "V" formed, are not only adjacent, but do also face each other, intersected by TM6, which forms a physical linkage to NBD1 (see figure 5.12). The TMs are completely straight between the top and mid residues, from which there is a light bending in the outer-membrane/cytoplasm helix part located at the mid residues. From this point the bottom residues are straight forward located close to ICL1 and ICL2, respectively.

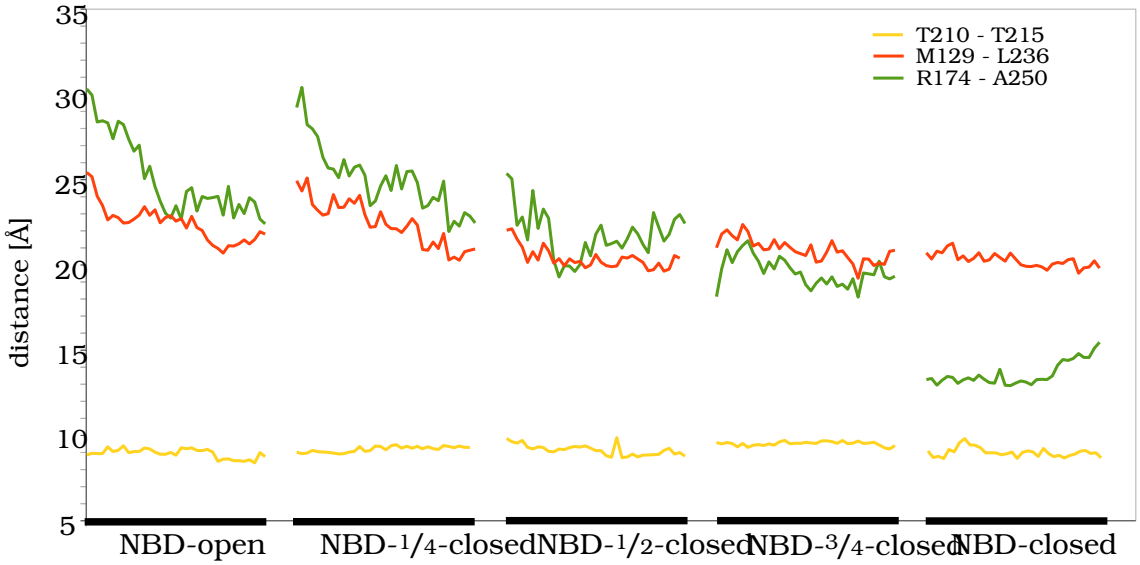
All conformations showed, that the top residues stayed at a constant distance of ∼9 Å.

The bottom residues of NBD-open and NBD-$^{1}/_{4}$-closed conformation decreased their distance drastically by almost 7 Å, while the mid residues reduced it only by a maximum of 4 Å, nearing each other to 20-21 Å. This distance stayed throughout the following conformations NBD-$^{1}/_{2}$-closed, NBD-$^{3}/_{4}$-closed and NBD-closed. Also the conformations, in which colchicine was docked in, showed a value of ∼20-21 Å, but with a steeper gradient, mainly by the NBD-$^{1}/_{2}$-closed and NBD-$^{3}/_{4}$-closed structures.

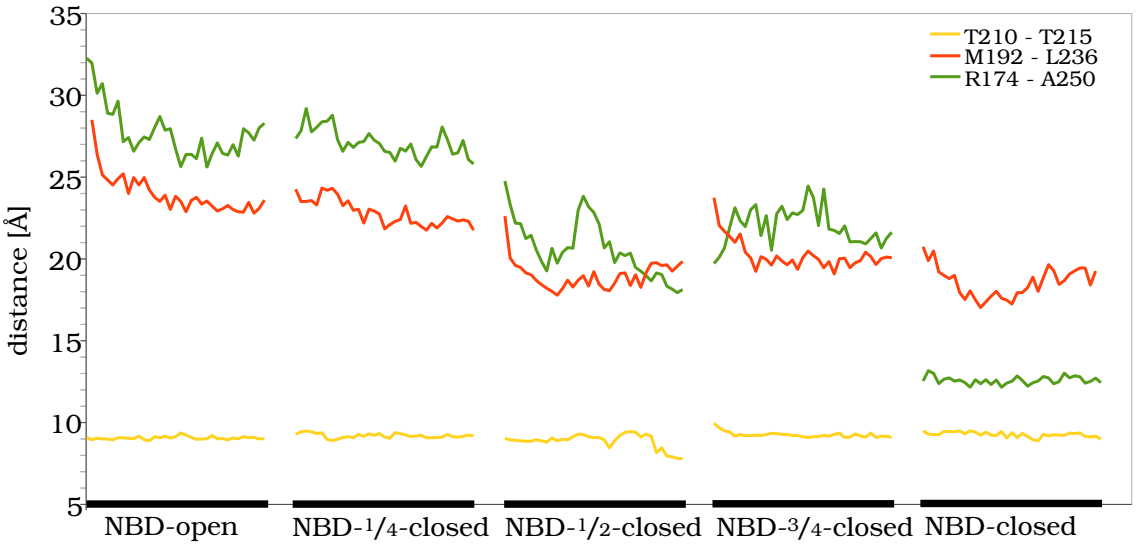
Arg174 and *Ala250* did not come as close together (∼27 Å) as with the conformations including colchicine, especially in NBD-open and NBD-$^{1}/_{4}$-closed structures, indicating, that although the colchicine binding-site was not close to the bottom residues, its presence influenced their movement towards each other.

The length of H3 and H4 showed in all P-gp conformations small fluctuations. Changes in stretching or compressing occurred only

in ranges of maximum 1-4 Å, concluding that the helices needed to stretch and contract before reaching their optimal size and conformation.



(a) Conformations without ligand.



(b) Conformations with docked colchicine.

Figure 5.14: Rearrangement of helix 3 and helix 4.

5.3.3 Helix 2 and Helix 4

Being directly connected to helix 3, the distances of helix 2 and 4 were analyzed to investigate the movement of the helix2-ICL1-helix3 complex. If the closure of the NBDs results in a straight movement of helix 2 and 3 towards helix 4, then the distances should be similar to H3 and H4 results in the above section 5.3.2.

The top residues of H2 and H4, *Gln132* and *Leu236*, increased their distance steadily from 26.5 Å to ∼30 Å. Conformations including colchicine in the drug-binding site showed an even higher distance of ∼35 Å. Only the NBD-$^{3}/_{4}$-closed and NBD-closed conformations showed fluctuation until they adapted to the resulting value, while all the other structures had an almost constant distance.

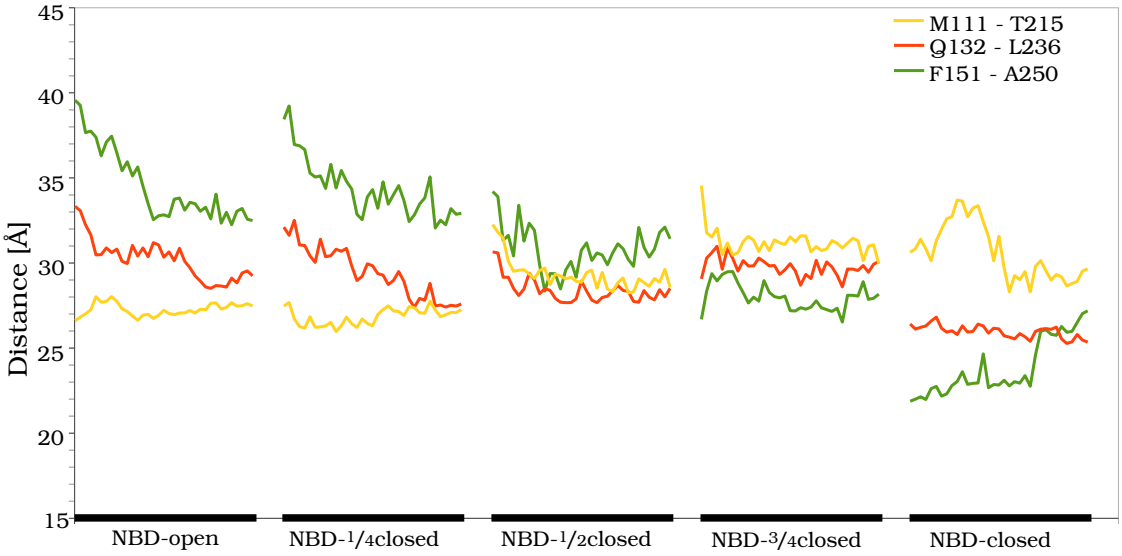
In contrast to the top and mid residues, *Phe151* and *Ala250* decreased their distance to each other not only during the simulation time, starting at ∼40 Å and reaching a distance of ∼28 Å, but also from one conformation to another.

Noticeably, *Gln132* and *Leu236* of the NBD-open and NBD-$^{1}/_{4}$-closed structure approached each other similar as the bottom residues, indicating that the helix part outside of the membrane (inside the cytoplasm) of H2 moved along with H4 in same direction, while the transmembrane parts came closer together.

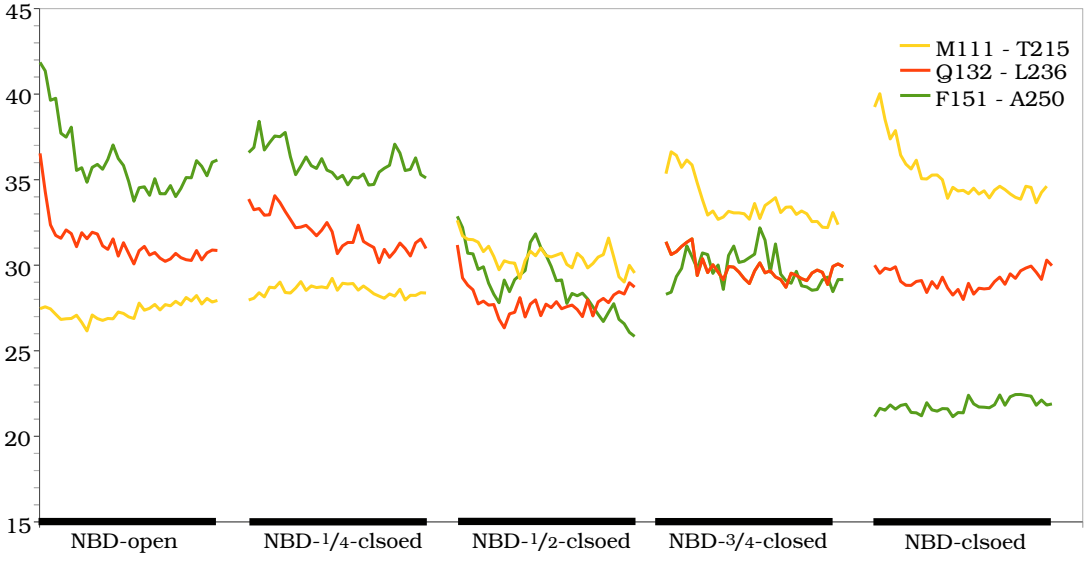
From the NBD-$^{1}/_{2}$-closed conformation on, these two residues stayed at a distance of ∼27 Å, while *Phe151* and *Ala250* continued to slowly decrease their distance reaching a value of 25 Å, suggesting that the helices 2 and 4 begin their movement approaching each other and from mid-point of the conformational change of P-

gp (NBD-1/2-closed), continued a twisting motion, corresponding to a angular rotation of ICL1 and a twisting of the NBDs.

Unlike the H3 and H4 movement (see section 5.3.2), colchicine seemed to have no effect in this motion of the two helices, since all conformations showed a similar pattern (figure 5.15a and 5.15b). This could indicate, that the movement of helix 2 is directly coupled to the movement of helix 3, while the movement of helix 3 is influenced by a ligand.



(a) Conformations without colchicine.



(b) Conformations with docked colchicine.

Figure 5.15: Rearrangement of helix 2 and helix 4.

5.3.4 Helix 4 and Helix 11

Helix 11 is located at the membrane faced side of P-gp, adjacent to helix 2, and is linked to ICL4. Due of its distance to helix 4 it does not have any interactions, but was investigated because, their movement, may be of interest, to clarify the movement of the two TMDs to each other. Unlike helix 2, it is not straight, but shows a kink, located close to the mid residue *Arg938*.

The top residues of both helices *Thr215* and *Arg958* fluctuated at an average distance of ∼30-31 Å. Ligand-free NBD-3/4-closed structure showed a higher value of 33.5 Å, but reached after 35 ns also 30 Å.

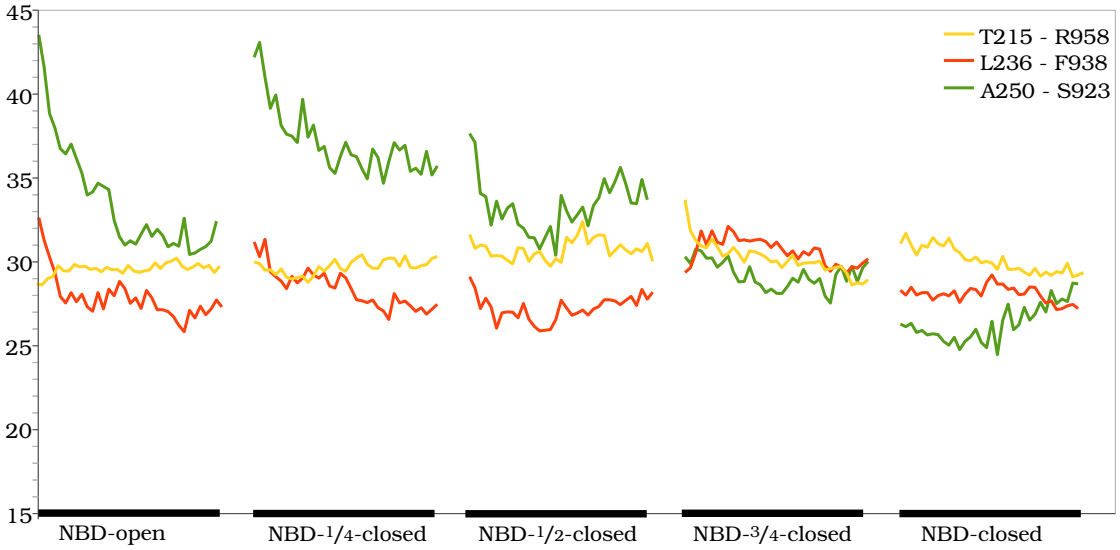
A slightly increasing distance by 3 Å was observed with the colchicine bound structures. Here too, the NBD-3/4-closed conformation had the highest distance.

Leu236 and *Phe938* fluctuated around a distance of ∼27-30 Å, having the smaller distances close to the NBD-open state and increasing it coming towards the NBD-closed conformation of P-gp. The, in comparison to all other mid residues distances, very high starting value of 32.5 Å of both NBD-open structures (ligand-free and ligand-bound) decreased within the very beginning of the molecular dynamic simulations (see 5.16a and 5.16b).

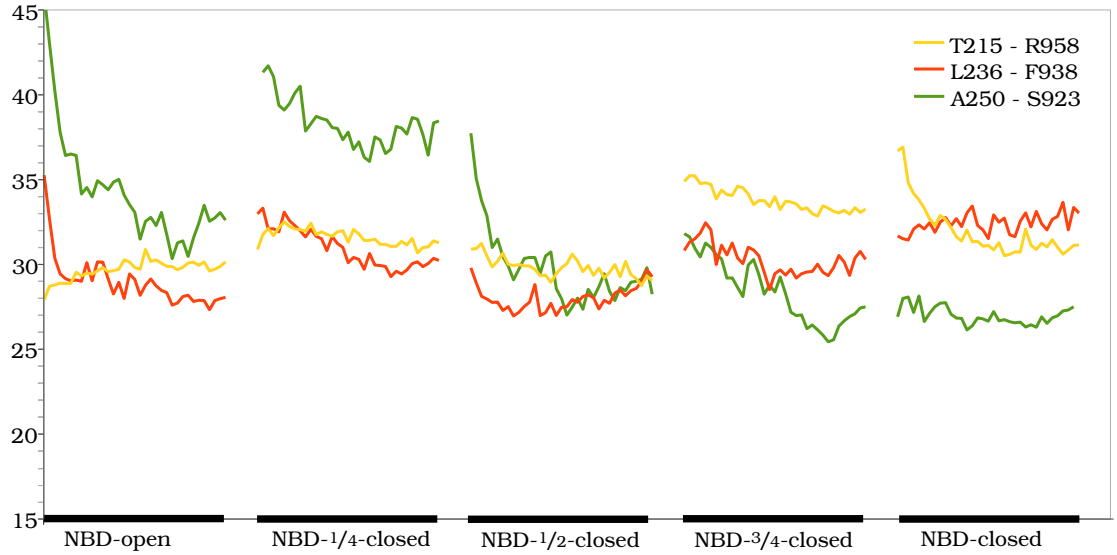
The highest changes showed the bottom residues *Ala250* and *Ser293*. Firstly the NBD-open conformation drastically reduced the distance by 11 Å. So did also the NBD-1/2-closed including colchicine conformation. While the other structures did not reduced their distances as much.

Interestingly colchicine seemed to influence the NBD-1/2-closed

conformation, since the ligand-free structures reduced their distances gradually from each structure to an other, reaching a distance of 30 Å by the NBD-closed conformation, while this distance was already reached by the NBD-1/2-closed conformation, which had colchicine in the drug-binding site.



(a) Conformations without colchicine.



(b) Conformations with docked colchicine.

Figure 5.16: Rearrangement of helix 4 and helix 11.

Considering the constant length of both helices throughout the simulations (figure 5.17), the small changes of the distances by 3 Å can be neglected. Therefore, it may be concluded, that while the transmembrane part of helix 4 and helix 11 keep their distance, the outer membrane, cytoplasm part of the two helices, which start close to the kink of helix 11 move, by closing the NBDs, towards each other.

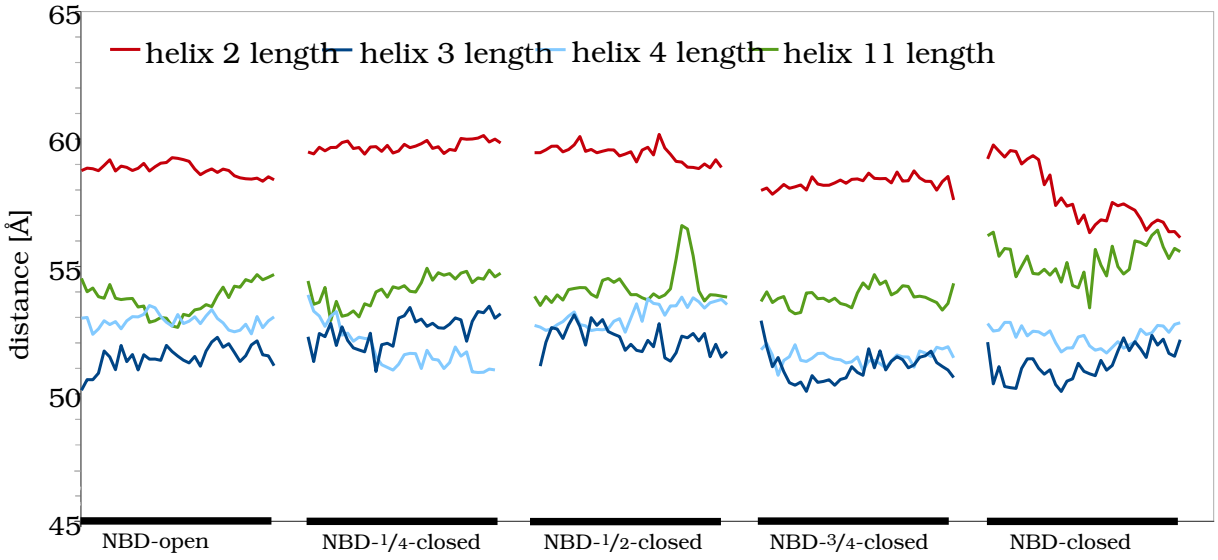


Figure 5.17: Length of helices. Distances were measured from top to bottom residues. The lengths were mostly consistent. Only NBD-1/2-closed conformation had a deviation, which was caused by an unwinding of a large sequence of helix 11.

5.4 Gate of TM4/TM6

The mP-gp crystal structure conformations (PDB ID: 2HYD and 4M1M) exhibit access to intracellular space through the inward-facing inner cavity. More interestingly two lateral portals are formed by TM4/TM6 and TM10/TM12, that are open wide enough to accommodate molecules and expose the drug-binding pocket to the inner leaflet of the bilayer [22]).

In the following section the P-gp conformations were analyzed to investigate, if the TM4/TM6 gate remains through the simulation and, if all generated conformations had this portal.

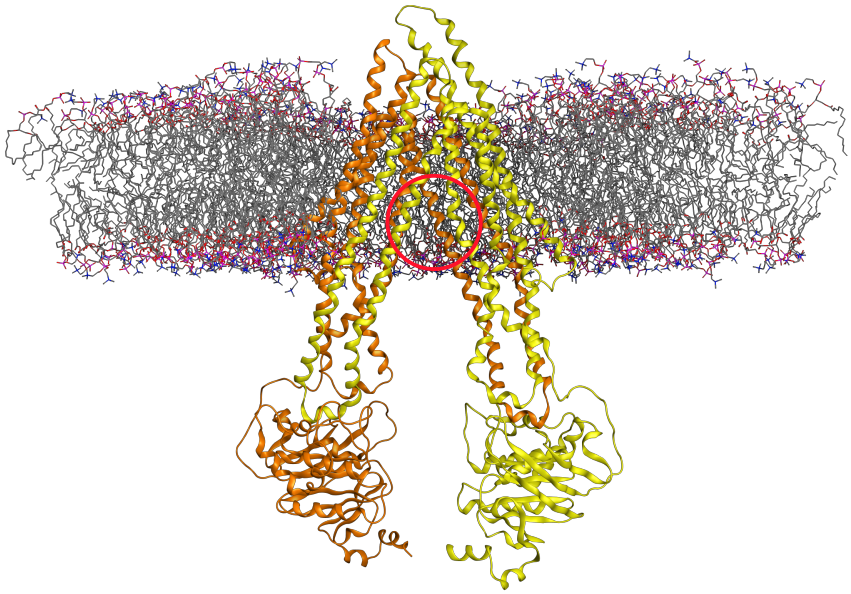


Figure 5.18: NBD-open conformation is embedded into the POPC membrane. The *red* circle shows the location of the gate. TMD1-NBD1 is shown in *yellow* and TMD2-NBD2 is shown in *orange*.

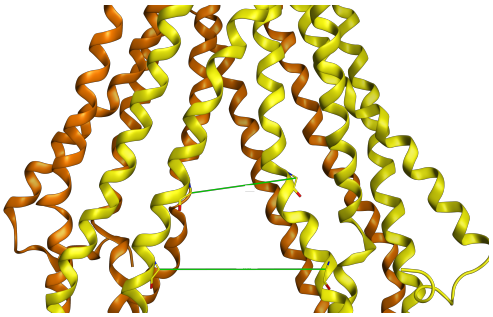


Figure 5.19: Closer look of the gate and location of the two picked residues *Ala230*, *Ser237* (TM4), *Ala348* and *Ala356* (TM6).
TMD1 is shown in *yellow* and TMD2 is shown in *orange*. The *green* lines represent the distance measurement.

5.4.1 Occupancy calculation by VMD volmap plugin

The Volmap plugin of VMD (visual molecular dynamics) [110] creates three dimensional grids based on the atomic coordinates, where each grid point is set to occupied or not-occupied, depending on whether it contains atoms or not. By averaging over all molecular dynamics frames, a fractional occupancy is provided.

(a) Ligand-free NBD-open conformation. **(b)** NBD-open P-gp including colchicine.

Figure 5.20: Grid points calculated by Volmap of VMD [110]. *Red* grid shows the occupancy of 0.6-0.8 and *yellow* is the occupancy of 0.4-0-6.

This calculated grids represent the structure of the protein. As shown in figure 5.20 the portal of TM4/TM6 is visible. The NBD-open conformation with colchicine inside the drug-binding site showed higher fluctuations and a larger gate, than apo P-gp. This higher fluctuations are visible by the color of the grid, where *yellow* shows an occupancy of 0.4 to 0.6 and *red* shows an occupancy during the simulation of 0.6 to 0.8, namely TM4 and TM6 were during most time of the simulation within the *red* region.

One problem of this method was, that the size of the portal was not computable. By calculating the number of grid points (see fig-

ure 5.21, there were small differences identifiable, which were not significant, and it was not possible to conclude from the number of grid points the size of the portal, or even if a gate was present.

The calculation showed that all but one conformation, where stabilized by colchicine, since they had less grid points than the apo conformations. Only NBD-$^{3}/_{4}$-closed had more grid points

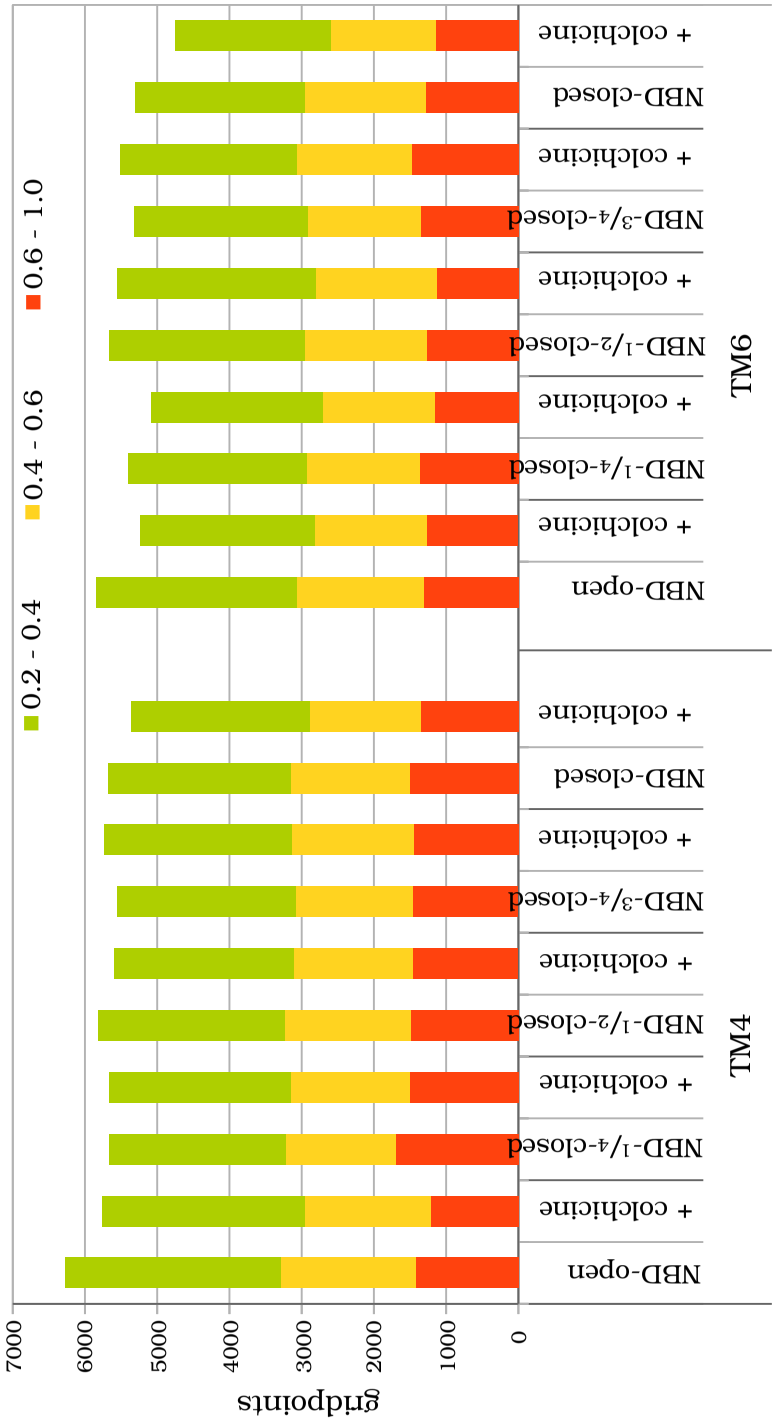


Figure 5.21: Occupancy of backbone atoms of TM4 and TM6.

5.4.2 Distance measurement of TM4 to TM6

To analyze the size of the portal, C_α-atoms of two residues of TM4 and TM6 were picked, respectively. The two residues *Ala230* (TM4) and *Ala348* (TM6) are located in the inner POPC leaflet, and represent the upper part of the entry portal, while the other two residues *Ser237* (TM4) and *Ala356* (TM6) represent the lower part of the portal and are located at the edge of the membrane to the inter cellular space (figure 5.18 and 5.19).

In the crystal structure set PDB-ID: 4Q9I-L a set of ligands induced a large conformational change (11 Å) of TM4 comparing corresponding C_α positions (see figure 5.22). This kinking begins at *Pro219* and ends at *Tyr243* close to NBD2 [111].

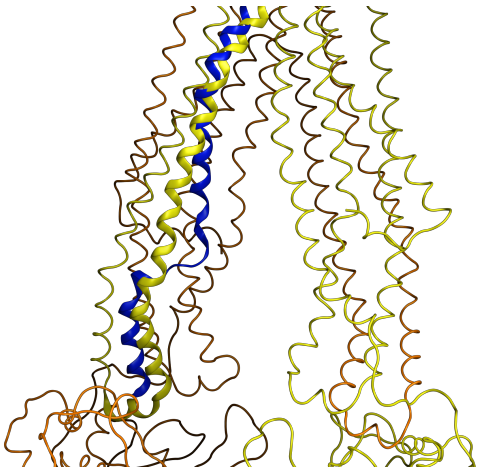


Figure 5.22: TM4 kink observed by *P. Szewczyk et al.* [111].

Therefore, to investigate the straight-helical conformation adopted by TM4 the residues *Ser237* (TM4) and *Ala999* (TM12) were measured additionally.

These two residues had in the NBD-open conformation a distance of 23 Å. The representing residues of the PDB-ID: 4Q9I, which were after an alignment at the same helical level, were *Ser233* and *Ala991* and showed a distance of 16 Å, since the kink of TM4 in this crystal structure is towards TM12.

Comparing the distance measurement results of the NBD-open conformation with the grid point occupancy results of section 5.4.1, the NBD-open conformation showed lower distances of *Ser237* to *Ala356*. This distance decreased from 17 Å to 11.6 Å, while the NBD-open conformation including the ligand colchicine showed only an insignificant small decrease of 2 Å, which resulted in a distance of still 17 Å (see figure 5.23), which is large enough for substrates to enter the inner cavity.

The distance of the residues located in the inner leaflet of the membrane was after 35 ns at 13 Å (apo P-gp) and 12 Å. While colchicine stabilized TM4, and decreased slowly the distance of *Ala230* to *Ala348* from 17 Å to ∼12 Å, ligand-free P-gp showed fluctuations of 4 Å. This fluctuations correspond very well to the higher amount of grid points and thus to the smaller looking TM4/TM6 portal of section 5.4.1.

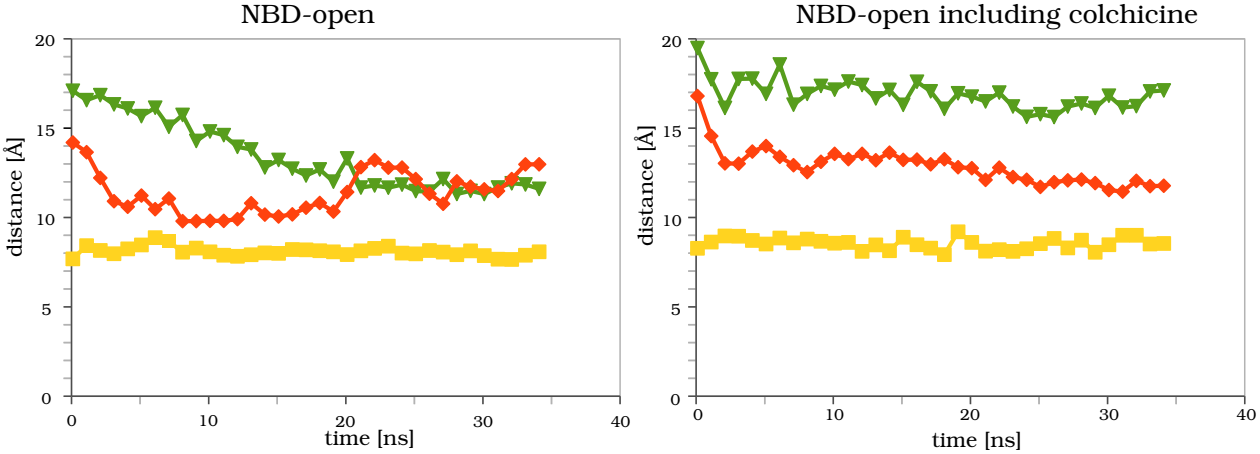


Figure 5.23: Distances of: **yellow**: *Leu219-Val338*, **red**: *Ala230-Ala348* and **green**: *Ser237-Ala356*.
While the ligand-free conformation decreased the distance of *Ser237* to *Ala356*, the colchicine including conformations kept the portal open

The NBD-$^1/_4$-closed conformations showed similar behavior as NBD-open P-gp. A portal existed in the colchicine structure, while the apo P-gp conformation reduced the distances as the NBD-open conformation (see figure 5.24).*Ser237* and *Ala356* had at the beginning of the simulation a distance of 16 Å, which kept the portal of TM4/TM6 open. *Ala230* and *Ala348* of the NBD-$^1/_4$-closed colchicine conformation reached after 35ns a value of ∼11 Å. A distance value close to 10 Å showed also the other morphing conformations: NBD-$^1/_2$-closed (including colchicine) as well both NBD-$^3/_4$-closed conformations.

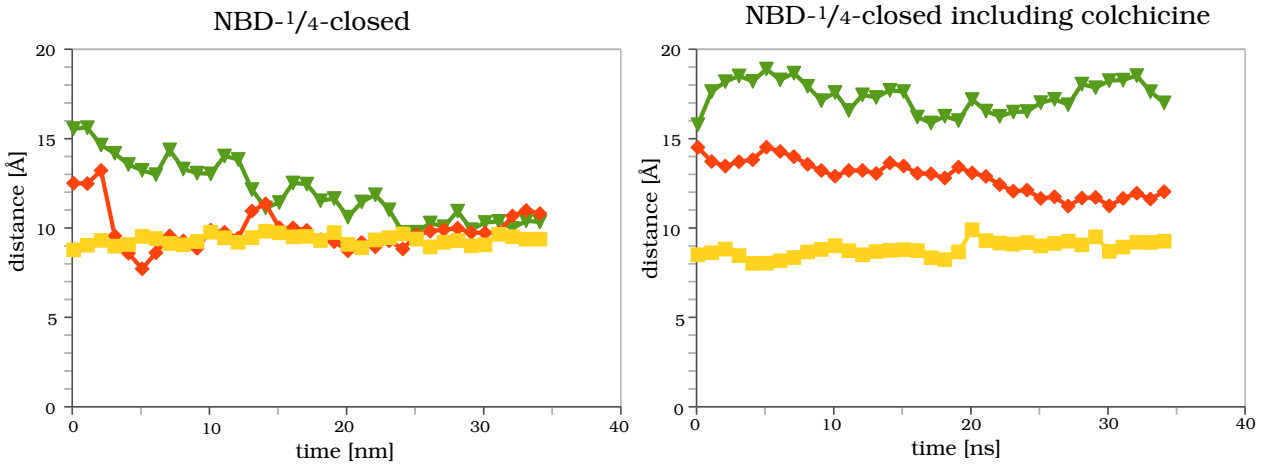


Figure 5.24: Distances of: **yellow**: *Leu219-Val338*, **red**: *Ala230-Ala348* and **green**: *Ser237-Ala356*.

Also in common was, that the three lingand-free generated morphing conformations (NBD-$^1/_4$-closed, NBD-$^1/_2$-closed and NBD-$^3/_4$-closed) decreased their distances in a similar way, reaching for all measured residues similar distances (see figure 5.25).

This distances of $\sim$10-12 Å are in contrary of the observed distances of the NBD-closed conformations. From the values of 14 Å of *Ala230* to *Ala348* a gate may be imaginable, but the very low distance of *Ser237* to *Ala356* show the opposite. A portal is unlikely, since the twisted motion of of TM4 and especially TM6, which is located within the drug-binding site, and blocking the portal to the drug-binding site.

Noteworthy the distance of *Leu219* to *Val338* stayed in all simulations constant at $\sim$8-10 Å.

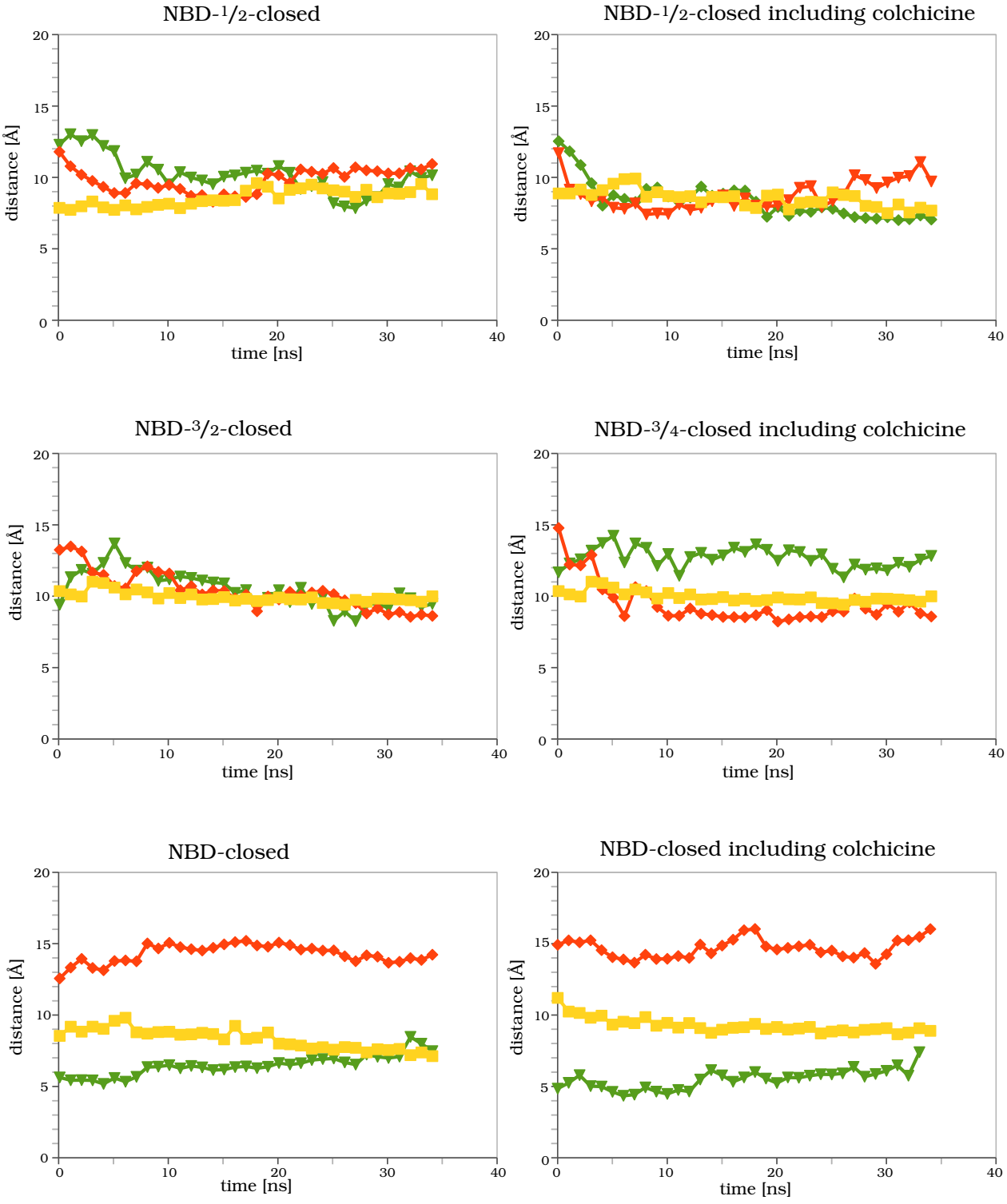


Figure 5.25: Distances of: **yellow**: *Leu219-Val338*, **red**: *Ala230-Ala348* and **green**: *Ser237-Ala356*.

5.4.3 TM4 Kink

The TM4 kink observed in the PDB-ID: 4Q9I crystal structure did not appear in any molecular dynamics simulation. All structures showed a straight helical conformation.

The distances of the two residues *Ser237* (TM4) and *Ala999* (TM12), stayed throughout the simulation constant (section 5.2.1).

Only the NBD-closed colchicine conformation showed higher distances, that occurred by a bending of TM4 toward the membrane.

The helical kink was only observed at the crystal structure PDB-ID: 4Q9I. Since no conformation did show any kink at this location and there were to date no other molecular simulation studies, in which a kink was reported, the TM4 kink may be a consequence of lattice contacts of the crystal form, or it may be induced by the ligand QZ-Ala which was co-crystallized [111].

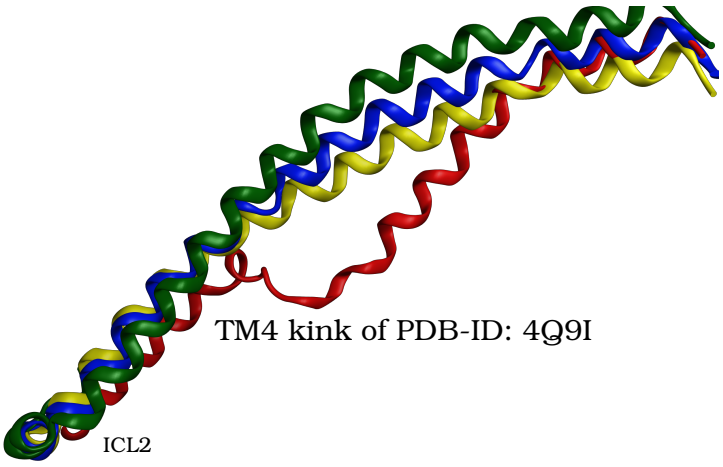


Figure 5.26: TM4 of PDB-ID: 4Q9I (**red**), NBD-open (**yellow**), NBD-1/2-closed (**blue**) and NBD-closed conformation (**green**).

5.5 Cross-linking

Cross-linking is a process of chemically joining two or more molecules by a covalent bond. This attachment results in a stabilization of the tertiary or quarternary structure. Proteins are typically cross-linked in a reaction involving a cross-linker and the side chains. The reactivity of a thiol (cysteine side chain) renders it as target. The cross-linker is usually a molecule with two reactive groups on either end, separated by a spacer. As a result, the cross-linker bridges a distance between residues at a distance influenced by the spacer (figure 5.27).

A mutation of the seven endogenous cysteines of P-gp to alanine has no effect on its activity [112]. In several cross-linking studies *Loo* and *Clarke* induced into this *Cys-less P-gp* cysteine residues into positions which were not only interesting for determination of the size of the drug-binding domain, but also for distances of the homologous halves which trigger ATP-ase activity.

In the following section the distances of several residues (figure 5.29) of all generated conformations of P-gp were analyzed during the molecular dynamics, to investigate if they correspond to cross-linking measurements.

All residues in the following section were not participating in the helical unwinding or structure loss of section 5.2.1. Only *Leu175* (TM3) and *Asn820* (TM9) were next to a hinge region.

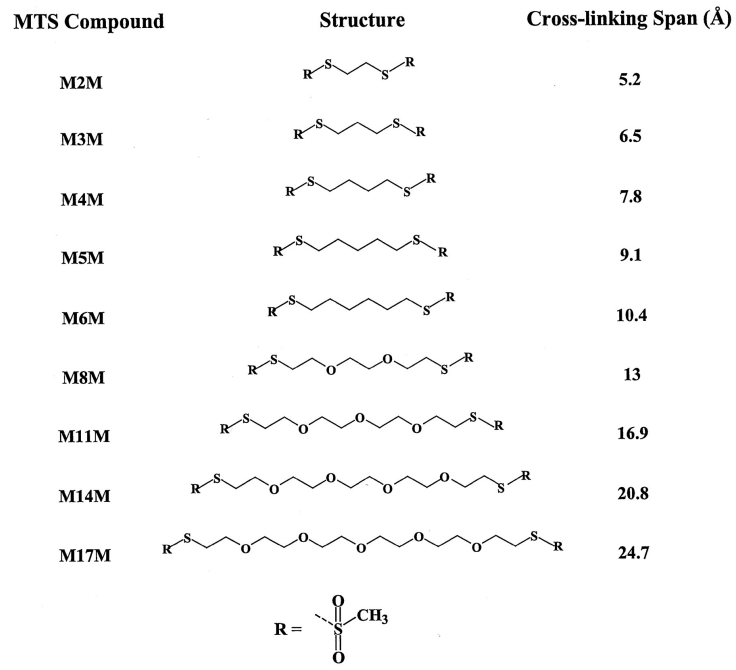


Figure 5.27: MTS (methanethiosulfonate) compounds and their spans in Å. For the distance of the C_α atoms of P-gp, 5.5 Å has to be added to the linker span [53] [65]. *This figure is taken from T. W. Loo and D. M. Clarke [65].*

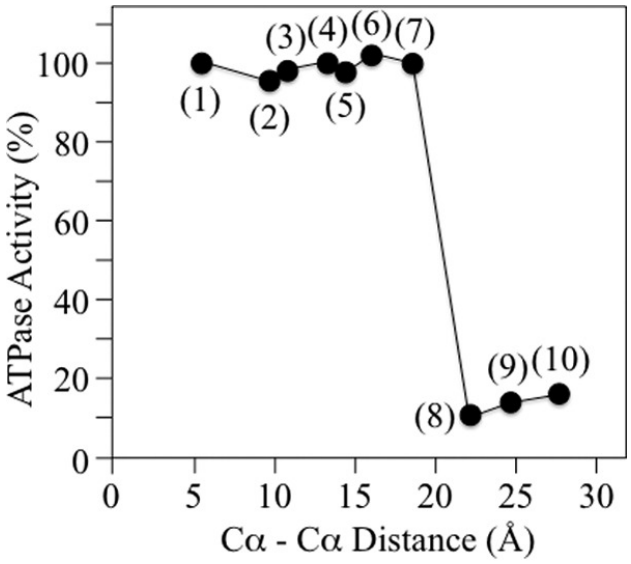


Figure 5.28: Correlation of ATPase activity of L175C/N820C treated with cross-linkers of various lengths and the distances (C_α atoms of P-gp). The cross-linkers were: M1M (2), M2M (3), M4M (4), M5M (5), M6M (6), M8M (7), M11M (8), M14M (9), or M17M (10).*This figure is taken from [113].*

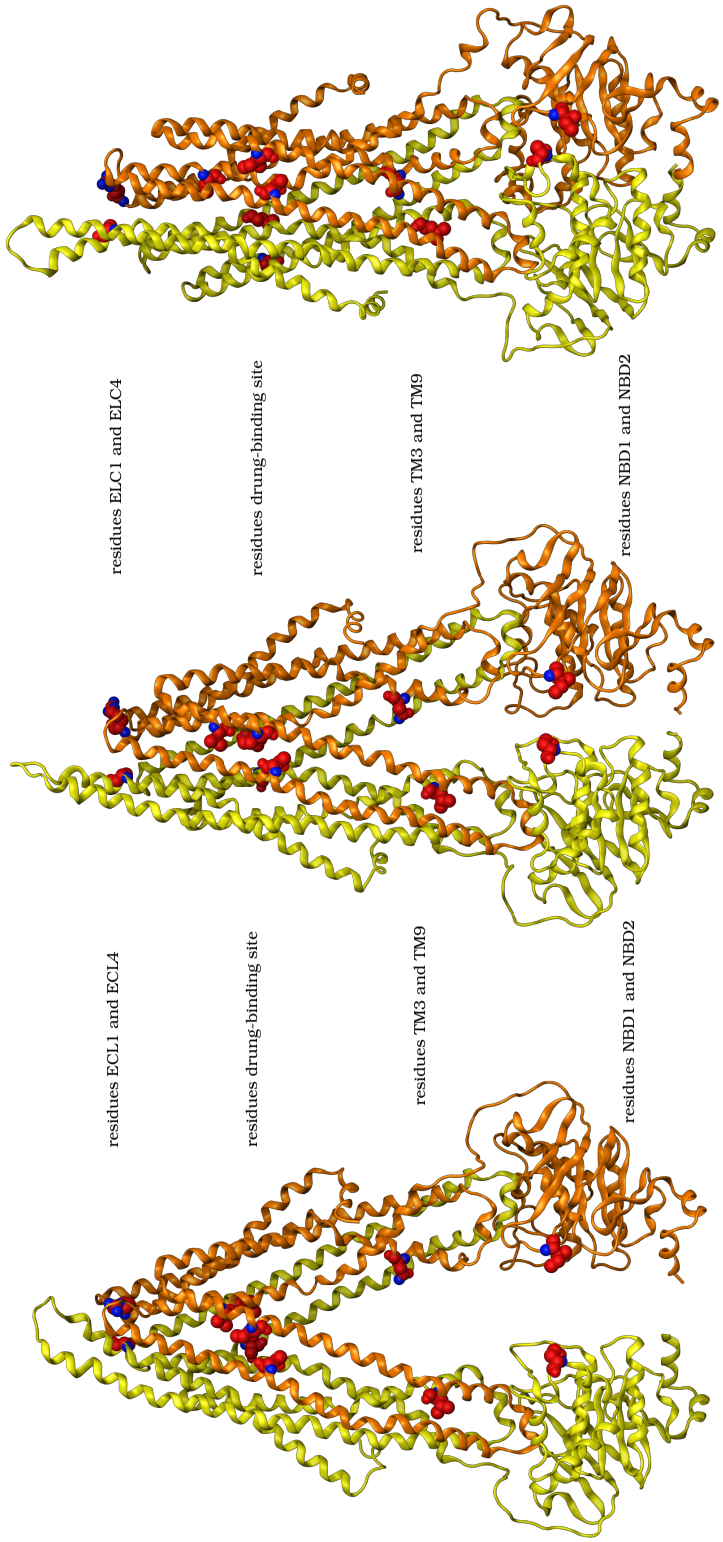


Figure 5.29: Residues (*red* marked) and their locations, which were compared to cross-linking data. **Left:** NBD-open P-gp; **middle:** NBD-1/2-closed P-gp; **right:** NBD-closed P-gp. TMD1-NBD1 *yellow* and TMD2-NBD2 *orange*.
The two NBD residues *Pro517* and *Ile1050* show clearly the twisting motion of the NBDs.

5.5.1 Distances of TM3 (L175) to TM9 (N820)

In cross-linking studies of *Leu175* and *Asn820* with cross-linkers of various lengths that separated the homologous halves of P-gp, T. W. Loo and D. M. Clarke observed that ATPase activity increased over 10-fold when the cysteines were cross-linked at distances between 6 and 19 Å, while cross-linking the two residues at distances greater than 20 Å yielded the activity at basal levels [114] (figure 5.28).

This observations correlate with the measured C_α distances of the P-gp morphing structures, where only the NBD-closed and NBD-^{3}/$_4$-closed conformation (approx. 18 Å throughout the simulation time) would represent a P-gp state, in which the ATPase activation switch appeared to be turned on (figure 5.30, *blue* and *green* graph).

Interestingly the NBD-^{1}/$_2$-closed conformation including the ligand colchicine, which represents the structure midway the open and closed P-gp structure, showed a decreasing distance averaging at 14 Å, and by this passing the ATP-ase switch of 19 Å, assuming the transporter closes to transport the substrate out of the cell, while the NBD-^{1}/$_2$-closed ligand-free protein fluctuated at an average of 19-20 Å(figure 5.30 and 5.31, *purple* graph). This behavior corresponds to T. W. Loo and D. M. Clarke suggesting a distance of 20 Å beeing the midway of the two known P-gp states.

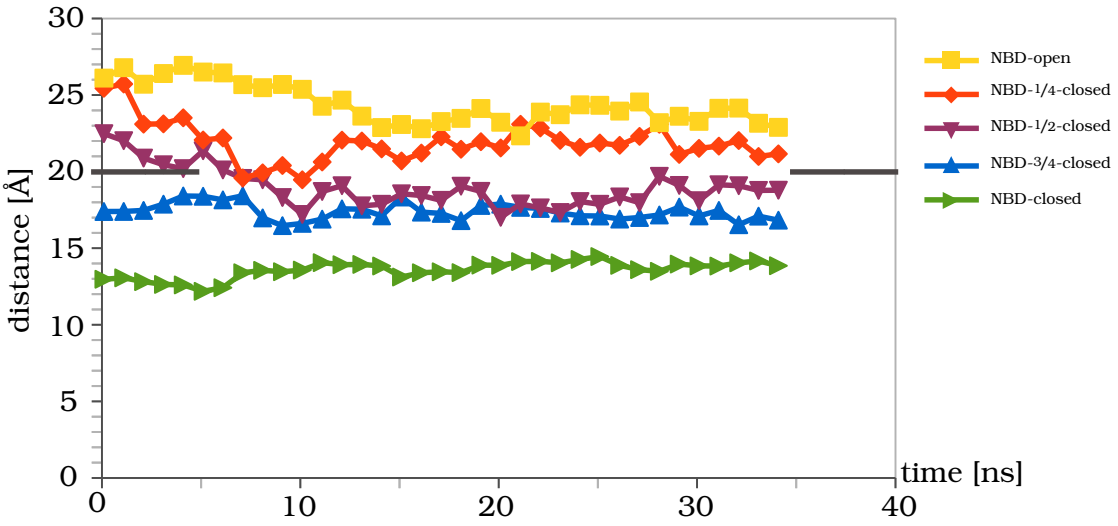


Figure 5.30: Distances of L175 to N820 of ligand-free P-gp conformations. The dark grey line marks the ATPase switch distance.

Also noticeable is, that all conformations which had colchicine within the drug-biding site, had lower values than the ligand-free conformations, suggesting that a ligand induces an approach of *Leu175* (TM3) to *Asn820* (TM9), which seems to be linked to the closure of the NBD (see section 5.2.1).

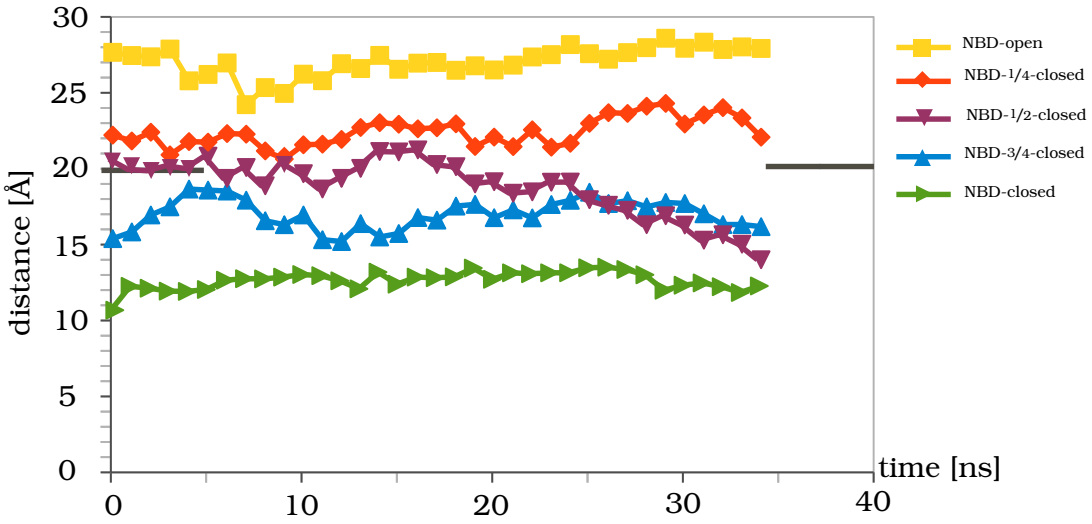


Figure 5.31: Distances of L175 to N820 of P-gp conformations with colchicine in the drug-binding site. The dark grey line marks the ATPase switch distance.

5.5.2 Distances of NBD residues P517 (NBD1) to I1050 (NBD2)

Trapping the central part of NBD1 (*Pro517*) and the N-terminal region of NBD2 (*Ile1050*) with a short linker (M4M) leaded into a close proximity of the NBDs (representing the NBD-closed conformation) and an activation of the basal ATPase activity of untreated P-gp by 14-fold, while a reduced ATPase activity was observed, after cross-linking these two residues with a larger linker (M17M) (linker sizes figure 5.27) [113].

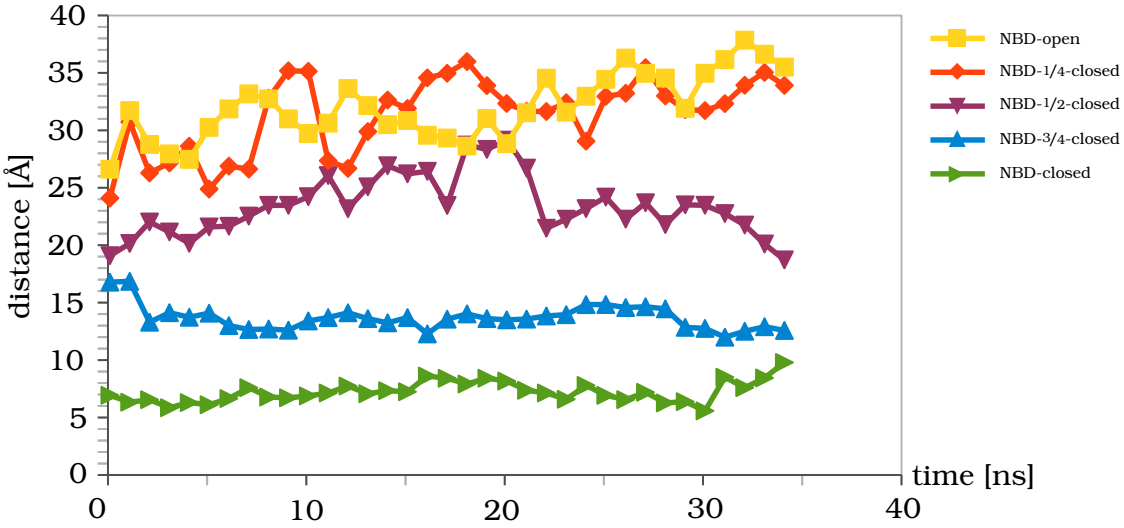


Figure 5.32: Distances of NBD residues P517 to I1050 of ligand-free conformations. While the NBD-closed (*green*) and NBD-3/4-closed (blue/) had constant distances, the other conformations showed fluctuations.

The distance measurements of all P-gp conformations showed, that only the NBD-closed conformation had the C_α distance of 8 Å, which increased after 31 ns to 11 Å (figure 5.32,*green* graph). The

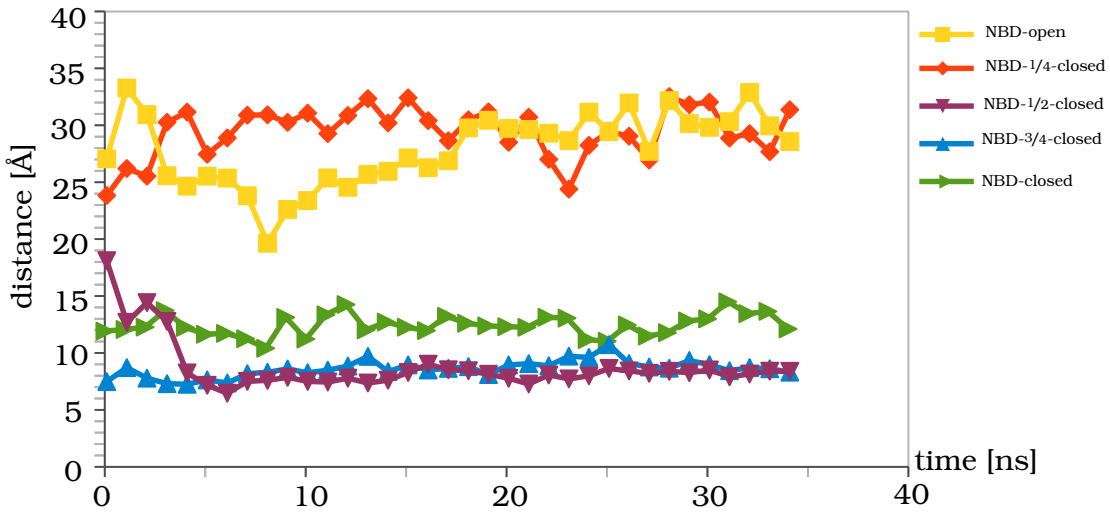


Figure 5.33: Distances of NBD residues P517 to I1050 of the colchicine included conformations.

NBD-3/4-closed fluctuated at a value of approx. 13 Å, which represents the maximum span of the M4M linker (figure 5.32, *blue* graph).

As soon as colchicine as substrate was in the drug-binding site of the NBD-3/4-closed and even NBD-1/2-closed conformation, the Walker A and LSGGQ sites moved in close proximity and therefore may increased by this movement the probability of generating the ATP-bound sandwich conformation. Both conformations reached an equal value of 8 Å. (figure 5.33, *blue* and *purple* graph)

This correlates very well with the results of section 5.5.1 were the NBD-1/2-closed including colchicine conformation passed the ATPase switch to initiate the transportation out of the cell.

Even tough, the conformations (NBD-open, NBD-$^3/_4$-closed), which showed a higher distance of *Pro517* to *Ile1050*, were at the beginning of the simulation within the measured range of a C_α distance of M17M of 28 Å, raised their distance even more throughout the simulation time (see figure 5.32), corresponding to the reduced ATPase activity, but contradicting the increased ATPase activity observed by T. W. Loo and D. M. Clarke after adding verapamil [113]. This observation may be explained by the M17M cross-linker, which still allowed the NBDs to move in close proximity or even increased the chance of doing so, by adopting gauche conformations [65], whereas during the simulation with no linker the NBDs could move freely, and thus moved more apart. The crystal structures PDB-ID 4Q9I-L and 5KPD: indicate, that a bigger separation between the two halves is probable [111, 115].

5.5.3 Distances of ECL1(A80) and ELC4 (R741)

The two residues *Ala80* and *Arg741* are located in the extracellular loop 1 (ECL1) and extracellular loop 4 (ECL4), respectively. Cross-linking studies showed that by trapping P-gp in a NBD open conformation the ATPase activity was inhibited [116]. Since the two residues are close together in a NBD-close conformation and widely separated in the NBD-open conformation, the the cross-linked P-gp would correspond to the NBD-open conformation.

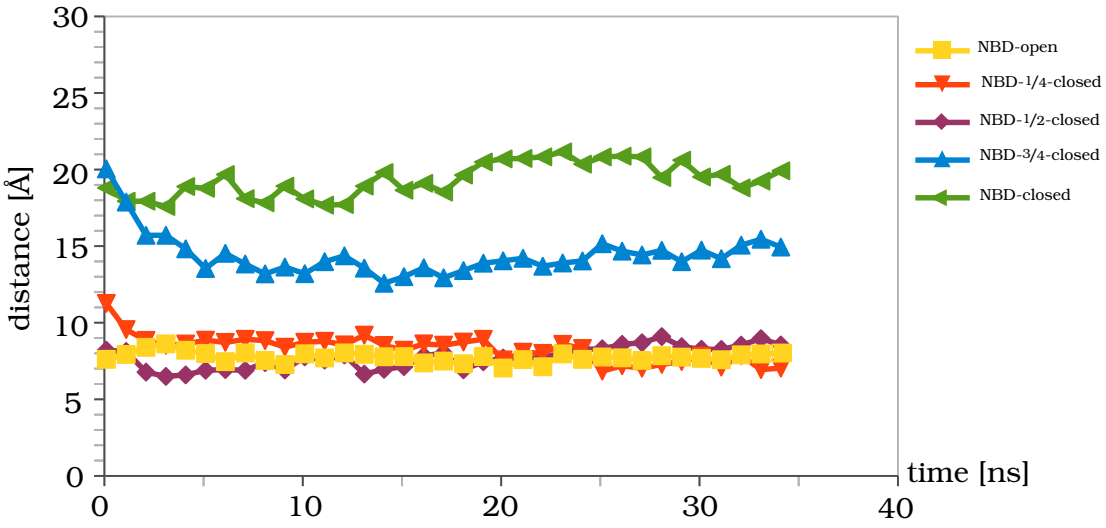


Figure 5.34: distances of A80 to R741 of the ligand-free P-gp conformations.

The molecular dynamics of the different conformations showed, that all measured distances were within the cross-linker spans of 6.9 Å and 30.4 Å (C_α atom distances). The NBD-open, NBD-1/4-closed and NBD-1/2-closed conformations had before the simulation a significant smaller distance than the other conformations

(figure 5.34, *yellow, red* and *purple* graph). This distance was constant, while the ECL1 and ECL4 of NBD-$^{3}/_{4}$-closed conformation approached each other (figure 5.34, *blue*), differing the measured distances of 30.4 Å for the NBD-closed conformation and contradicting the controversial of P-gp existing in cell with the NBDs in close proximity [117]. If P-gp would exist in cells preferring the NBD-closed conformation, the NBD-$^{3}/_{4}$-closed conformation should have adopted the NBD-closed ECL1 and ELC4 distance, since the starting structure had the same value of 20.4 Å.

Interestingly, the NBD-$^{1}/_{2}$-closed including colchicine conformation (fig. 5.35, *purple*) displayed a similar behavior as both NBD-$^{3}/_{4}$-closed conformations (ligand free and including colchicine), which corresponds to the similar distances of the NBDs of these conformations (see section 5.5.1 and 5.5.2). Also this conformation being on the midway of NBD-open to NBD-closed P-gp, had surprisingly constant distances, were one would expect to see a favorable trend to either NBD-open or NBD-closed conformation.

Neither NBD-$^{1}/_{4}$-closed, nor the NBD-open conformation changed distances during the simulation, independently if ligand-free or not. While the NBD-closed homology model had the referencing distance of 30 Å, but only the structure, which had colchicine in the drug-binding site.

The ECL1 is the most flexible part of the protein with the highest fluctuations throughout the simulations, this might be a reason that the distances of the two residues decreased in all conformations, where they were not close to each other.

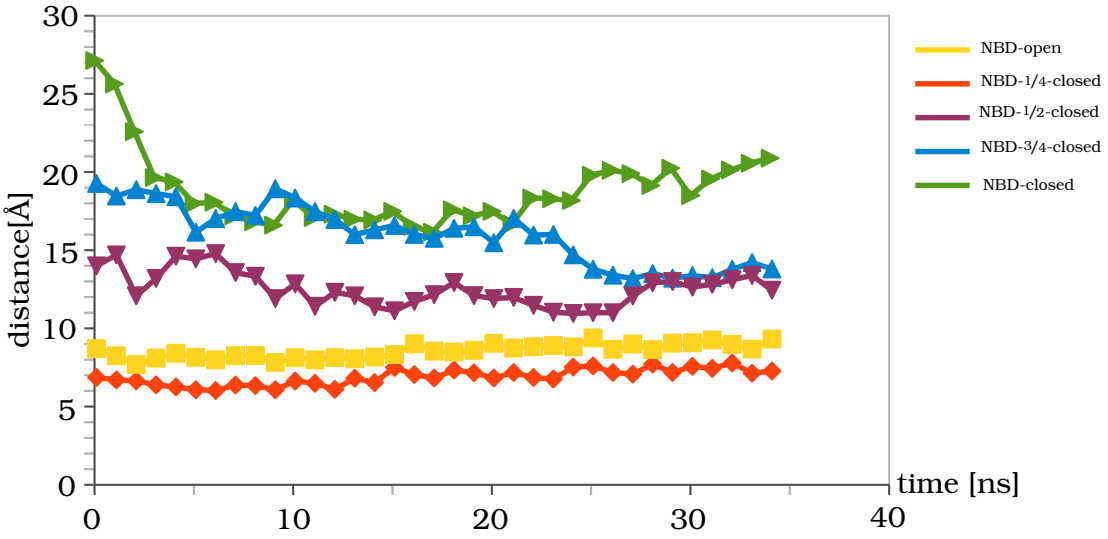


Figure 5.35: Distances of A80 to R741 of the colchicine included conformations.

The position taken by ECL1 and ECL4 in the P-gp conformations from NBD-1/2-closed towards NBD-closed may be the most comfortable for the two loops. An indication could be the NBD-1/2-closed including colchicine and both NBD-3/4-closed conformations reaching an almost equal value. Colchicine seems to facilitate the two loops to decrease their distance, likewise the NBDs and TM3 and TM9 residues *Leu175* and *Asn820* (see section 5.5.1 and 5.5.2).

5.5.4 Dimensions of the drug-binding site

Several residues, that were successfully cross-linked, were blocked by large substrates as vinblasitine and clyclosporin A. This results indicated that the cross-linked residues must line in a common drug-binding site, which may expand or contract to accommodate different sizes of substrates (figure 5.36) [65].

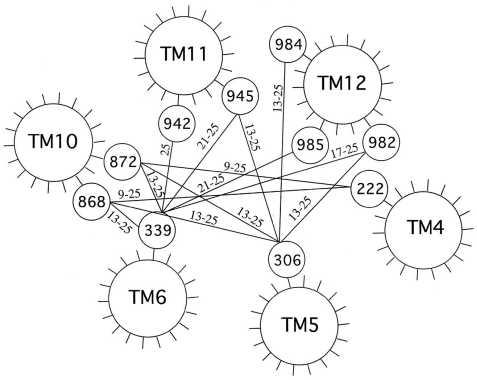


Figure 5.36: Top view of P-gp, where the residues of the TM segments are arranged as α-helical wheels. The circles refer to the positions of residues that line the drug-binding site of P-gp. The numbered lines represent the sizes [Å] of the cross-linker arms used to cross-link the two residues. *Figure taken from [65].*

The distances of *Ile306* (TM5) to *Ile868* (TM10), *Thr945* (TM11) and *Val982* (TM12) as well as the distance of *Ser222* (TM4) to *Ile868* (TM10) were measured during the 35 ns simulation time. All showed throughout the whole simulation time constant results (see table 5.4).

***Ile306* (TM5)** - ***Ile868* (TM10)** with a cross-linker distance span of 18-30.5 Å

Ligand free P-gp conformations	average distances of MD [Å]	Colchicine conformations	average distances of MD [Å]
NBD-open	25.2 ±0.49	NBD-open	25.3 ±1.14
1/4-closed	27.7 ±0.56	1/4-closed	27.2 ±0.55
1/2-closed	27.1 ±0.40	1/2-closed	29.2 ±0.37
3/4-closed	30.8 ±0.50	3/4-closed	31.7 ±0.93
NBD-closed	30.8 ±0.40	NBD-closed	30.0 ±0.89

***Ile306* (TM5)** - ***Thr945* (TM11)** with a cross-linker distance span of 18-30.5 Å

Ligand free P-gp conformations	average distances of MD [Å]	Colchicine conformations	average distances of MD [Å]
NBD-open	22.8 ±1.07	NBD-open	23.5 ±1.20
1/4-closed	25.3 ±0.53	1/4-closed	25.3 ±0.53
1/2-closed	24.5 ±0.56	1/2-closed	28.0 ±0.55
3/4-closed	28.9 ±0.40	3/4-closed	29.9 ±1.06
NBD-closed	29.6 ±0.29	NBD-closed	29.3 ±0.60

***Ile306* (TM5)** - ***Val982* (TM12)** with a cross-linker distance span of 18-30.5 Å

Ligand free P-gp conformations	average distances in MD	Colchicine conformations	average distances in MD
NBD-open	19.4 ±0.62	NBD-open	19.1 ±0.70
1/4-closed	20.1 ±0.71	1/4-closed	19.1 ±0.34
1/2-closed	17.4 ±0.51	1/2-closed	20.0 ±0.39
3/4-closed	19.8 ±0.46	3/4-closed	23.2 ±0.42
NBD-closed	18.5 ±0.98	NBD-closed	21.2 ±1.30

***Ser222* (TM4)** - ***Ile868* (TM10)** with a cross-linker distance span of 14.5-30.5 Å

Ligand free P-gp conformations	average distances in MD	Colchicine conformations	average distances in MD
NBD-closed	31.3 ±0.48	NBD-closed	32.9 ±0.67
3/4-closed	33.6 ±0.48	3/4-closed	34.3 ±0.81
1/2-closed	33.4 ±0.60	1/2-closed	33.6 ±0.79
1/4-closed	35.4 ±0.97	1/4-closed	36.5 ±0.98
NBD-open	32.7 ±0.71	NBD-open	37.5 ±0.81

Table 5.4: Dimensions of drug-binding site. The residues face the inner P-gp cavity. The distances are the average values of 35 ns simulation time.

All conformations showed distances which correlate with the higher values of the M11M, M14M and M17M (C_α-C_α distance of 27.5 Å) cross-linkers. The NBD-closed and NBD-3/4-closed conformations were at the upper limit of the cross-linking results. T. W. Loo and D. M. Clarke observed no evidence of cross-linking with the smaller analyzed cross-linkers M2M to M6M. The maximum distance of the M6M linker is 16.9Å (C_α-C_α distance), which means that the drug-binding site of P-gp seems to have a size of at least 16.9 Å. (see table 5.4).

The TM4 to TM10 distances of all conformations were 1-6.5 Å above the experimental results. Especially the two residues *S222* and *Ile868* of these transmembrane helices were also cross-linked with smaller reagents as the M8M linker [65].

Therefore one explanation could be that the TMs are quite flexible [118]. Since the linker sizes varied from M5M to M17M (spacer sizes 9.1-24.7 Å), it may be possible that due to their mobility TM4 and TM10 can be cross-linked by different linker sizes, but did not showed this behavior during the molecular dynamics simulation, because they did not had any reaction partner (linker) to decrease their distance. Also there are no experimental results with cross-linker spacer larger than 27.5 Å (C_α-C_α distances) to analyze if larger distances are probable.

The positions of the residues in the homology models did all face the inner P-gp cavity and would be thus accessible for cross-linking. *Ser222* of the NBD-closed conformation pointed into the direction of *Ile868*, but TM6 and TM12, which were located within the inner pore in this P-gp conformation would prevent a successful cross-linking in this conformation.

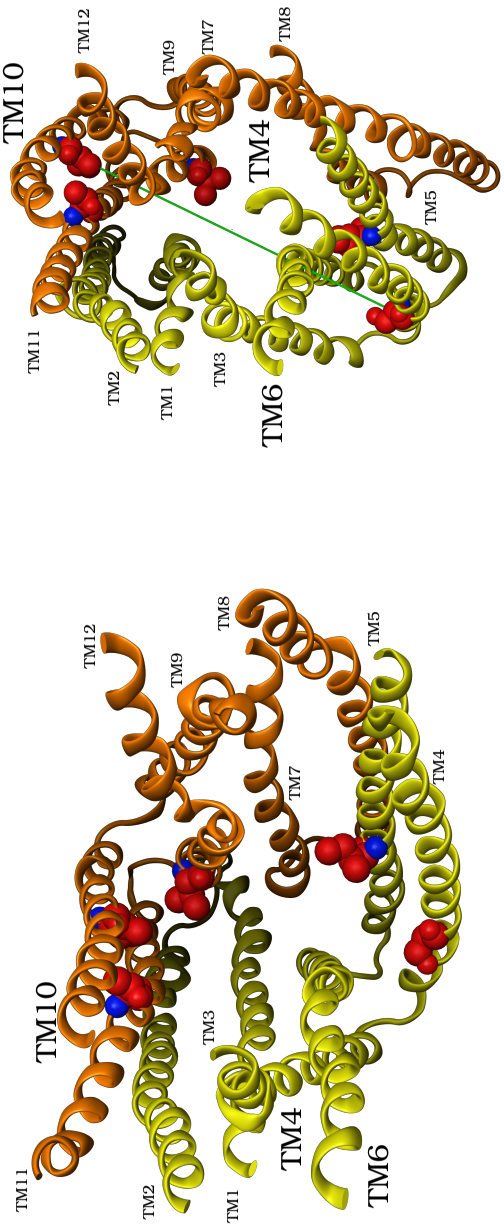


Figure 5.37: Residues which were taken to measure the dimensions of drug-binding site.**Left:** NBD-open P-gp; **right:** NBD-closed P-gp. TMD1-NBD1 marked *yellow* and TMD2-NBD2 marked *orange*. The *green* line of the NBD-closed conformation shows the direct connection of *Ser222* and *Ile868*, which is blocked by TM6.

Chapter 6

NBD-open conformation in complex with other ligands

P-gp transports a wide range of structurally diverse substrates, of which colchicine is in comparison to others a relatively small ligand, therefore additional ligands were docked into the NBD-open conformation.

This P-gp conformation was chosen, because it is the predicted high affinity state of P-gp and of the pre-defined drug-binding site of the co-crystallized ligand QZ59-RRR (section 4.3.1.1). The docking was performed identically to the colchicine docking (section 4.3.2). The position overlapping the colchicine docking position was selected.

The ligands docked were:

1. **Elacridar**, a *second generation* oral bioenhancer, that inhibits P-gp. Due to severe pharmacokinetic interactions, its development is assumed to have been discontinued.

2. **Tariqudiar**, a *third generation* P-gp inhibitor, that inhibits P-gp with a high affinity.

3. **Daunorubicin**, a toxic anthracycline used in treatment of leukemia and other neoplasms is a substrate of P-gp

4. **Hoechst 33342**, a fluorescent bisbenzimidazole derivative has been found to be transported by P-gp [119] and is used in cell based assays.

5. **cis-(Z)-Flupentixol**, an antipsychotic neuroleptic drug, was found to inhibit drug transport of P-gp [120, 121]

6. **trans-(E)-Flupentixol**, in contranst to its cis isomer, trans-(E)-flupentixol inhibits not only the drug transport of P-gp, but also the ATPase activity [122]

The movement of the NBDs (as in section 5.1) was monitored. Colchicine showed a drastic effect on the fluctuation of the NBDs of the NBD-open P-gp conformation, which indicated, that a ligand induced the NBDs to approximate each other.

The aim was to see, if all ligands show a similiar effect on P-gp or if the inhibitors elacridar, tariquidar and both flupentixols influenced P-gp differently.

In addition, the root mean square deviation of P-gp and the corresponding ligand versus the starting structure (energy minimized structure) were calculated to analyze, if the ligands escape and dissociate from the cavity of P-gp during the molecular dynamics simulations and if the different ligands have an impact on the stability of P-gp.

Due to high computational cost, this molecular dynamics simulations were performed only once.

6.1 Distances of NBDs

Comparing the distances of the NBDs at the end of the 35 ns molecular dynamics simulation, there was no significant difference between the P-gp substrates and the inhibitors. While colchicine showed a distance of 37.5-40 Å, daunorubicin had a value of 40 Å and Hoechst 33342 showed a value of ∼35 Å. Interestingly, daunorubicin and Hoechst 33342 displayed a similar behavior, by first increasing the NBD distance and then reducing it (figure 6.1 right). In contrast, colchicine decreased the NBD distances directly, so that an equilibrium was already reached after 10 ns (figure 6.1 left).

Since colchicine is a smaller molecule than daunorubicin and Hoechst 33342, it may have adapted easier and faster to its final position within the binding cavity of P-gp, indicated by reaching very fast the equilibrium state within the first 10 ns, while the other two bigger and more immobile substrates needed longer to position themselves, and therefore trigger the NBD closure.

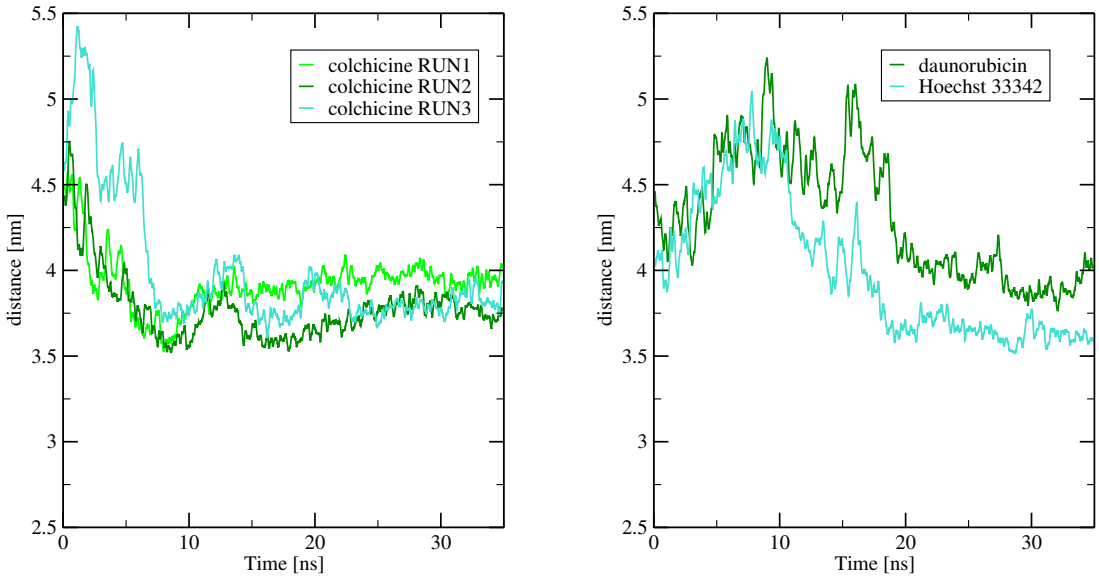


Figure 6.1: NBD-distance measurement of the NBD-open conformation for the substrates **colchicine (left)** and **daunorubicin and Hoechst 33342 (right)**.

The two inhibitors elacridar and tariquidar showed similar behavior to each other. Both, reduced slightly their distance from 45 Å to 38 Å and 35 Å, respectively (figure 6.2).

The slightly different starting points of the NBD-distances visible in the elacridar/tariquidar and daunorubicin/Hoechst 333432 graph (figure 6.1), as well as in the flupentixol graph (figure 6.2) occur from the equilibration phase before the actual simulations, in which in the very last step, position restrains, that held the Cα atoms of P-gp fixed, were released for 100 ps. Even in this extremely short time period, the protein adapted to its environment, and decreased the distances.

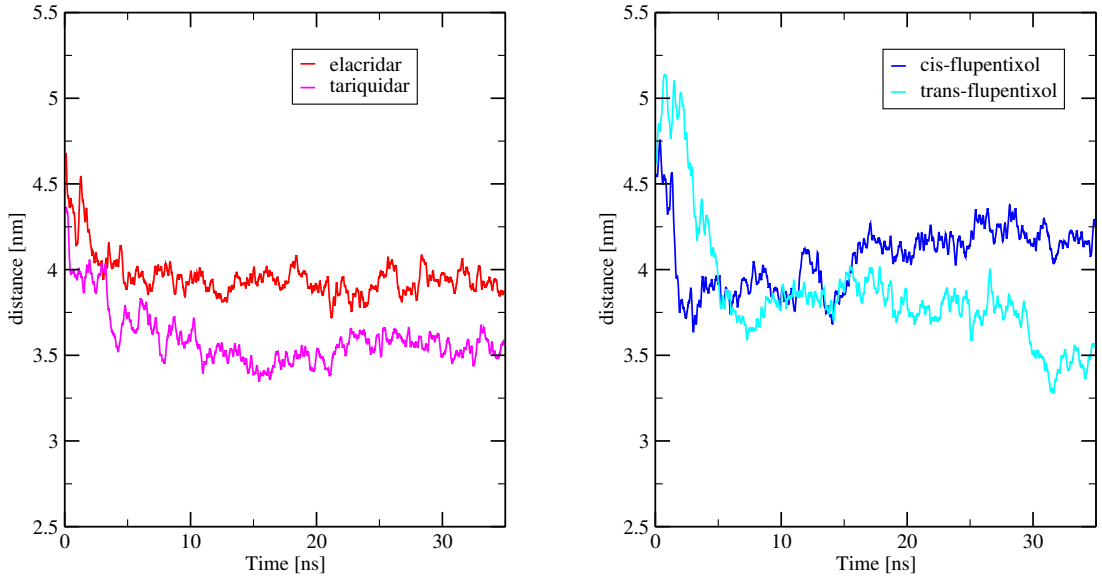


Figure 6.2: NBD-distance measurement of NBD-open conformation of inhibitors **elacridar and tariquidar (left)** and **cis- and trans-flupentixol (right)**.

Cis- and trans-flupentixol decreased the distance of the NBDs as well. In contrast, to the inhibitors elacridar and tariquidar, which showed an equilibrium state of the NBD distance, cis- and trans-flupentixol showed a drastic movement within the first 5 ns (figure 6.2), similar to colchicine. These two inhibitors, also, did not reach an equilibrium state, and while cis-flupentixol slightly increased, trans-flupentixol decreased throughout the remaining simulation time of 30 ns.

This results indicate, that the NBD-open P-gp conformation moves slightly towards the NBD-closed conformation, with no discrimination of the ligand bound in the drug binding site, since substrates as well as inhibitors, favored an approximation of the NBDs. Only the *third generation* inhibitor tariquidar and trans-flupentixol, reached a NBD distance close to the NBD-$^1/_2$-closed P-gp conformation (36 Å, see section 5.1).

The two inhibitors cis- and trans-flupentixol indicated the limitations of the short simulation time of 35 ns. Here, an extension of the simulation would be appropriate to investigate, if movement of the NBDs continues, or reaches an equilibrium.

6.2 Ligand movement during the simulation

The backbone RMSD of the performed simulations and the different ligands indicated, that the protein structure equilibrated quickly. In contrast, the RMSD of the ligands was more variable, showing their mobility within the binding pocket.

The three MD simulations of the NBD-open conformation with docked colchicine showed different behaviors of colchicine, while P-gp reached its end value of 6-7 Å within the first 5 ns. In the first and second simulation run, colchicine moved up to 9 Å within the protein, in the third molecular dynamics simulation it showed almost similar RMSD as the protein, but still with larger fluctuations (figure 6.3 and 6.4). It seemed colchicine had not the optimal position and interactions in the docked binding cavity and needed especially in the second and third run over 10 ns to reach its final location in the transporter. In agreement with the above section 6.1, when the NBD-distance equilibrated after 10 ns, here too colchicine reached its steady state after 10 ns, but still showed fluctuations (figure 6.4).

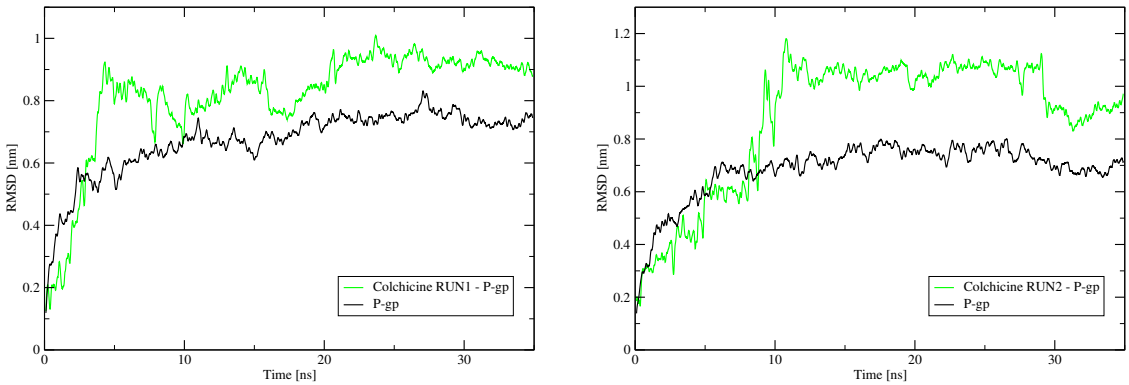


Figure 6.3: RMSD of NBD-open conformation and **colchicine**.

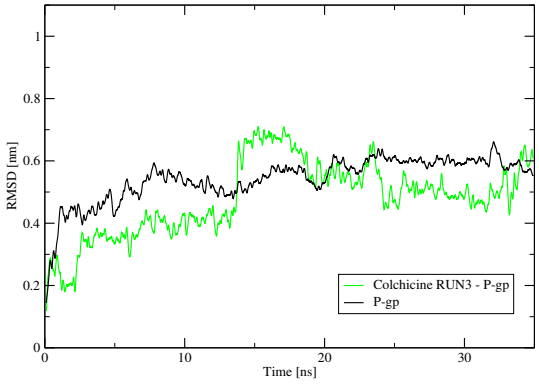


Figure 6.4: RMSD of NBD-open conformation and **colchicine**.

A RMSD, which is comparable to P-gp, showed also the substrates daunorubicine and Hoechst 33342. While daunorubicine and P-gp showed after 20 ns an almost identical RMSD rising from 5 Å to 6 Å, Hoechst 33342 jumped in the 25 ns from 2 Å to 8 Å indicating a translocation within the drug binding cavity (figure 6.5). This translocation of Hoehst 33342 emerged after the NBDs moved to their final state, with almost not existent fluctuations of the protein RMSD. Interestingly, P-gp displayed only with this ligand an improved RMSD value of 5 Å.

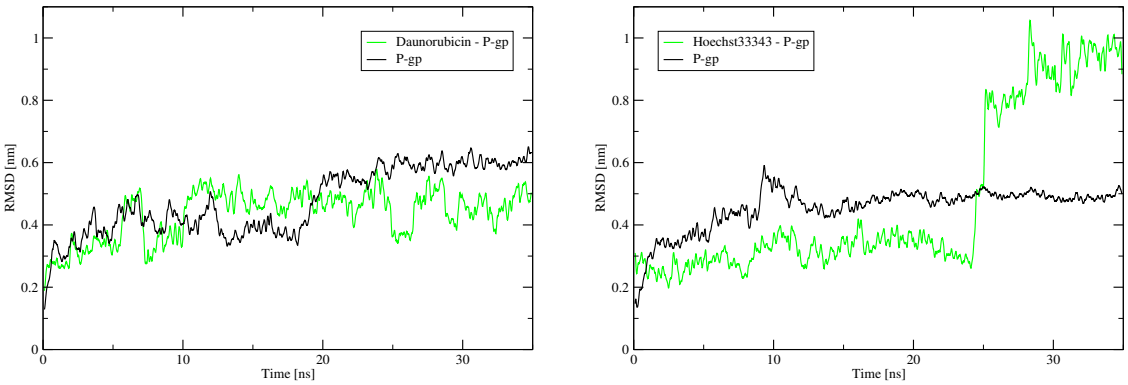


Figure 6.5: RMSD of NBD-open conformation and **daunorubicin (left)** and **Hoechst 33342 (right)**.

The closure of the NBDs of P-gp including daunorubicin is visible in the RMSD plot of the protein (figure 6.5 left). From the first state of a RMSD of 4 Å (12-18 ns) to a value of 6 Å in from the 18th ns onward. Why this is only with daunorubicin so clear visible is could not be clarified.

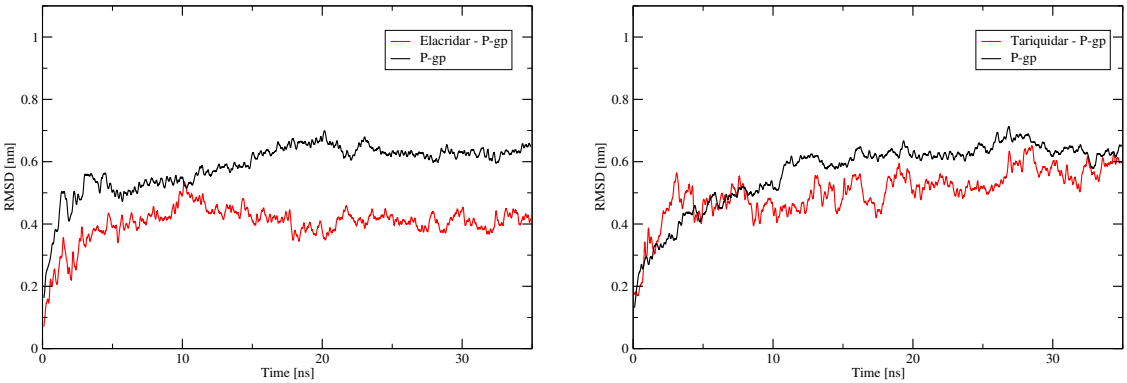


Figure 6.6: RMSD of NBD-open conformation and **elacridar (left)** and **tariquidar (right)**.

While P-gp itslef showed very low fluctuations, both inhibitors elacridar and tariquidar showed an even lower RMSD than the protein, suggesting, that they stayed fixed within the binding pocket (figure 6.6).

Cis-flupentixol and trans-flupentixol appeared to have a higher mobility in P-gp, than elacridar and tariquidar. Their RMSD was similar to the substrates. Comparing cis-flupentixol, which is a P-gp inhibitor that activates the ATPase activity, to trans-flupentixol, that inhibits P-gp without activating ATPase activity, small differences are detectable (figure 6.7). The intensity of fluctuations of cis-flupentixol is higher comparable to the substrates, while trans-flupentixol showed lower fluctuations after 15 ns similar to elacridar and tariquidar.

As with the NBD closure in the above section (section 6.1), a differentiation of substrate or inhibitor was not possible. All ligands showed a similar RMSD value for P-gp. This RMSD of 6-7 Å corresponds to the RMSD of P-gp without ligand.

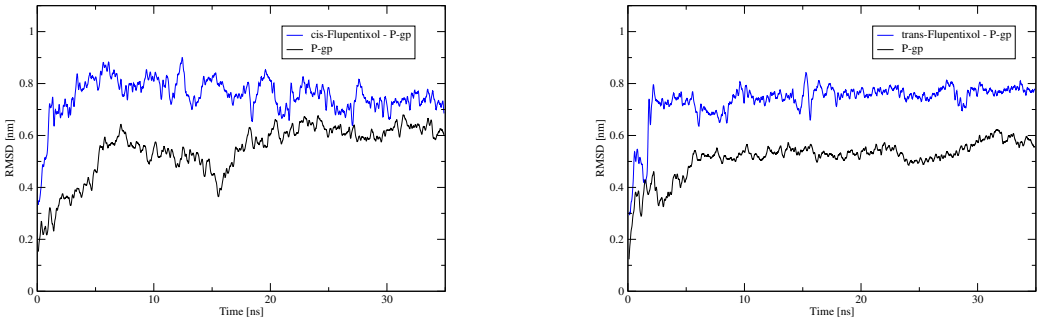


Figure 6.7: RMSD of NBD-open conformation and cis-flupentixol (**left**) and trans-flupentixol (**right**).

Differences of the RMSD of the ligands, could be caused by the ligand size. A small molecule, as colchicine, showed a higher mobility than a bigger and bulky molecule as daunorubicin and tariquidar. The jump of Hoechst 33342 after 25 ns may be caused by reduced interactions (H-bonds, aromatic interactions) of the ligand and the binding cavity. However, more investigations are needed for a statement (section 6.3).

The simulations were performed only once. Several and especially longer simulation runs (up to 50-100 ns) would be beneficial.

6.3 Drug-binding sites of ligands

As mentioned in section 1.3.4, to date, the high-affinity substrate binding site has not been conclusively resolved. To address this, the drug-binding cavity of the NBD-open conformation of P-gp with the different ligands was analyzed before, and after the 35 ns molecular dynamics simulations.

6.3.1 Colchicine

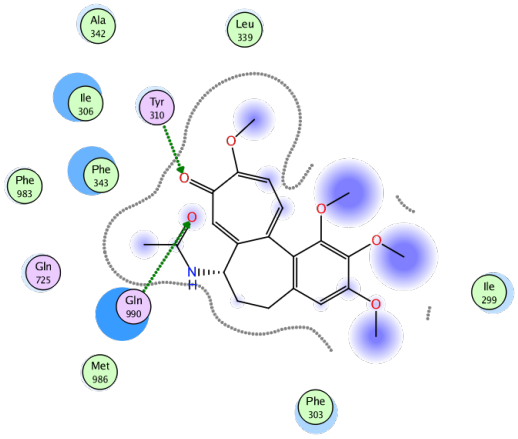
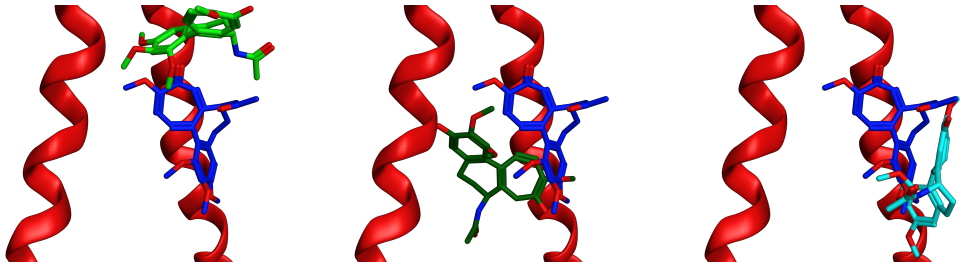


Figure 6.8: Ligand interaction of P-gp with **colchicine** before MD simulations.

Colchicine was docked into the putative H-binding site of P-gp located between TM5, TM6, TM7 and TM12 (see section 4.3.2). All residues, except (*Met986*), interacting with colchicine (figure 6.8) are predicted to face the drug-binding area or to be involved in drug-binding.

The three RMSD plots of colchicine in section 6.2 (figures 6.3 and 6.4) indicated the different movement of colchicine during the 35 ns simulations. While in the first and second run, colchicine fluctuated to a different position within the drug-binding cavity, in the third simulation the ligand stayed close to its initial position (figure 6.9).



(a) Colchicine RUN1. **(b)** Colchine RUN2. **(c)** Colchicine RUN3.

Figure 6.9: Postion of colchicine after 35 ns. In **blue** the initial position of colchicine is shown.

Each MD of colchicine showed movement to a different direction. While in the first and third simulation cholchicine moved towards the inner P-gp cavity, in the second run, it moved towards the lateral portal of TM4/TM6 and was "sandwiched" between TM4, TM5 and TM6 (see figure 6.11, left). By interactions with mostly residues of TM5 and TM6, the mobility of colchicine was reduced, this was visible by the very low fluctuations of the RMSD value of colchicine to the protein in from 10-28 ns and 30 ns onward (figure 6.3 right). Residue *Gly346* contributed two interactions with P-gp.

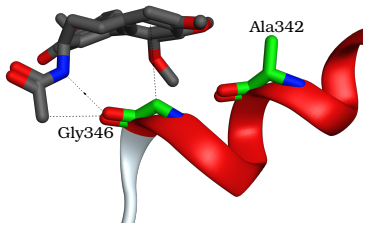
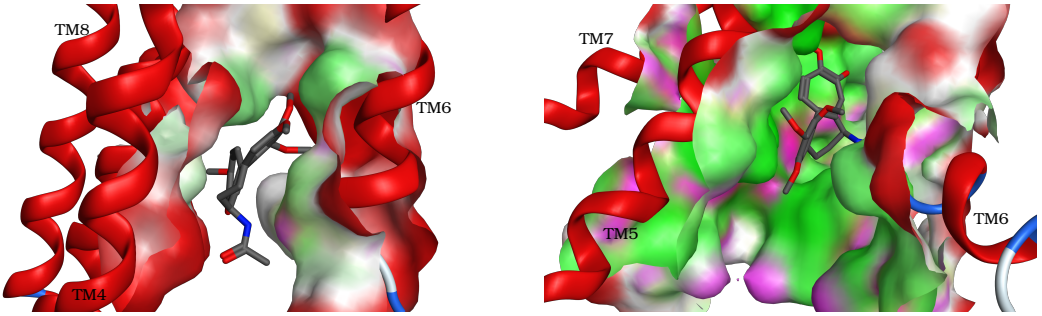


Figure 6.10: Orientation of Gly346 and Ala343.

No information about drug-binding of this residue is known. However it lies exactly one helix pitch next to *Ala342* (figure 6.10), which showed by substitution a resistance to colchicine [73].

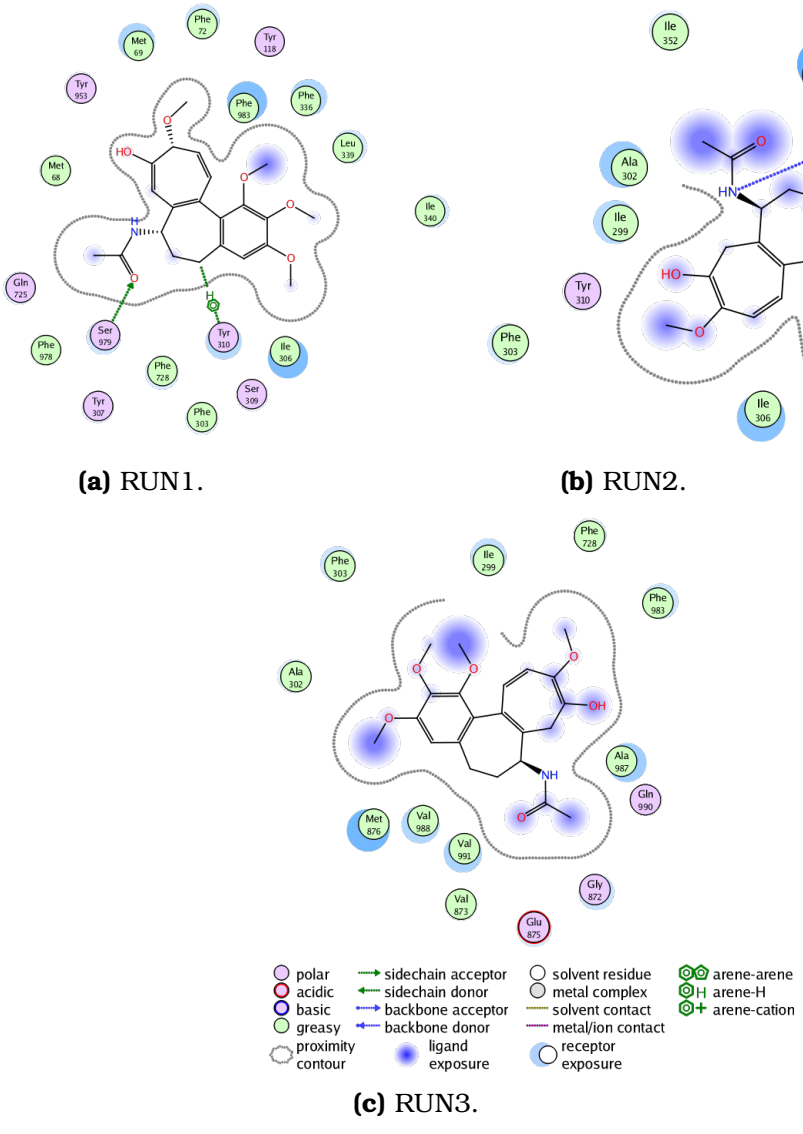
Likewise the visible tight holding of colchicine between TM4, TM5 and TM6, the exposure of the ligand into the inner P-gp cavity was displayed in higher fluctuations of the RMSD values (figure 6.3) and although colchicine did not show any H-bond interaction at the end of the simulation (third simulation run), the ligand did not escape far from its starting position (figure 6.12), concluding a compact binding cavity (see figure 6.11, right), and high affinity.



(a) RUN 2. **(b)** RUN 3.

Figure 6.11: Colchicine in the binding pocket. Views are from the lateral portal into the inner cavity of P-gp. On the left, between TM4 and TM8, barely visible lies TM5. Hydrophobic surface is shown in **green** and polar surface in **purple**.

Colchicine moving towards the inner P-gp cavity (first and thrird MD run) showed throughout both simulations interactions with residues *Gln725* and *Phe728* of TM7 and *Phe983* and *Ala987* of TM12 (see figure 6.12). Especially *Gln725* and *Phe728* seem to play an important role in drug-binding, since a mutation of one of these two residues led to a loss for substrates to inhibit the P-gp labeling with IAAP ([125I]iodoarylazidoprazosin) [68]. *Ile306*, which appeared to interact with the ligand in all three performed simulations, is suggested to be part the of signal cascade for the ATP hydrolysis [66].



(a) RUN1.

(b) RUN2.

(c) RUN3.

Figure 6.12: Ligand interaction of **colchicine**.

6.3.2 Daunorubicin

Residues interacting with daunorubicin (figure 6.14) were also present in the binding cavity of the first or the third simulation run of colchicine, corresponding to the twisting movement towards the inner protein cavity. In contrast two H-bonds existed with daunorubicin. The first H-bond by *Tyr310* showed in an arginine mutatinesis study of *Loo* and *Clarke* an inhibit of tariquidar-stimulated ATPase activity [64]. The second H-bond was formed by *Gln347*. This residue, as well as *Phe336, Leu339, Ile340* and *Phe343* (all TM6) are suggested by *Loo* and *Clarke* to face the drug pathway of P-gp, since mutagenesis studies showed a promoted maturation of the G251V P-gp mutant (see section2.4.2) [58].

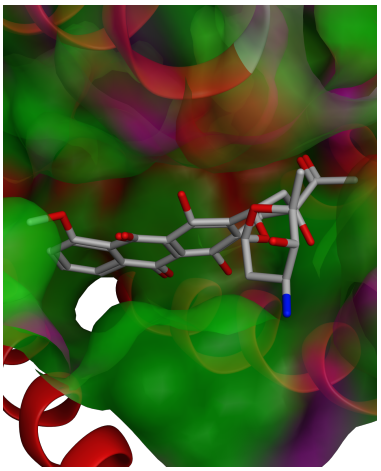


Figure 6.13: Daunorubicin in binding pocket of P-gp. Hydrophobic surface is shown in **green** and polar surface in **purple**.

Although daunorubicin did not escape as far as colchicine into the inner cavity of P-gp, suggesting strong interactions between the ligand and the protein, it showed fluctuations in the RMSD values. The strong interactions could be the result of the two H-

bonds holding the ligand into position and the surface of the planar tetracene moiety. The interaction surface within the binding cavity was not as tightly around the ligand, as e. g. the sandwiched colchicine, leaving enough freedom for daunorubicin to fluctuate, but still holding it back not to escape (figure 6.13).

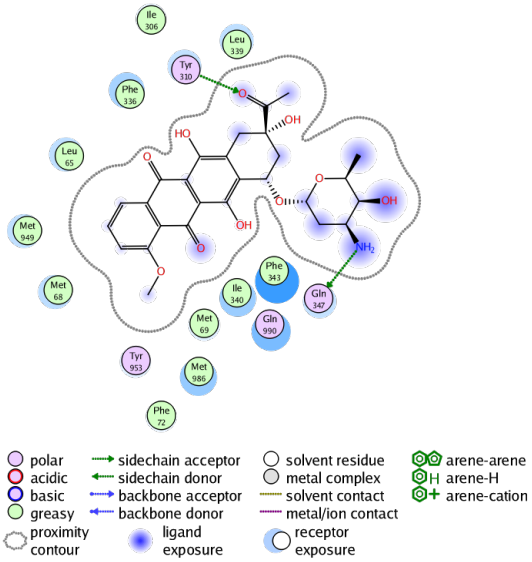


Figure 6.14: Ligand interactions of **daunorubicin** after 35 ns.

6.3.3 Hoechst 33342

The same mutagenesis study of *Loo* and *Clarke*, mentioned above, showed *Asn296*, which formed a H-bond with Hoechst 33342, to face the drug-binding area, too. Hoechst 33342, due to the length of the molecule, spanned the inner lipid bilayer and moved towards the transmembrane helices, than rather to the inner cavity. This may be caused by the span of the ligand witch leads to a bigger surface for the residues of P-gp to interact with. Also this movement may be the reason for the RMSD jump seen in figure 6.5.

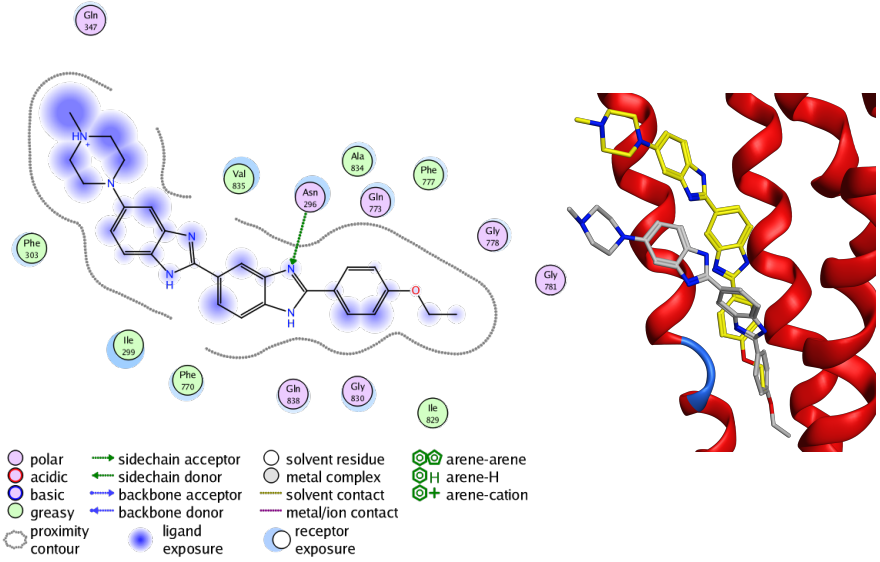
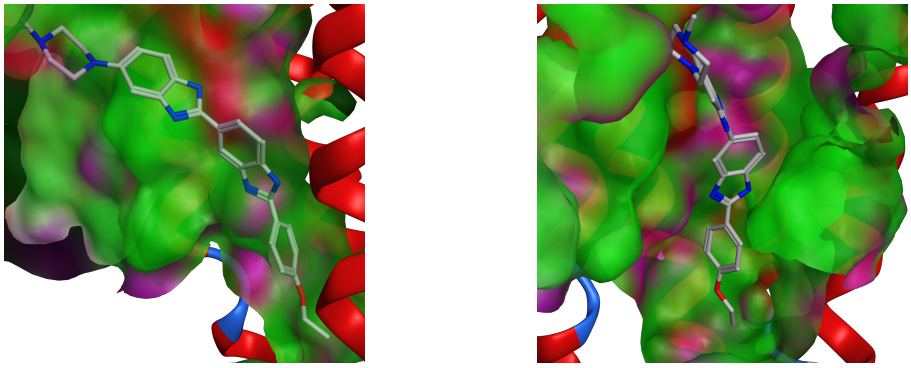


Figure 6.15: Left: Ligand interactions of **Hoechst 33342**.
Right: Movement of Hoechst 33342. **Yellow** shows the starting position and **gray** the position after 35 ns. The helices visible are TM5, TM8 and TM9.

Because of its position close to TM8 and TM9, it was the only substrate to interact with residues of these helices (figure 6.15). The one residue, that interacted also with other ligands was *Phe770*



(a) Front view **(b)** side view

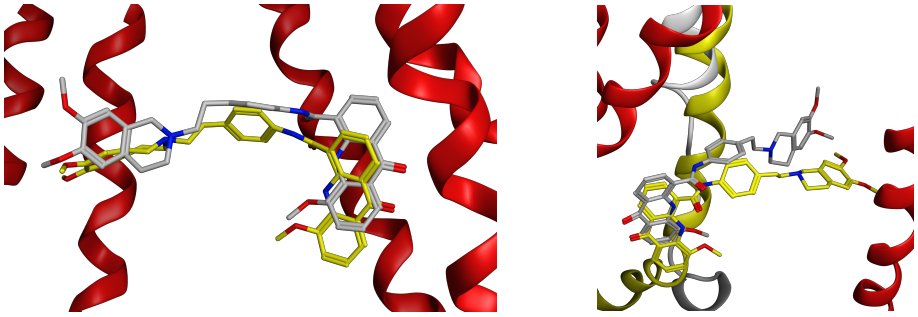
Figure 6.16: Hoechst 33342 in the binding pocket. Hydrophobic surface is shown in **green** and polar surface in **purple**.

of TM8, which is predicted to face the drug translocation path [58]. For all other residues from TM8 and TM9, that interacted with Hoechst 33342 no literature about drug-binding could be found. Only the piperazine moiety of the ligand faced the inner cavity, while the rest of the molecule was embedded between TM5, TM8 and TM9 (see figure 6.16).

The different movement within P-gp, caused Hoechst 33342 not to show as many interacting residue cross-overs with the other ligands. But one of the few cross-overs was *Phe303*.

6.3.4 Elacridar

Elacridar was docked into the binding cavity with the acridine moiety binding into the putative H-site (section 2.4.3). Two H-bonds were formed with *Asn271* and *Gln838*, both residues not mentioned in literature to be involved in drug-binding (figure 6.18). However *Asn721* is only one helix pitch away of *Gln725*. *Chufan et al.* mentioned *Gln725* to be in the primary drung-binding site, since a triple substitution with *Tyr307* and *Val982* led to a loss the ability to inhibit labeling of P-gp by the transport substrate [125I]-Iodoarylazidoprazosin (IAAP) [68]. All three mentioned residues showed interactions with elacridar throughout the molecular dynamics simulation.

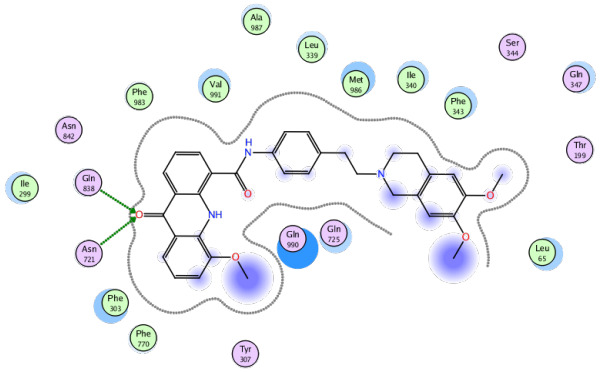


(a) front. **(b)** back.

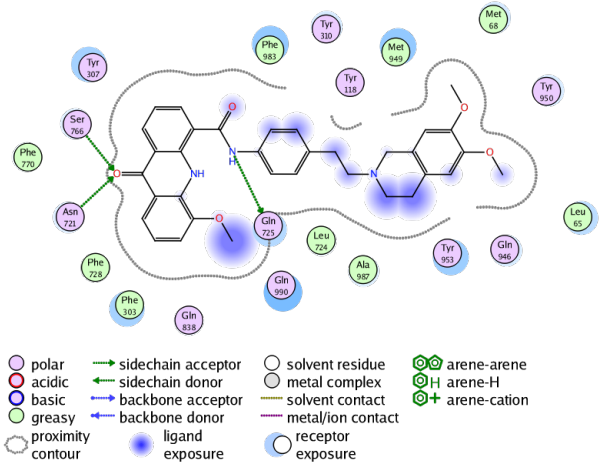
Figure 6.17: Movement of **elacridar**. Elacridar moved towards TM12 (b) and pushed it. **Yellow** before, **white** after 35 ns.

Residues from TM12 were: *Phe983, Met986, Ala987* and *Gln990*. Additionally *Val991, Ser993* and *Phe994* showed ligand interactions after docking elacridar into P-gp (figure 6.18). Taking these into account, TM6 had during the 35 ns MD simulation the biggest effect on the ligand. This could be a reason, why elacridar did not move towards the inner cavity and moved rather, because of

the strong and/or the amount of interactions, towards TM6. And therefore, contributed to the helical structure loss (figure 6.17). Further the low RMSD value from section 6.2 (ligand movement, figure 6.6) may be explained by elacridar moving along TM6 and staying close to the starting drug-binding pocket, rather than escaping to the inner protein cavity. Also the low fluctuation in comparison to the substrates could be caused by the strong interactions of two H-bondings to the acridine moiety.



(a) Start.



(b) After 35 ns.

Figure 6.18: Ligand interactions of **elacridar**.

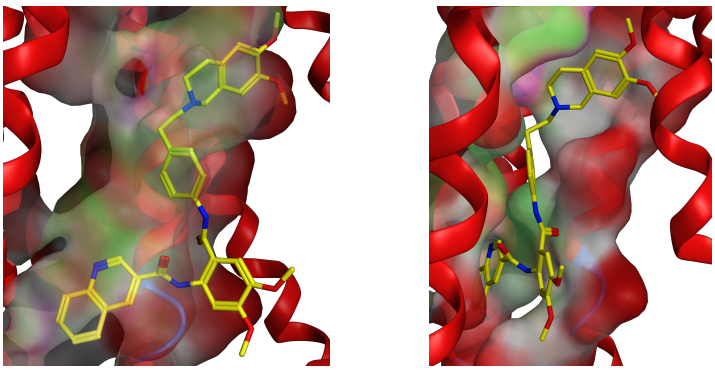
Because of the length of elacridar as molecule, the dimethoxy-isoquinoline moiety facing the inner cavity showed already in the beginning interactions with TM1 (*Leu65*). An introduction of arginine at the *Leu65* position disrupted interactions of tariquidar with P-gp, and is therefore suggested to be involved in drug-binding.

Residues from TM5 that interacted with elacridar were the same interacting with the substrates, namely *Phe303*, *Tyr397* and *Tyr310*.

6.3.5 Tariquidar

Similar to Hoechst 33342, tariquidar was docked into the putative H-site spanning the inner POPC leaflet. The dimethoxy-isoquinoline moiety was located close to the elacridar one, and the dimethoxyphenyl moiety was protruding the inner P-gp cavity. The size of the cavity led to the assumption of being responsible for the RMSD fluctuation, section 6.2 (figure 6.19).

The residues interacting with tariquidar showed cross-overs with elacridar (due to the similar docking position of the moieties) and Hoechst 33342, because of the size and location. The ligand was interacting mostly with residues of TM5 and TM12 (figure 6.20). As with all ligands (except elacridar and daunorubicin), *Ile299* was located in the binding cavity. This residue was prediceted by *Loo* and *Clarke* to be located at the distinct hydrophilic face of TM5, and consequently in the drug translocation pathway [58]. In addition *Phe303* and *Ile306* remained interacting with tariquidar throughout the simulation time.

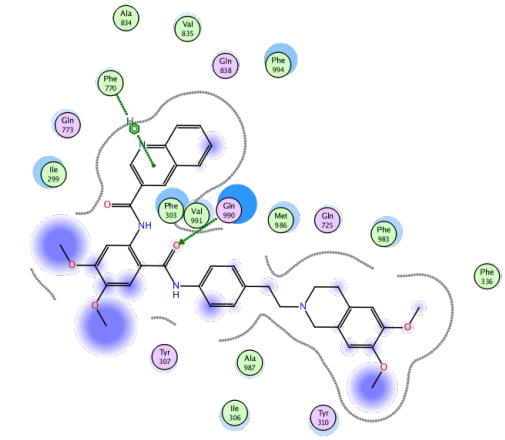


(a) Front view. **(b)** Side view.

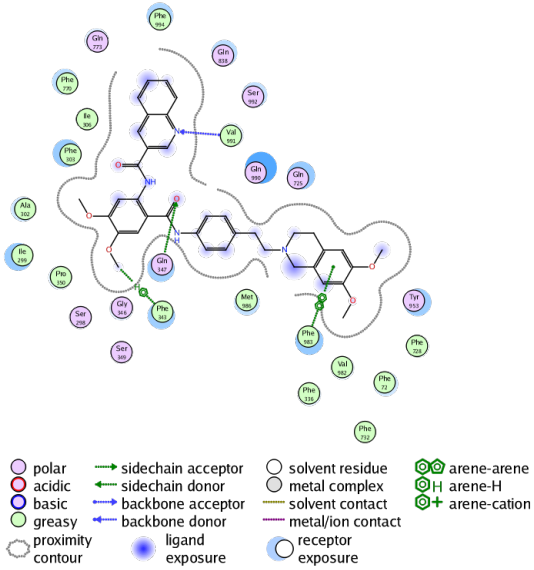
Figure 6.19: Binding pocket of **tariquidar**. Hydrophobic surface is shown in **green** and polar surface in **purple**.

Due to the movement of the inhibitor towards the inner P-gp cavity, interactions with residues from TM6 appeared during the molecular dynamics, namely *Phe336*, *Phe342*, *Gly346*, *Gln347*, *Ser349* and *Gly350*.

As with elacridar TM8 showed a H-bond interaction with tariquidar. The interacting residue was *Phe770*, which appeared also in the Hoechst 33342 binding pocket. A second H-bond was formed with *Gln990* from TM12. Both did not last during the 35 ns, but



(a) Start.



(b) After 35 ns.

Figure 6.20: Ligand interactions of **tariquidar**.

the residues remained interacting with tariquidar. At the end of the simulation the ligand appeared to form H-bonds with *Phe343*, *Gln347* (both TM6), *Phe983* and *Val991* (both TM12). These H-bonds, as well as the high number of residues interacting with the ligand of TM6 and TM12, which represent the linker of TMD to NBD, may attribute to the communication and stabilization of the protein, and lead consequently to a favorable NBD approximation of P-gp.

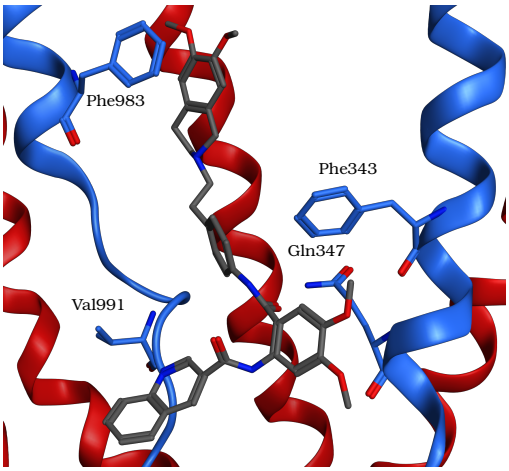


Figure 6.21: *Tariquidiar* in drug-binding site. TM6 on the left and TM12 on the right are shown in **blue**.

The RMSD fluctuation of tariquidar can be caused by the higher flexibility of the ligand in comparison to the relatively rigid elacridar. With three H-bonds holding tariquidar at the quinoline and dimethoxyphenyl moiety, the molecule is still pretty flexible to fluctuate in the inner P-gp cavity (figure 6.21).

6.3.6 Cis- and trans-flupentixol

Flupentixol was docked into the putative binding pocket, with the thioxanthen moiety facing the inner P-gp cavity. In contrast to the other ligands it did not show H-bond interactions (figure 6.23). The residue interactions were mostly for both flupentixols the same, with the difference, that trans-flupentixol showed additionally an interaction to *Met949*, which was not favorable to interact with cis-flupentixol due to the significant electronegativity of the trifluormethyl moiety pointing towards this residue. The other two differences occurred by latter reaching out to *Phe978* and *Val982*.

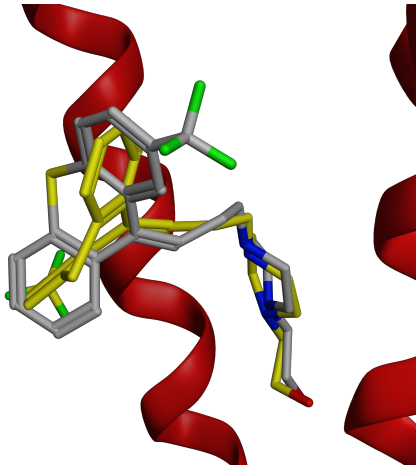
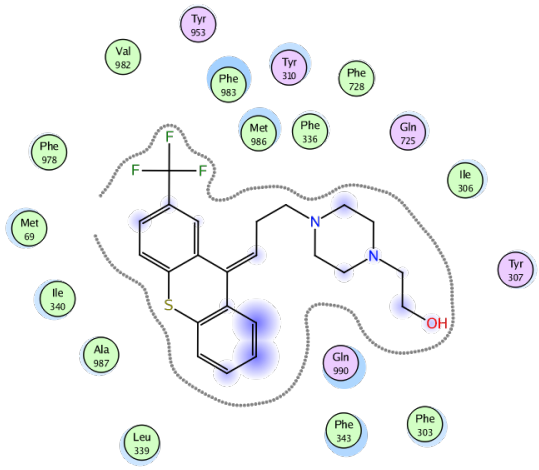


Figure 6.22: Docking position of **cis-flupentixol (grey)** and **trans-flupentixol (yellow)**. (*TM5 is visible behind the ligand, TM12 at the edge of the figure*).

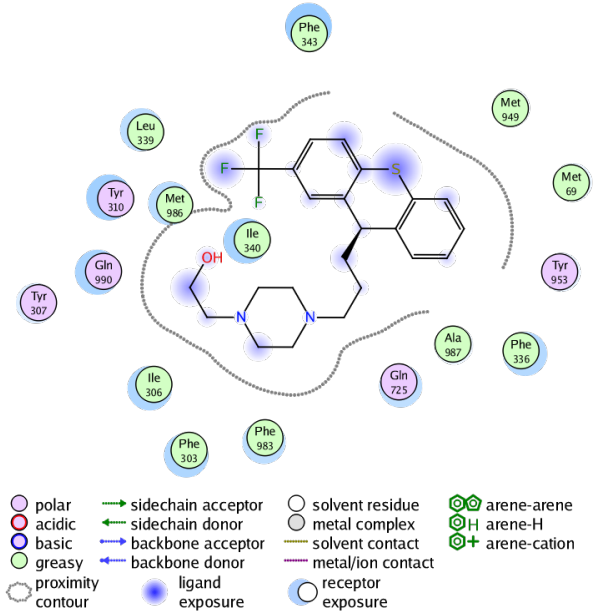
The residues of the docking pose of TM5 (*Phe303, Ile306, Tyr307* and *Tyr310*) corresponded to elacridar, and colchicine of the second and thrid MD simulations, as well as the residues of TM6 (*Phe336, Leu339, Ile340* and *Phe343*) to daunorubicin.

Gln725 and *Phe728* which were also identified to be in the drug-binding site of the other inhibitors elacridar and tariquidar, showed

only interactions before the MD simulations.



(a) Cis-flupentixol.



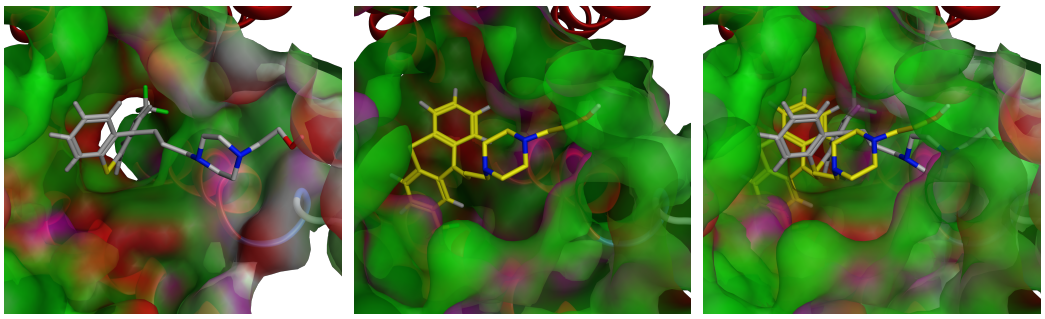
(b) Trans-flupentixol.

Figure 6.23: Ligandinteractions with P-gp before the MD simulations.

Both, cis-flupentixol and trans-flupentixol, moved during the molecular dynamics simulation towards the outside of the cell, whereby trans-flupentixol corresponding to the higher RMSD value (section 6.2, covered a farther distance. Therefore the residues interacting differed. By the twisting motion of trans-flupentixol, *Phe759* of TM8, was identified to interact with the ligand. To date, no information of this residue being involved in drug recognition is known.

In a cross-linking study, in which *Loo* and *Clarke* used several cross-linkers with different lengths to map distances in the drug-binding domain, *I868* was successfully cross-linked with *I306*, and therefore suggested to line in a common drug-binding site [65] Interestingly, this residue (*I868*), which did not show any interactions with the other ligands appeared in the cis-flupentixol binding site (figure 6.25).

Mutation of *Phe983*, performed by *Gottesman et. al.* indicated that this residue played an important role in stimulation and inhibition of ATP hydrolysis and [125I]IAAP labeling cis- and trans-flupentixol, respectively [122]. *Phe983* was not only permanent present in the flupentixol drug binding site but also in almost every other ligand.

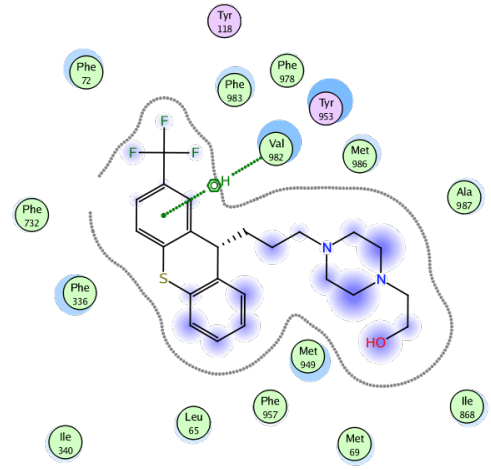


(a) Cis-flupentixol. **(b)** Trans-flupentixol. **(c)** Both.

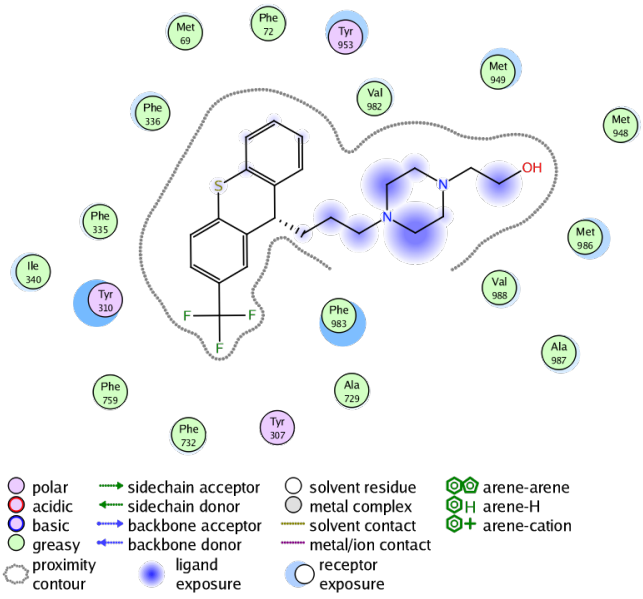
Figure 6.24: Binding pocket after 35 ns. **cis-flupentixol** is shown in **grey** and **trans-flupentixol** is shown in **yellow**. Hydrophobic surface in **green** and polar surface in **purple**.

A H-bond was formed by *Val982* and cis-flupentixol. This residues was present in almost all ligand interactions, and seems to play an important role in drug binding. Although cis-flupentixol showed this H-bond, it was fluctuating stronger in the drug-binding site than trans-flupentixol.

Both cavities showed space, for the ligands to fluctuate, therefore the reason for trans-flupentixol being more stabilized must be the thioxanthene moiety having more interactions with P-gp and consequently limit the mobility of the ligand (figure 6.25).



(a) Cis-flupentixol.



(b) Trans-flupentixol.

Figure 6.25: Ligandinteractions with P-gp after 35 ns.

These results of the ligands indicate, that P-gp does not have distinct binding pockets, since neither substrates nor inhibitor escaped the docked cavity. They rather adjusted themselves within the protein. If P-gp had different binding sites for the substrates and inhibitors, the fluctuations and the movement of one or the other should have been visible, which was not the case. This may be confirmed by colchicine which showed in three MD runs, three different ending locations. The high affinity of P-gp holding colchicine close to the docking positions, still could not prevent colchicine to move within the transporter. If a favorable binding site of colchicine existed in P-gp, the ligand should have moved towards it. However, to clarify this assumption, docking studies of different ligand poses and binding sites combined with molecular dynamic simulations would be beneficial.

Some residues seem to play a crucial role in drug-binding, e.g. *Met69* (TM1), *Phe303, Tyr310* (TM5), *Gln725* (TM7) and *Phe983* (TM12), because they not only appeared to be present in almost all ligand interactions, but also experimental data, as mutagenesis and cross-linking is known.

No distinction could be made between substrate or inhibitor. It was more a manner of size, if the ligand spans through the cavity reaching the opposite TMs as elacridar, tariquidar and daunorubicin did. Or if the ligand is relatively small to escape from the docking site as case of colchicine and flupentixol. Or it depended by the positioning within the binding pocket, as Hoechst 33342 and tariquidar (spanning the inner POPC leaflet).

The interactions with P-gp were mostly hydrophobic with very

few H-bonds. This may be also the reason for the inner cavity not differentiating the ligands, since most residues have non-polar sidechains. However the residues, that showed interactions with almost all ligands were polar, namely glutamine and tyrosine.

6.4 Gate of TM4/TM6

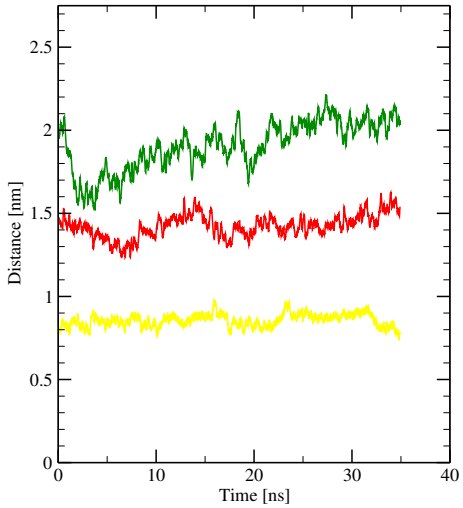
Analog to section 5.4 distances of Cα atoms of residues *Ala230* (TM4) and *Ala248* (TM6) located in the inner POPC leaflet and *Ser237* (TM4) and *Ala356* (TM6) located at the edge of the membrane to the inner cellular space were measured.

The residues *Leu219* and *Val338* were measured too, but will not be mentioned anymore in this section, since they stayed as in the previous investigation of section 5.4 constant throughout all simulations.

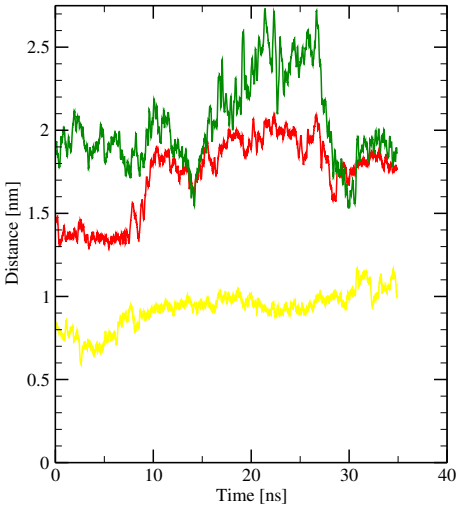
The NBD-open conformation showed in its ligand-free state a distance decrease from 17 Å to 11.6 Åfor the residues located at the edge of the membrane, while colchicine showed an insignificant small decrease of 2 Å, which resulted to a distance of still 17 Å (see figure 5.23), which should be large enough for substrates to enter P-gp.

The residues *Ala230* and *Ala248* in the inner membrane leaflet, had a similar spacing of 12-13 Å, but the ligand-free conformation showed higher fluctuations in comparison to the colchicine including conformation (figure 5.23).

While colchicine showed a marginally closure of the portal, the distances of *Ser237* and *Ala356* of the daunorubicin NBD-open conformation reached a value of 16 Å within the first 5 ns, only to equilibrate after 20 ns at their starting distance of 20 Å (see figure 6.26a). The other two residues *Ala230* and *Ala248* fluctuated around their initial distance of 15 Å. In comparison to the ligand-free fluctuations of $\sim$4 Å this fluctuations of *Ser237* and *Ala356*, as well as *Ala230* and *Ala248* were negligibly small (figure 6.26).



(a) Distances of portal residues for the NBD-open including **daunorubicine** conformation



(b) Distances of portal residues for the NBD-open including **Hoechst 33342** conformation

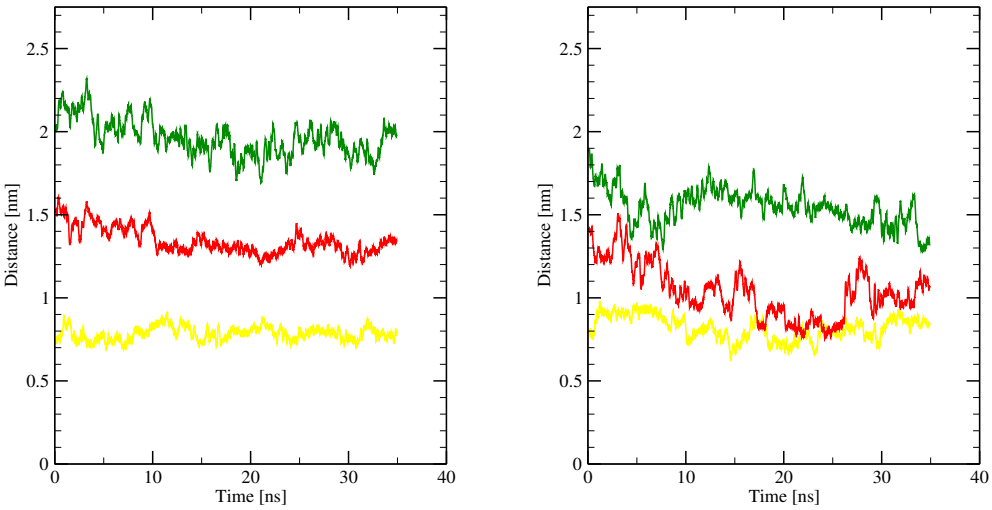
Figure 6.26: Distances of: **yellow**: *Leu219-Val338*, **red**: Ala230-Ala348 and **green**: Ser237-Ala356. While the ligand-free conformation decreased the distance of Ser237 to Ala356, the substrates kept the portal open. *(Although the distances are referred in the text in Ångström, due to the used graphical program, they are shown in nano meters).*

Unlike daunorubicin which seemed to stabilize the TM4/TM6 portal of P-gp similar to colchicine, Hoechst 33342 showed an impact on the distances. Both residue couples increased their distance opening the gate. The "lower" residues *Ser237* and *Ala356* lying at the edge of the membrane showed extreme fluctuations with a minimum distance of 15.5 Å in the 15th nano second, followed by a maximum distance of ∼25 Å for a duration time of almost 10 ns, before they decreased the distance to a value of 18 Å for the remaining simulation time of 5 ns (see figure 6.26).

Ala230 and *Ala348* increased their distance to each other simultaneously to the above mentioned residues (8 ns), reaching a

value of ∼19 Å, which they kept for over 15 ns, only to reduce the distance to ∼18 Å after a sudden decline in the 15th nanosecond to a minimum distance of 16 Å.

Again, elacridar influenced P-gp as ligand in a similar way as the substrates colchicine, daunorubicin and Hoechst 33342 did. The portal between TM4 and TM6 remained "open" stabilized by the inhibitor. Both residue couples, *Ala230-Ala348*, as well as *Ser237-Ala356* kept their distance at an equilibrated value of 13 Å and 20 Å, respectively (figure 6.27).



(a) Distances of portal residues for the NBD-open including **elacridar** conformation.

(b) Distances of portal residues for the NBD-open including **tariquidar** conformation.

Figure 6.27: Distances of: **yellow**: *Leu219-Val338*, **red**: Ala230-Ala348 and **green**: Ser237-Ala356. While elacridar kept the distances, tariquidar reduced them to a closure of the portal *(Although the distances are referred in the text in Ångström, due to the used graphical program, they are shown in nano meters).*

In contrast, tariquidar showed a reduced distance. *Ala230* and *Ala348* defining the upper part of the portal, had a value of ∼10 Å.

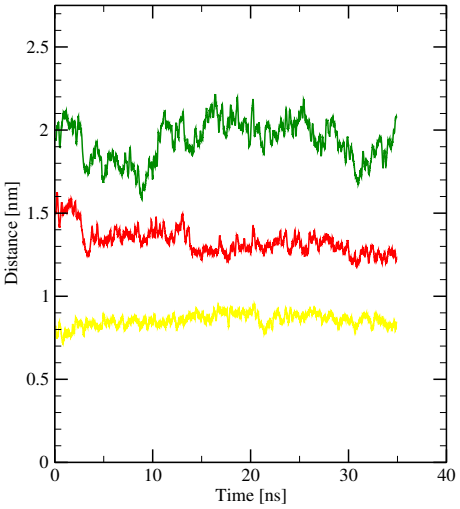
While *Ser237* and *Ala356* reduced their distance steadily to 13 Å. These two distances are comparable to the ligand-free NBD-open conformations, with the only difference of the upper residues coming by 3 Å closer to each other (figure 6.27).

In discrepancy to each other, elacridar seemed likewise colchicine to stabilize the gate, while tariquidar seemed to induce a closure of the gate. The distances of the residue couple measured should be 14 Å and higher as reached by Hoechst 33342.

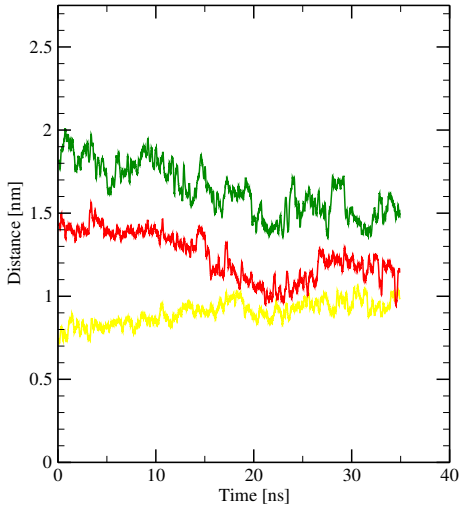
Interestingly, cis-flupentixol and trans-flupentixol showed a similar behavior to each other as elacridar and tariquidar.

While cis-flupentixol reduced the distance of *Ala230* to *Ala348* by ∼2Å, and showed fluctuations of *Ser237-Ala356* around the starting value of 20 Å, the inhibitor trans-flupentixol, which also inhibits ATPase activity of P-gp, reduced the distances of both measured residues couples steadily from a starting value of 14 Å to ∼11 Å and from 19 Å to 15 Å, respectively (figure 6.32).

To enable ligands to pass the lateral portal of TM4/TM6, it should show distances of at least 14 Å for *Ala230-Ala348* and *Ser237-Ala356*, since the distances were measured from the Cα atoms of the residues (figure 6.29). The partially bulky sidechains of residues forming the gate may obstruct the way into the cavity of P-gp, and therefore minimize the gateway into the protein. While colchicine is in figure 6.29 clearly visible and appears to be capable to pass the gate, in figure 6.30, which shows the surface of TM4 and TM6, respectively, the way for the ligand into the inner cavity seems more demanding.



(a) Distances of portal residues for the NBD-open including **cis-flupentixol** conformation.



(b) Distances of portal residues for the NBD-open including **trans-flupentixol** conformation.

Figure 6.28: Distances of: **yellow**: *Leu219-Val338*, **red**: Ala230-Ala348 and **green**: Ser237-Ala356. While cis-flupentixol kept the distances, trans-flupentixol decreased them. *(Although the distances are referred in the text in Ångström, due to the used graphical program, they are shown in nano meters).*

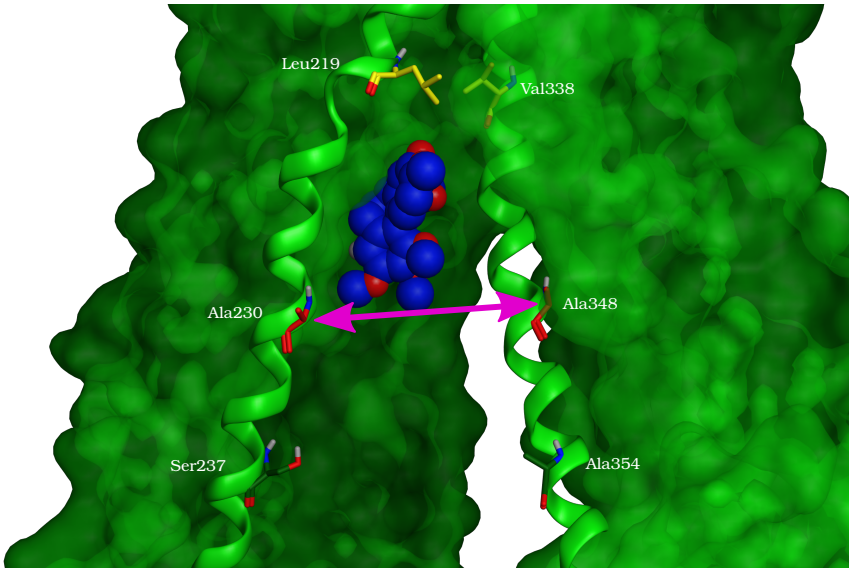


Figure 6.29: The **purple** arrow shows the dimensions of the initial portal of TM4/TM6, measured by Cα atom distances. Inside the P-gp cavity lies **colchicine** (shown in **blue** spheres.) The residues are shown in **yellow**: *Leu219-Val338*, **red**: *Ala230-Ala348* and **dark green**: *Ser237-Ala356*.

The importance of this minimum distance of $\sim$13.5 Å for *Ala230* and *Ala348* was also indicated by flupentixol as ligand (figure 6.31). Cis-flupentixol showed a lateral portal between TM4 and TM6 (figure 6.32). The snapshot after 35 ns of the surface of P-gp including trans-flupentixol showed an almost not existent portal. The small gap located at the edge of the membrane is smaller than 11 Å (measured Cα atom distance of *Ser237* and *Ala356*), which may be adequate for small molecules. Substrates are suggested to enter P-gp from the membrane and not directly from inside of the cell [25], and therefore the distance of TM4 and TM6 to each other inside of the cell should not influence the substrate entry.

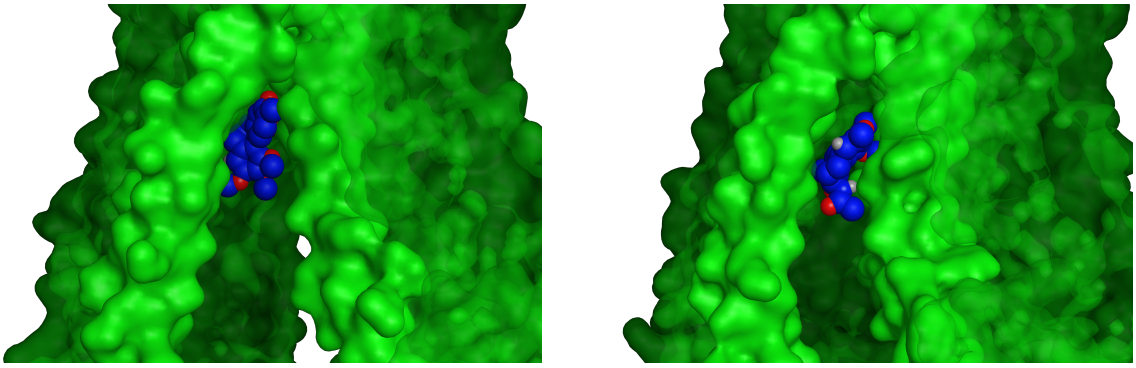


Figure 6.30: Surface of P-gp. Colchicine (shown in **blue** spheres), located inside the cavity, is visible through the TM4/TM6 portal. On the **left** the initial state, and on the **right** after 35 ns.

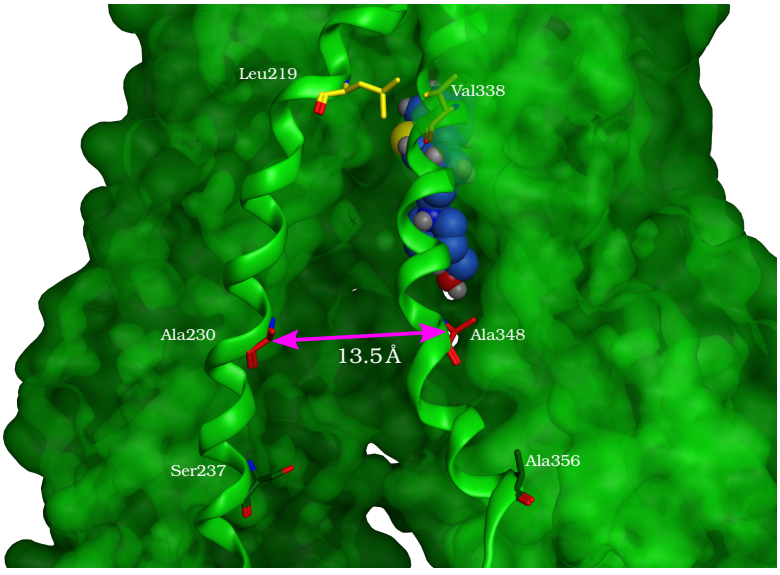
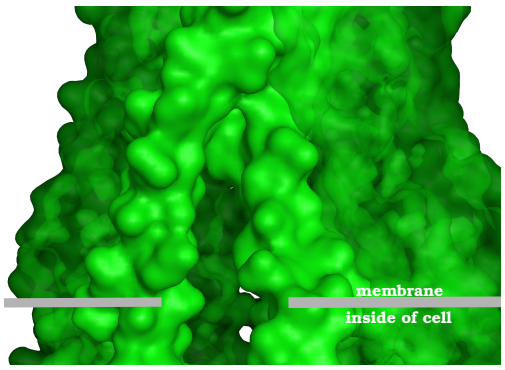
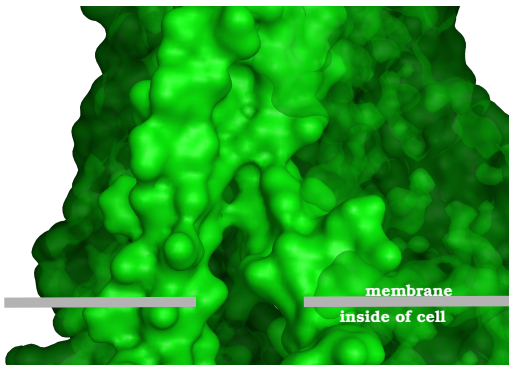


Figure 6.31: Minimum required distance of *Ala230-Ala348* (**red**) for existent TM4/TM6 portal. Inside the P-gp cavity lies **cis-flupentixol** (shown in **blue** spheres.)The other measured residues are shown in **yellow**: *Leu219-Val338* and **dark green**: *Ser237-Ala356*.



(a) P-gp surface after 35 ns. **Cis-flupentixol** inside the cavity is not visible.

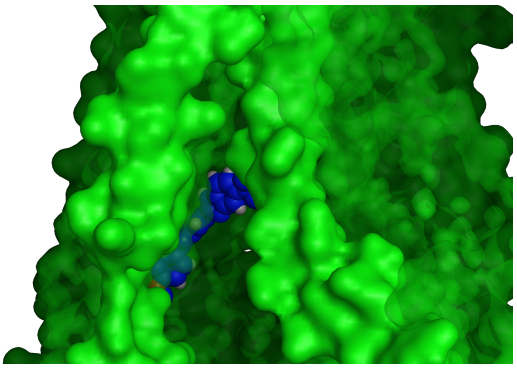


(b) P-gp suface after 35 ns.**Trans-flupentixol** inside the cavitiy is not visible.

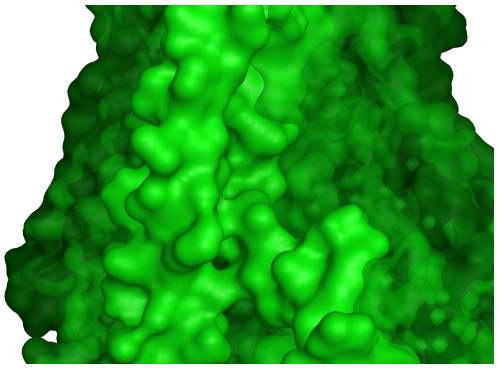
Figure 6.32: Lateral TM4/TM6 portal of P-gp.

The ligands could stabilize the protein and reduced the fluctuations of the measured distances. A differentiation of the docked ligands in inhibitor and substrate was not possible. Substrates (colchicine, daunorubicin) as well as inhibitors (elacridar, cis-flupentixol) kept the portal at its initial distance. While Hoechst 33342 opened the gate (figure 6.33a), tariquidar (figure 6.33b) and trans-flupentixol (figure 6.32b) initiated a closure.

With this measurements, a prediction of a present gate may be made, but to clarify the opening and closure of the portal further investigations are needed.



(a) Pgp surface after 35 ns. **Hoechst 33342** is located inside the cavity.



(b) P-gp surface after 35 ns. **Tariquidiar** is not visible.

Figure 6.33: Surface of P-gp. While Hoechst 33342 (shown in **blue** spheres) on the left figure, is visible through the TM4/TM6 portal, the gate of P-gp including tariquidar is closed (right figure).

Chapter 7

Principal component anylsis

Performed on a set of structures or from a MD trajectory, principal components (PCs) describe concerted atomic displacements and can therefore highlight major conformational changes between the structures. Since these motions are often essential for the function of the protein [123], the dynamics in this low-dimensional subspace are also termed “essential dynamics” [124].

The overall translational and rotational motions in the molecular dynamics trajectory were eliminated by a least squares fit superimposition onto a the starting conformation as a reference structure [125]. By using a linear transformation on a Cartesian coordinate space the configurational space can then be reconstructed into a covariance matrix. The following diagonalization of the matrix generates eigenvectors for each component, that indicate the direction of motion. Each eigenvector has a corresponding eigenvalue representing the energetic contribution to the motion. Therefore the eigenvector with the largest eigenvalue (e.g. first Principal Component) accounts for the highest variance within the data.

Conformation	production time used for PCA
NBD-open	2-35 ns
NBD-open + colchicine	10-35 ns
NBD-open + daunorubicin	20-35 ns
NBD-open + Hoechst 33342	12.5-35 ns
NBD-open + elacridar	17-35 ns
NBD-open + tariquidar	12.5-35 ns
NBD-open + cis-flupentixol	20-35 ns
NBD-open + trans-flupentixol	5-35 ns
NBD-1/4-closed	15-35 ns
NBD-1/2-closed	18-35 ns
NBD-3/4-closed	7.5-35 ns
NBD-closed	12.5-35 ns

Table 7.1: Production time. *(The equilibration and production time may be seen in section 6.2).*

To remove the models equilibration period, PC analysis was performed after monitoring the RMSD. Consequently the time frames considered as production time, were chosen differently for each conformation and ligand. The time frames resp. time that was used for PC analysis is listed in table 7.1

Since initially large-scale motions are of interest, side-chain local motions were excluded, by using only the coordinates of P-gp's Cα-atoms. This procedure did not only reduced the data analyzed into an amount that was computationally tractable, but functioned also as an initial filter, that removed the side-chain motion.

In the following sections the different structures will be analyzed and compared to the NBD-open conformation of P-gp. Therefore, at first, the NBD-open conformation and afterwards mostly the differences of the other structures are discussed.

7.1 Residue-wise loadings

7.1.1 NBD-open conformation

The resulting PC analysis plot (figure 7.1), indicated the proportion of variance against the eigenvalue rank of P-gp. The first PC accounts for more than one third of the overall variance, strongly dominating the overall variance. The first three components together contributed 64.4 % of the complete variance. Additionally, each successive component is responsible for a decreasing amount of this variance.

Projecting the trajectories on the first PCs (figure 7.1), a partial visualization of the multidimensional phase space on a plane is visible. Throughout the molecular dynamics simulation, the phase space explored by the MD increased, as the PC1-PC2 subspace (visible by the changing color from blue to white to red).

The residue-wise loadings for PC1, which contributed almost half (41.4 %) of the overall variance showed, that one nucleotide binding domain dominated this component (figure 7.2). Highest loadings were not only observed in NBD2, but also in the helices, which are in close contact to NBD2 (TM4/TM5 and TM8/TM9, respectively). The loop (ECL1), connecting H1 and H2, was also contributing to the PC1, as a result for its high flexibility.

Especially this residue contribution for PC1, resembled the overall residue fluctuations, suggesting, that the NBD2 and connecting helices contribution dominated P-gps movement.

The movement can be plotted on the NBD-open structure and therefore visualizing its direction. While in PC1 NBD1 moved apart

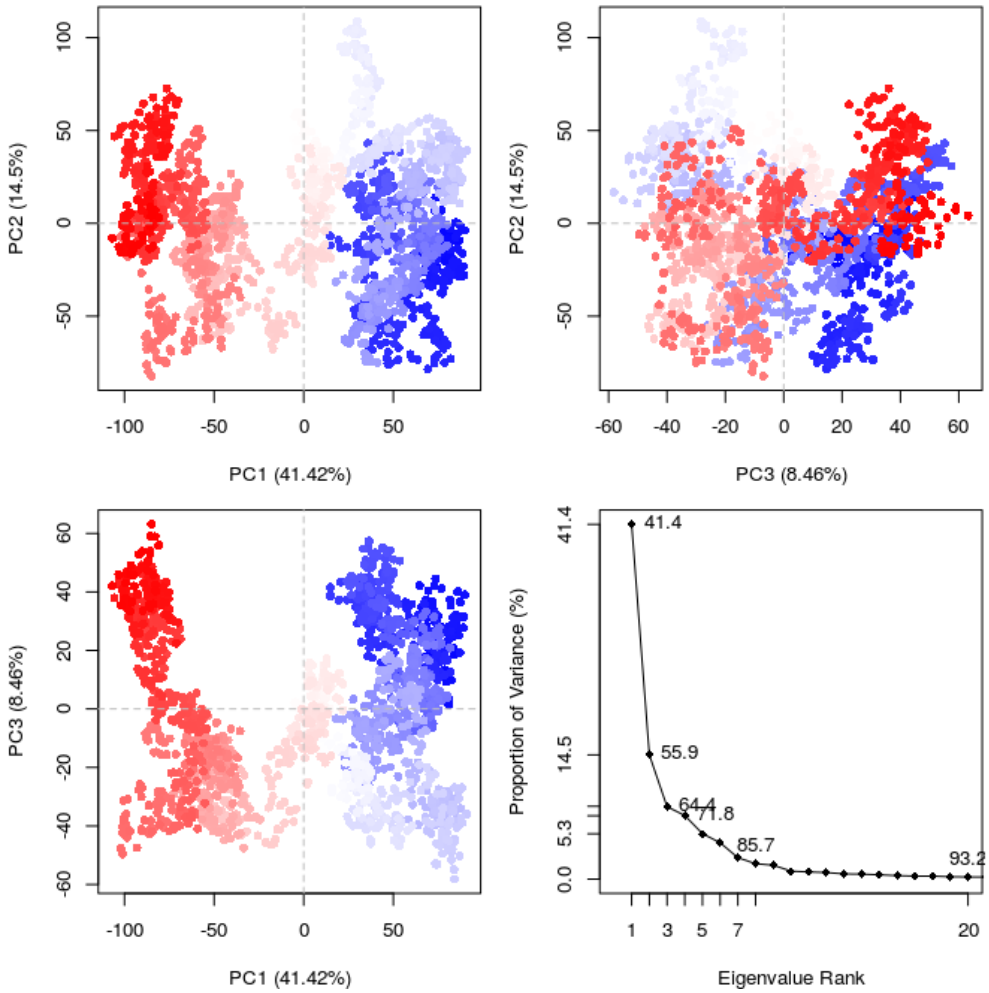


Figure 7.1: PCA NBD-open. The movement through space is shown from blue to white to red. Each dot corresponds to a different time frame of the MD simulation. Different cluster of dots correspond to different conformations.

of NBD2 in a twisting-like motion, in PC2 both NBD1 and NBD2 moved straight towards each other. NDB2 kept the twisting motion, however towards NBD1 (figure 7.3).

The residue-wise loadings for PC2 differentiated by the high contribution of NBD2 as well as NBD1 to this component. Interestingly the helices forming ICL1 and ICL2, did not show an increased contribution, as were H5 and H6 in PC1, indicating a higher independence of NBD1 of the corresponding connecting helices.

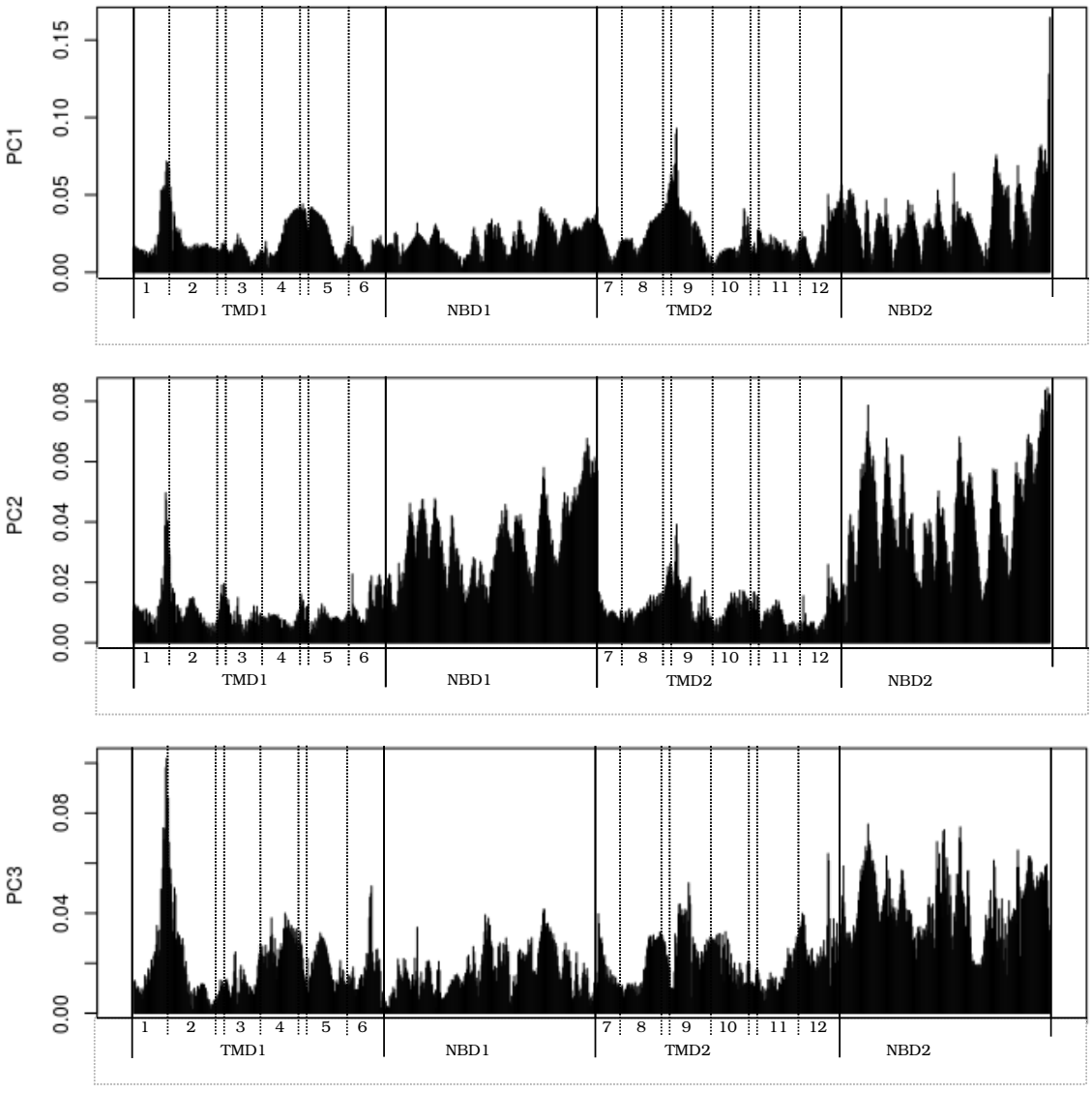
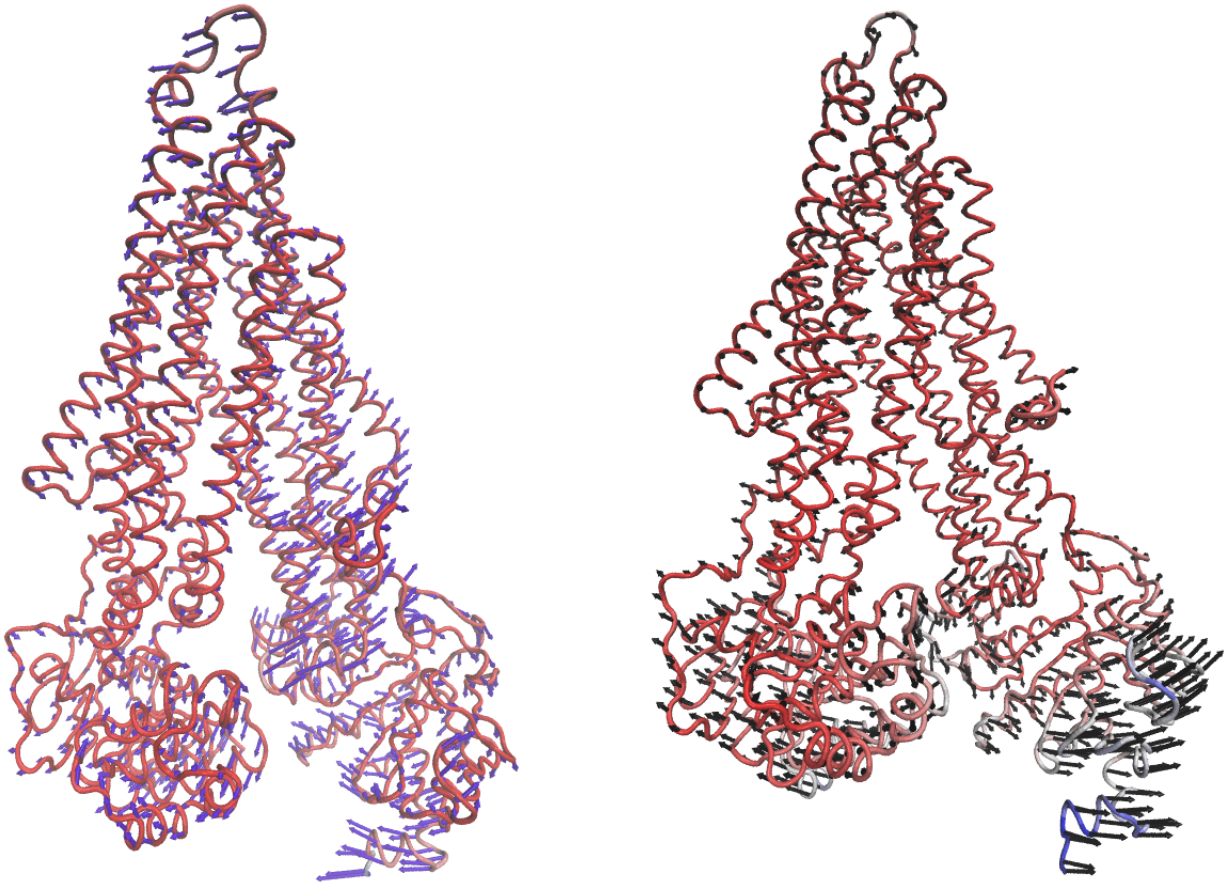


Figure 7.2: Residue-wise loading of first three PC of NBD-open



(a) PC1. **(b)** PC2.

Figure 7.3: The calculated movement of P-gp is shown by arrows.

7.1.1.1 PCA without NBDs

To investigate the influence of the NBDs to the residue-wise loading of P-gp and especially the possibility of the high contribution of the NBDs to mask helical movement, a Principal Component Analysis was performed without NBDs.

This analysis showed, that H2/H3 as well as H8/H9 and H11 showed the highest loadings for PC1 (figure 7.4). Comparing with the PCA including the NBDs, the results differ by the highest residue-wise loadings. H4/H5, as well as H8/H9 contributed most to PC1, whereas in the PCA including the NBDs H4/H5 and H8/H9 showed the highest loadings.

However, the distribution resp. helices contributing to PC1 were the same, showing that the NBDs did not mask any helical movement. No helical regions appeared in this PCA, that were not also visible as contributing parts in the PCA with the NBDs.

Helices connected by ICLs showed the highest residues wise loadings. Interestingly, in this analysis no suggestion to possible movements of the NBDs can be made, since no ICL pair (ICL1 and ICL4 vs. ICL2 and ICL4) clearly dominated PC1, while in the PC1 including the NBDs, ICL2 and ICL3 showed higher residue-wise loadings.

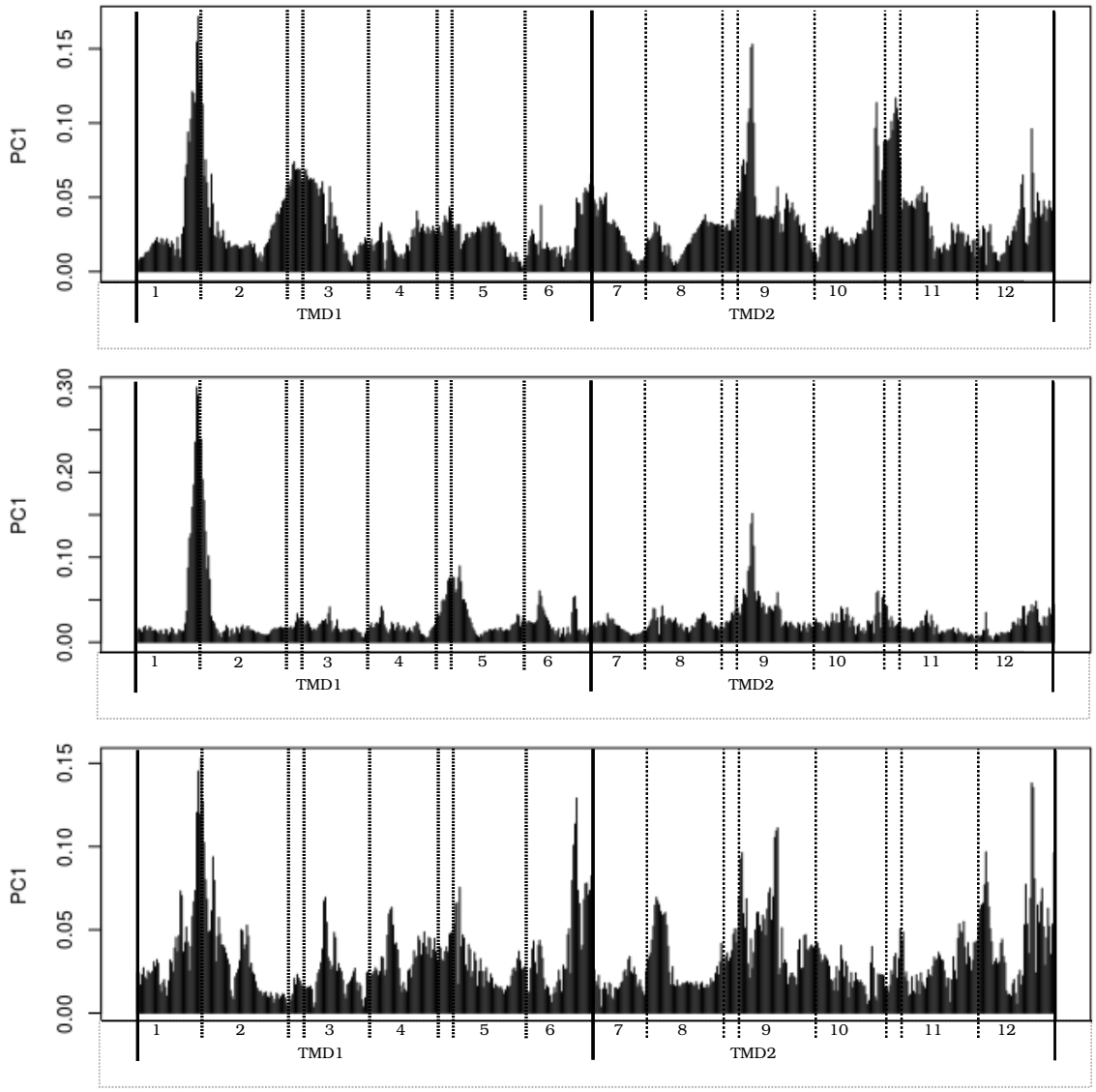


Figure 7.4: Residue-wise loading of first three PC of NBD-open

7.1.2 Morphing structures

The residue-wise loading of PC1 of the morphing structures differed especially in NBD-1/2-closed conformation, where NBD1 and NDB2 were dominating, whereas in the NBD-1/4-closed conformation the extracellular loop dominated the distribution. Although the NBDs were dominating, the ICLs, mostly of TMD1 (figure 7.6), did not show a high contribution, which is in good conformity with the constant NBD-distance. However H3 and H5, as well as H9 and H10, that are next to ICL1 and ICL3/ICL4, respectively, showed higher residue-wise loading (figure 7.5).

This contribution of the residue-wise loadings of the helices indicate the flexibility of this transition state of the transporter.

The other conformations, which are closer to an open or closed conformation, showed less flexibility in the helices and may indicate P-gp needing a ”stiffness” in the helices to move the NBDs and subsequently to change the conformations.

On the other side, H10 showed a high peak in this NBD-1/2-closed conformation and this helix had continuously in all conformations a loss of structure. In section 5.2.1 H10 is suggested to have a hinge and to facilitate P-gp's conformational change.

NBD-1/4-closed and NBD-3/4-closed, both did not appeared to show regions dominating the first principal component (figure 7.5). Especially NBD-3/4-closed appeared to show only a few motions.

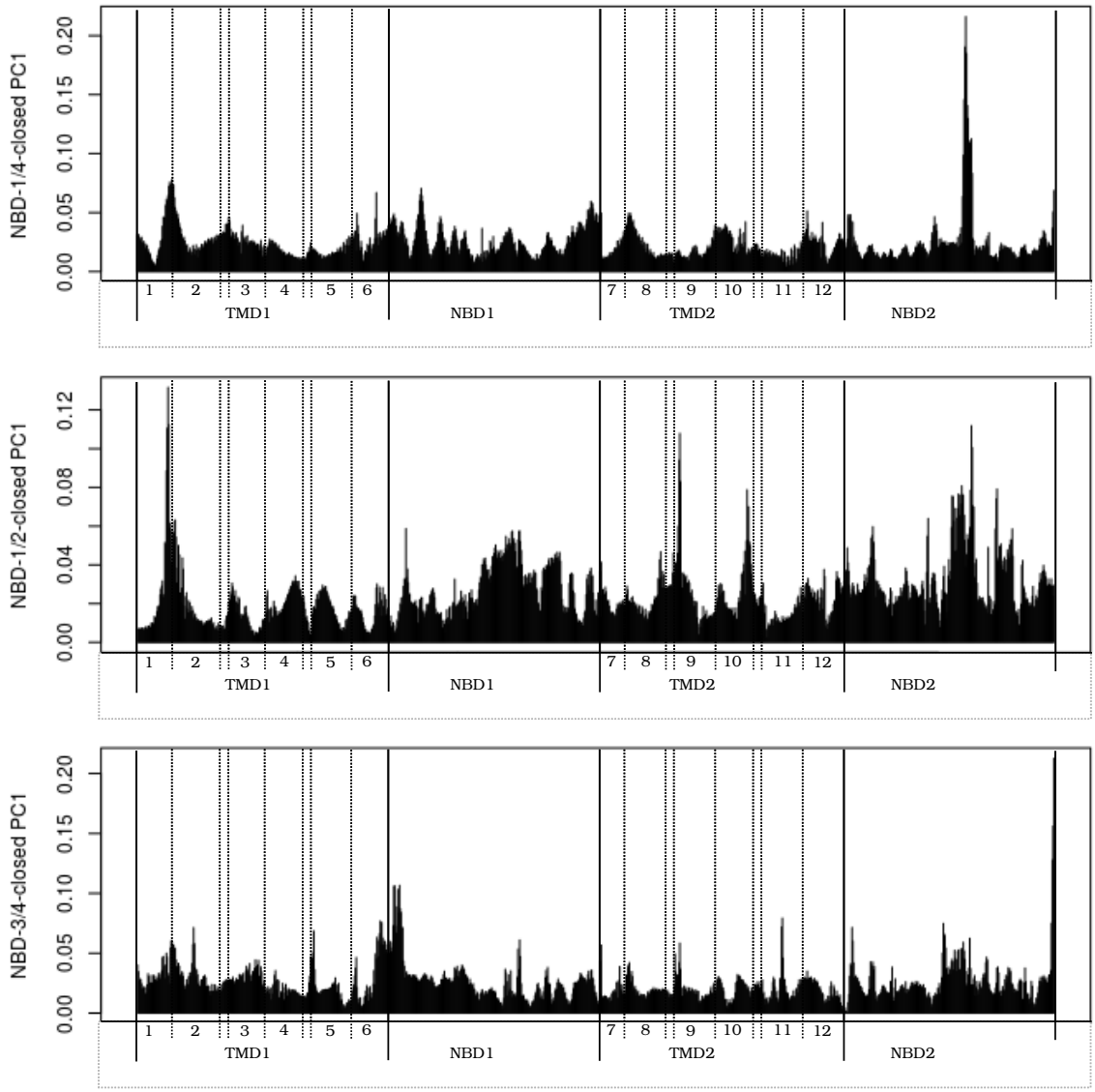
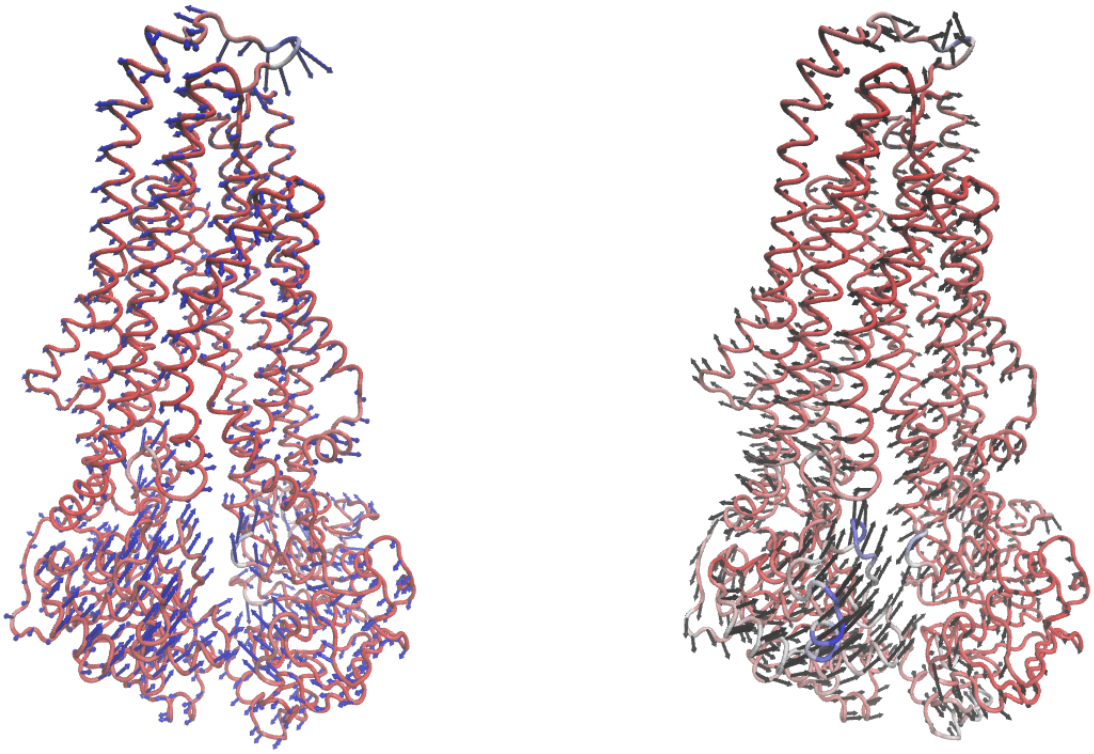


Figure 7.5: Residue-wise loadings of morphing structures.



(a) PC1. **(b)** PC2.

Figure 7.6: NBD-1/4-closed movement.

TNBD-3/4-closed and NBD-closed resembled to each other during the molecular dynamics simulations by the constant distance of the NBDs. Comparing the residue-wise loadings, however, NBD-closed showed, that NDB2 and a small part of NBD1 were dominating, next to almost all helices. Interestingly not the ICLs themselves, but rather the sequence of helices next to the intracellular loops showed a high contribution. Underlying the suggestion, that only by movement of the NBDs, the ICLs show high residue wise-loadings, and influence the transporter to change transition. This wide flexibility of the helices was comparable to the fluctuations of NBD-1/2-closed, although the contributions were not as high.

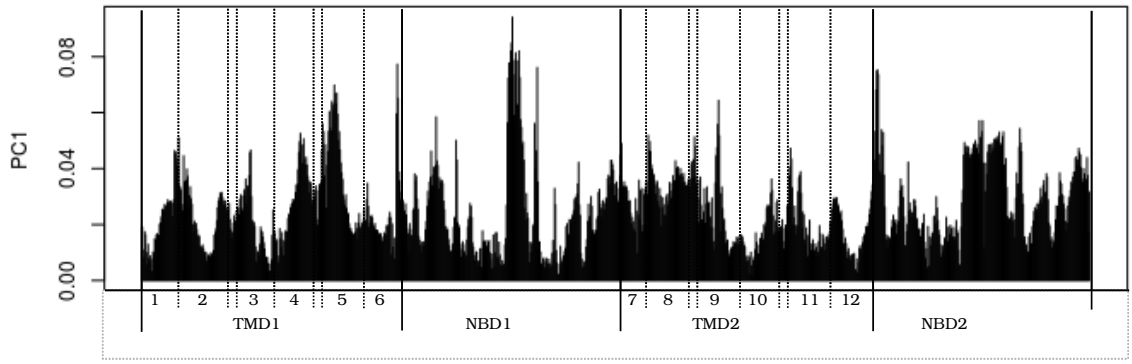


Figure 7.7: Residue-wise loadings of NBD-closed.

No connection of the helices to any movement was distinguishable in any conformation. From the distribution of the residue-wise loading, only NBD-open showed a clear pattern, indicating the tight connection of NBDs and ICLs. The other conformations, showed contributions, none comparable to an other. However, if taking into account the high contribution of helices in the conformations with no particular movement of the NBDs (e.g. NBD-closed, NBD-3/4-closed and NBD-1/2-closed), it seems that the helices show higher

fluctuations, when the protein does not move, resp. does not close or open the NBDs.

7.1.3 Substrates

Comparing the residue-wise loadings of the substrate bound to apo P-gp, the substrates influenced the contribution (figure 7.8).

Due to its movement towards the inner P-gp cavity, colchicine had interactions with TM1, TM2, TM5 and TM6 of TMD1 as well as TM7, TM11 and TM12 of TMD2. The residue-wise loadings, however, did not show to be affected by any interactions of colchicine. Only H5 showed a contribution, which may be connected to an interaction with colchicine. H9 and H10 with comparable high contributions to H5, did not had any interactions, and are not neighboring helices to helices, that were involved in drug-binding. This indicates that the network of interactions and communication is very complex.

Daunorubicin and Hoechst 33342 showed a similar NBD movement, due to their different behavior inside the binding cavity. Interestingly, Two peaks of NBD1 are similar, which may be connected to the NBD-movement. Although, NBD2 showed higher residue-wise loadings that NBD1 in the daunorubicin structure, ICL1, ICL2 and ICL3 showed a higher contribution. ICL2 and ICL3 are communicating with NBD2, but ICL1 is communicating with NBD1 and therefore this contribution is not really explainable. The contributions of H4/H5 as well as H8/H9 correspond to the movement of NBD2, since they are the connecting helices to ICL2 and ICL3, respectively. By the residue-wise loadings it is not evident

which helices were interacting with daunorubicin.

Hoechst 33342 move not to the inner P-gp cavity but stayed in a pocked defined by TM5, TM8 and TM9. Highest contribution was as with daunorubicin by NBD2, however without ICL2/ICL3 sharing this contribution. The peak of H4 is at the sequence height of H8, that was interacting with the ligand.

It was clearly visible, that all the substrates changed the contribution of the residues, however how they were affected could not be seen.

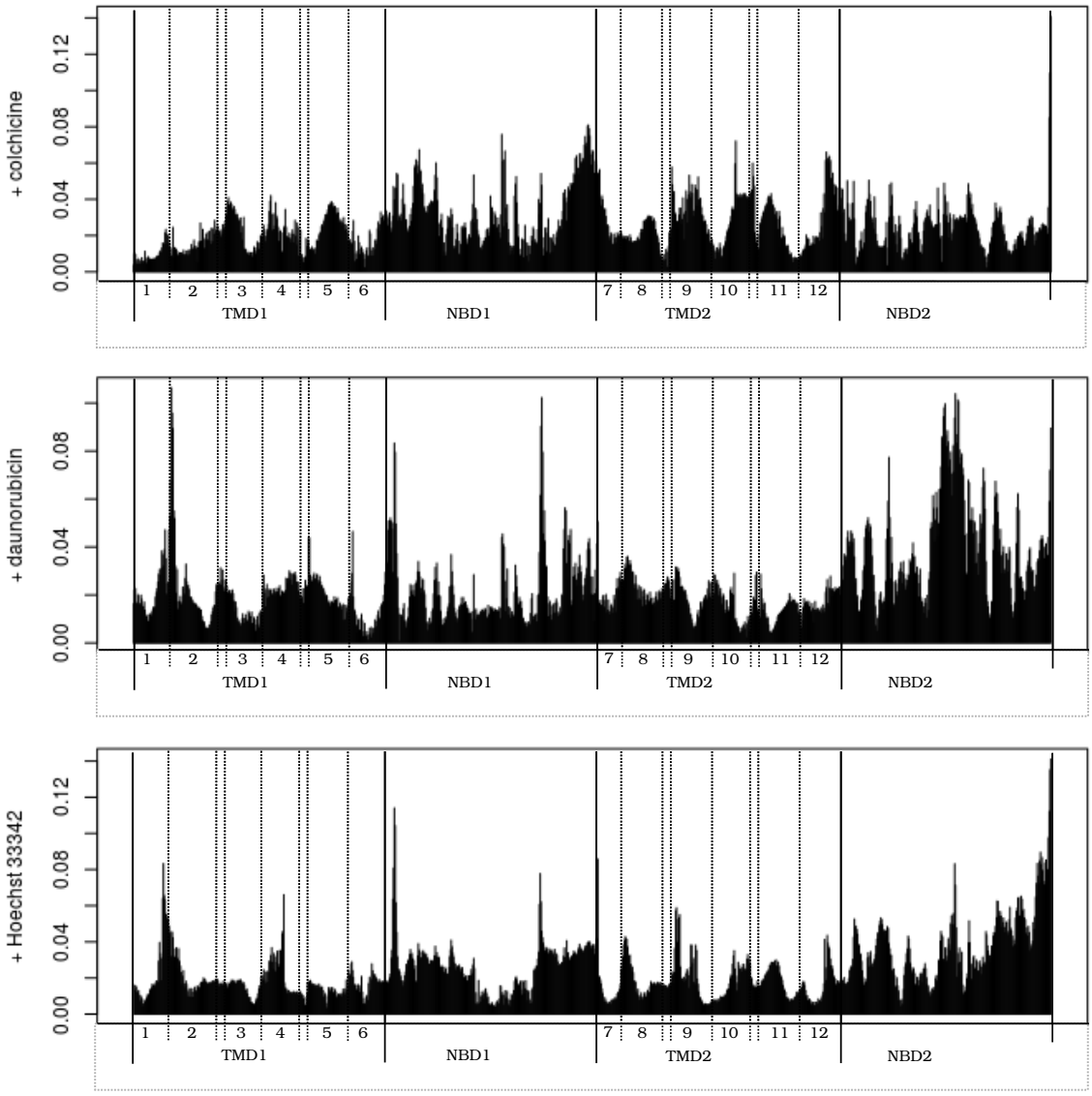


Figure 7.8: Residue-wise loadings of PC1 of the substrates.

7.1.4 Inhibitors

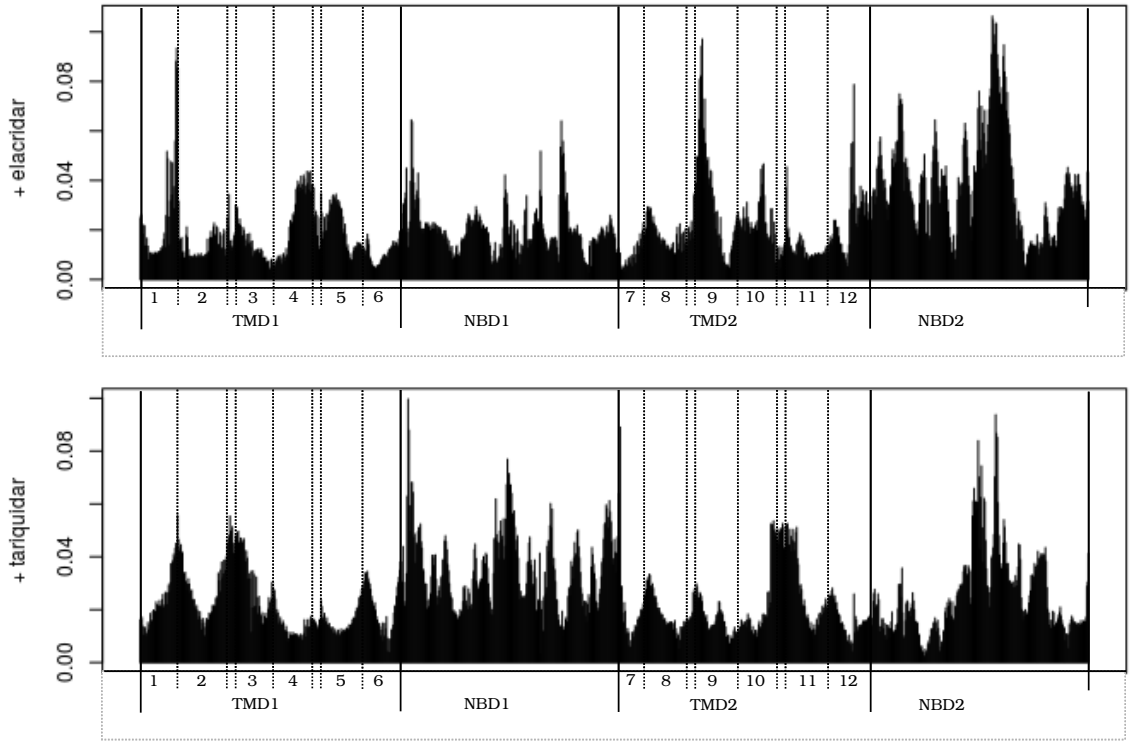


Figure 7.9: Residue-wise loadings of PC1 of elacridar and tariquidar

Likewise the substrates, the inhibitors also changed the distribution of the contributions. Besides elacridar, the inhibitors showed very high residue-wise loadings of NBD1 and the communicating ICL1 and ICL4.

The interactions of elacridar with the binding cavity of P-gp were not discernible. One of the highest residue-wise loadings was contributed by H9. This part of the helix showed a loss of structure, and therefore this contribution is a result of this helical unwinding. The high contribution of H4/H5 might be connected to the NBD2.

Tariquidar, stayed similar to Hoechst 33342 in the binding site, spanning due the its size the membrane. But unlike the substrate tariquidar showed a very high contribution of NBD1. And although NBD2 showed as well high residue-wise loadings, ICL2 and ICL3 did not appear to have a big influence.

Cis- and trans-flupentixol showed the most contrary behavior. While cis-flupentixol had very high residue-wise loadings indicating a high motion, trans-flupentixol showed only within NBD1 a high contribution, while only H4 and H5 showed an additional movement.

Like the substrates, the inhibitors changed the contribution pattern. However it was not discernible how and why the residue-wise loadings changed.

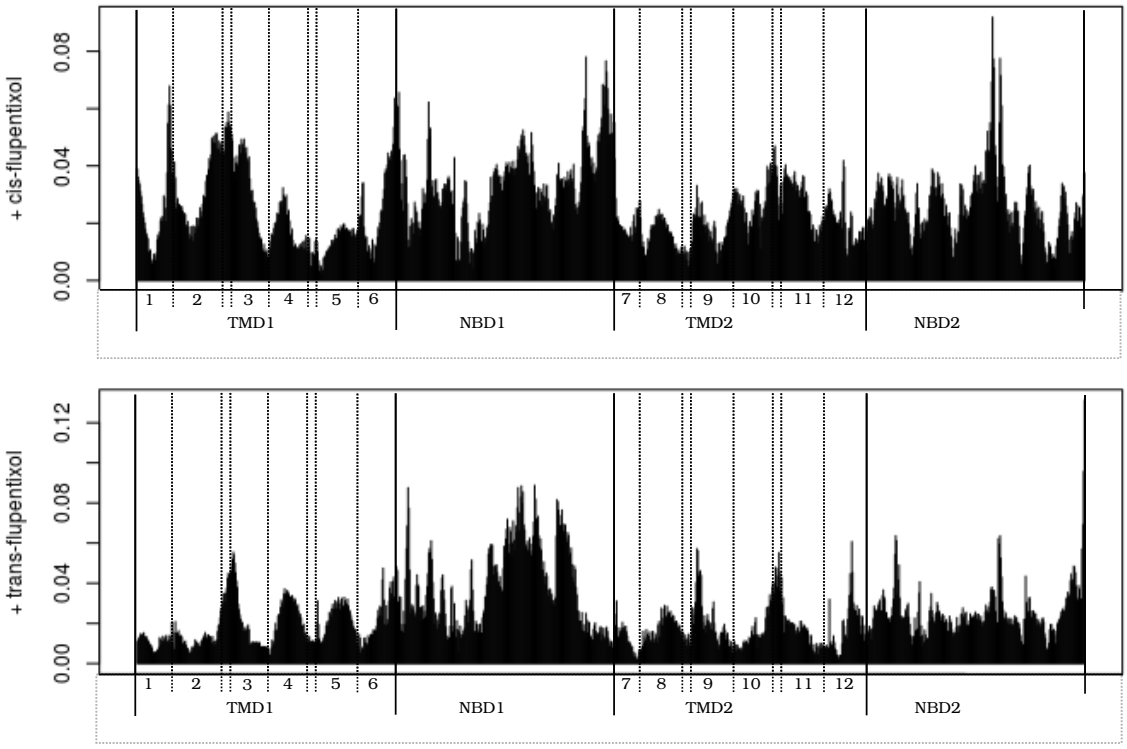


Figure 7.10: Residue-wise loadings of PC1 of cis- and trans-flupentixol

7.2 Dynamic cross-correlation maps (DCCM)

To investigate dynamic coupling between the residues within P-gp, a covariance matrix was computed and visualized as dynamic cross-correlation map (DCCM). The extent of the correlations or anti-correlations are color coded: red represents a strong correlation whereas dark blue represents a strong anti-correlation.

The DCCM of P-gp in the NBD-open conformation showed, that the secondary elements within the NBD1 were positively correlated, whereas NBD2 showed also anti-correlations (see figure **??**).

The motion within TMD1 displayed both positive and negative correlations. Especially H1 to H2 to H3, as well as H4 to H5, showed strong positive correlations between each other and anti-correlations to each other. The corresponding helices of TMD2 behaved equally to TMD1, only stronger. Namely, H7 to H8 to H9 showed very strong correlations, as did H10 and H11. Also the correlation of H7/8 and H9 to H12 was very clearly visible and was much stronger than the corresponding H1, H2 and H3 to H6.

The dynamic coupling of the two transmembrane domains, showed a coupled movement of H4/H5 to H7, H8 and H9. These five helices are connected by ICL2 and ICL3, respectively, and therefore are the communication to NBD2. Their stronger movement in comparison to the dynamic coupling of H1, H2 and H3 to H10 and H11, connected by ICL1 and ICL4, respectively, may be explained by the stronger movement of NBD2.

The divided correlation of H12 including correlations and anti-correlations, can be explained by H12 having a part of the helix, inside the membrane, interacting with the other helices, and the second part, outside the membrane, being the physical link to NBD2.

In general, this correlation pattern seen in the NBD-open apo structure was present in all dynamic cross-correlation matrices:

- **a correlation between H1 and H2**
- **a correlation between H2 and H3, as well as between H6 and H10/H11:** H2 and H3 are connected by ICL1 and H10 and H11 by ICL4, respectively. Subsequently these four helices are all together on one side of P-gp and interact with NBD1. The additional correlation of H2/H3 to H6 may be explained, that they are adjacent to each other.

- **a correlation between H4 and H5, as well as between H8/H9 and H12*:** H4 and H5 are connected by ICL2, and H8 and H9 are connected by ICL3 to each other. Likewise the above mentioned helices, these helices are on the other side of P-gp and communicate with NBD2. The strong correlation to TM12 may result form TM12 lying adjacent to H8/H9. However, comparing these correlations of TM12 to the ones of TM6 and the "NBD1 helices", they are much stronger. Also H4/H5, that are not adjacent to H12 showed very strong correlations to TM12.

- **a correlation between H7 and H8/H9, as well as between H8/H9 and H12:** H7 lies in between H8 and H9, therefore the correlations are strong.

- **a correlation between H8 and H9, as well as between H8 and H12:** H8 and H9 are connected by ICL3 and are communicating with NBD2. Also TM12 lies adjacent to these two helices.

- **a correlation of H10 to H11:** These two helices are connected by ICL4.

- **an anti-correlation between H1/H2 and H4/H5, as well as between H7/H8 and H12:** These anti-correlations are to the part of P-gp that interact with NBD2. The strong interaction between one side of P-gp (NBD1-side) result into anti-correlations to the "other side" (NBD2-side) of P-gp

- **an anti-correlation between H4/H5 and H6 as well between H4/H5 and H11:** Here too, H4 and H5 communicating by ICL2 with NBD2, did show anti-correlations to H6, being

the connection to NBD1 and H11, that communicates by ICL4 to NBD2.

- **an anti-correlation of H6 to TMD2**

* Only the NBD-open including colchicine structure did not show a correlation between H4/H5 and H12

Summarizing (table 7.2), P-gp showed a coupled movement of the helices communicating with the NBDs and lying adjacent to each other. H2/H3 move along with H10/H11, as well as with H6. All five helices are connected or communicate by ICLs to NBD1. However, although the helices interacting/connected to NBD2, namely H4/H5, H8/H9 as well as H12 did show also a coupled movement, they did not anti-correlate to each other. H6 showed anti-correlations to all helices of TMD2. Whereas only H1/H2 were anti-correlating to H7/H8 as well as to H4/H5 and H4/H5 to H11.

In the next section differences of the structures are mentioned.

	Helices											
	1	2	3	4	5	6	7	8	9	10	11	12
1		+		-	-		-	-				-
2	+		+	-	-					+	+	-
3		+				+				+	+	
4	-	-			+	-		+	+		-	+*
5	-	-		+		-		+	+		-	+*
6		+	+	-	-		-	-	-	-	-	-
7	-	-				-		+	+	-	-	+
8				+	+	-	+		+			+
9				+	+	-	+	+				+
10		+	+			-	-				+	
11		+	+	-	-	-	-			+		
12	-	-		+*	+*	-	+	+	+			

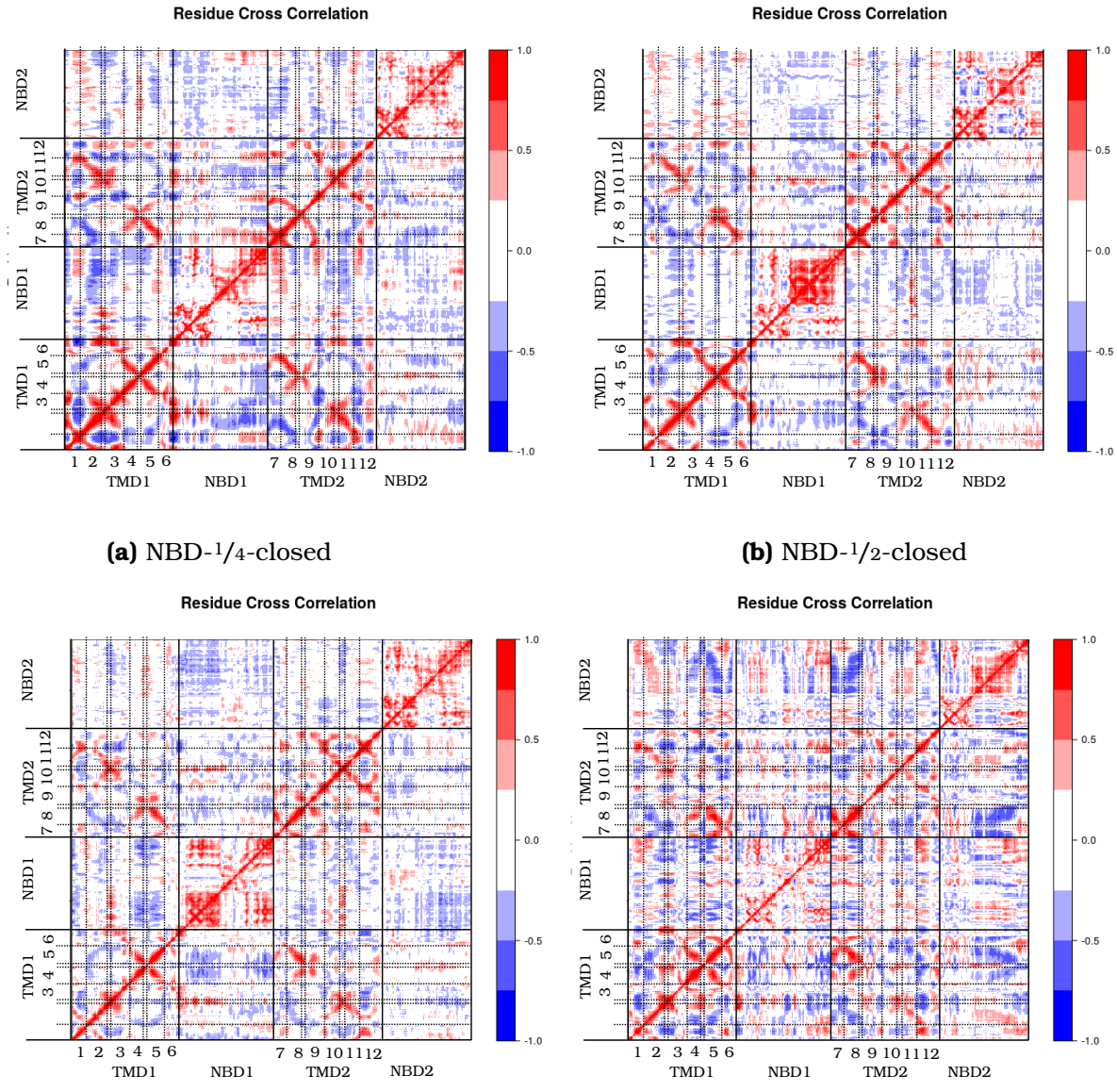
Table 7.2: Cross-correlation pattern. + correlation, - anti-correlation. * Only the NBD-open including colchicine structure did not show a correlation between H4/H5 to H12

7.2.1 Morphing structures

Overall, the correlations seen in the morphing structures were not as strong as in the NBD-open conformation. The strong correlations of NBD-open ICLs helices, where the whole helix was included in the movement, was not present in the morphing conformations, where mostly the ICLs themselves with the part of the helix close to the ICL showed coupled movement and the other part of the helices showed mostly an movement with no correlation.

The coupled movement of H2/H3 (ICL1) to H6 increased in the NBD-$^1/_4$-closed structure, whereas the other morphing structures did not show any effects. Also a strong anti-correlation was present between ICL1 and ECL1. Additionally an anti-correlation to the upper part of H12 and a correlated movement with the lower part of

H12 was visible. This is contradicting the coupled motion of the two parts of P-gp, of which ICL1 and ICL4 have strong correlations to NBD1 and anti-correlations to ICL2/ICL3 and NBD2, and vice versa. Although, the production run time from 15 ns on wards was taken, NBD-$^1/_4$-closed showed still fluctuations in the NBD distances (section 5.1). This fluctuation may result to this coupled movement. However NBD-open also fluctuated, and did not show this coupled motion.



(a) NBD-1/4-closed

(b) NBD-1/2-closed

(c) NBD-3/4-closed

(d) NBD-closed

Figure 7.11: Cross-correlation maps of P-gp and morphing structures.

Changes were mostly within the NBDs (best visible NBD1 of NBD-$^1/_4$-closed and NBD-$^3/_4$-closed, figure 7.11b and d). However ICL1 changed the motion from an independent, or correlated motion in NBD-open to an anti-correlated motion in NBD-$^1/_4$-closed, NBD-$^1/_2$-closed and NBD-closed conformations. Within the nucleotide binding domains the coupled movement dominates, whereas, the domains to each other have an anti-correlation.

The cross-linking measurements, showed that from NBD-$^1/_2$-closed towards the NBD-closed conformation, P-gp is in the ATPase active state, however no indices were seen in the dynamic cross-correlation maps.

NBD-open differentiated from the other conformations by NBD1 showing fragmented correlation/anti-correlation pattern.

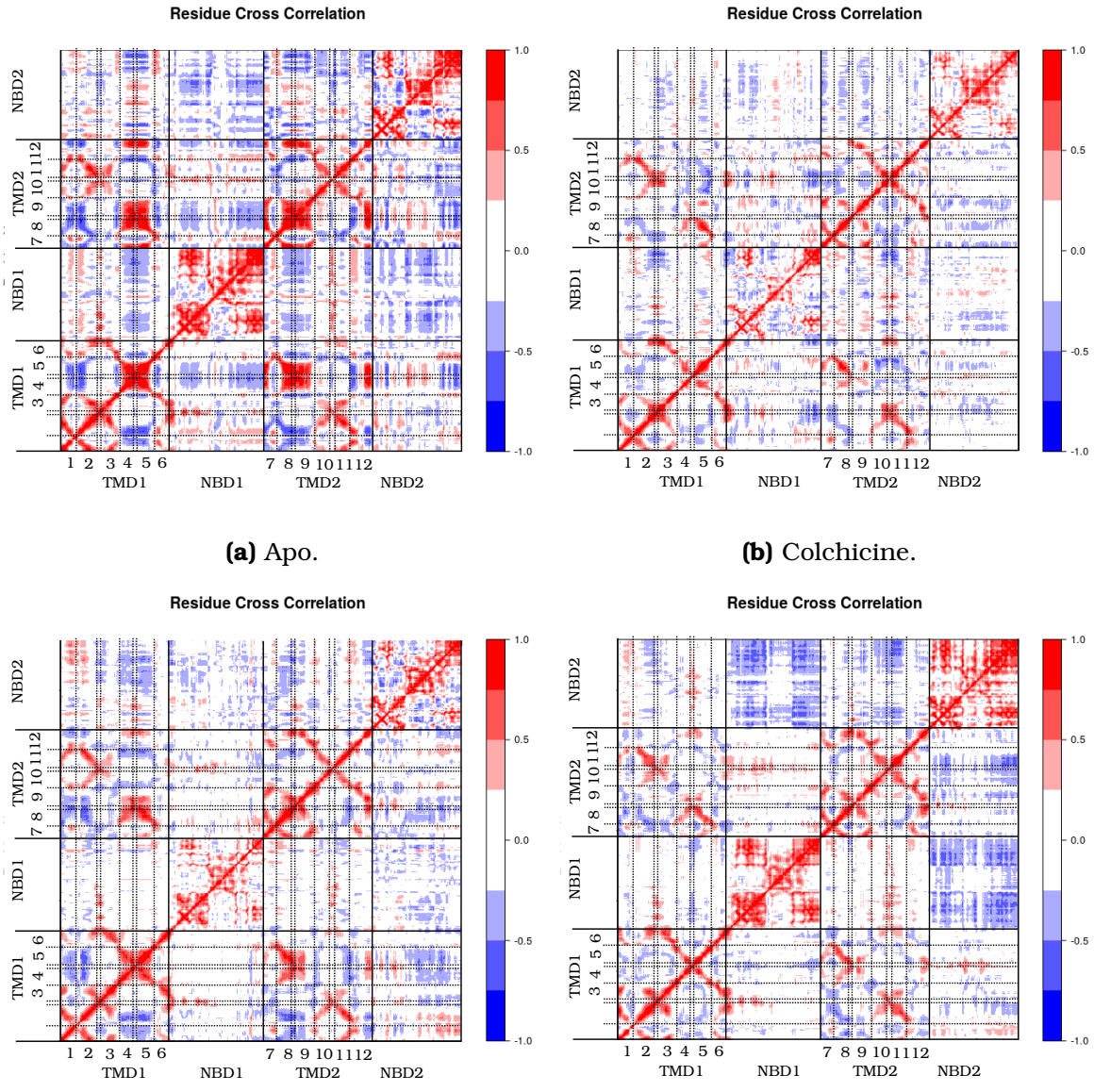
All four ICLs showed coupled motion to each other. This may be explained by P-gp having adapted in its conformation with the NBDs in close proximity. Since, the transporter does not need to show any movement anymore, the ICLs harmonize. Also a correlated movement showed the helices H2/H3 to H4/H5 and H9/H10, that are connected to ICLs. The other helices, did neither show a correlation nor anti-correlated motion.

7.2.2 Substrates

The correlation pattern of NBD-open with colchicine in the binding pocket shows the positive coupling dominating TMD1, while the apo-structure had also strong anti-correlations between H4/H5 and H1/H2. Colchicine had interactions with TM4 and TM5, and therefore this interactions seemed to reduce this anti-coupling. This reduction of anti-correlation was also visible in TMD2, namely H8/H9, to which colchicine had as well interactions to. Especially H9, did not appear to have as much correlations to TMD1 (besides to H5) as it had in the apo structure (figure 7.12a and b).

Another general observation is that NBD2 did not show as strong coupled movement to itself, as it did in apo P-gp Therefore, it can be said, that colchicine did not change the inter domain residue coupling (e.g. from a positive correlated coupling to a negative anti-correlated coupling), but it reduced the intensity of the correlations and anti-correlations, respectively.

Daunorubicin, kept in contrast to colchicine the anti-correlations of H1/H2 to H4/H5 as well as to H7/H8. Comparing both cross-correlation matrices of apo P-gp and P-gp including daunorubicin they did not show much difference, except the correlations, independently positive or negative, were not as strong as in apo P-gp. While NBD1 did not show any correlation or anti-correlation, NBD2 decreased the coupled movement to itself (figure 7.12c).



(a) Apo.

(b) Colchicine.

(c) Daunorubicin.

(d) Hoechst 33342.

Figure 7.12: Cross-correlation maps of P-gp and substrates

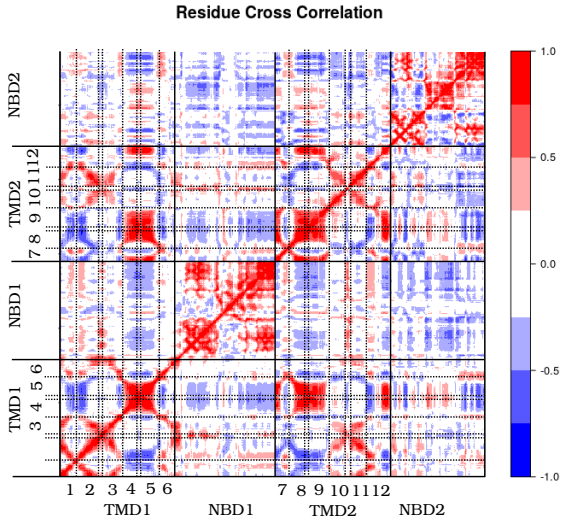
Hoechst 33342 was the only ligand, that moved towards the transmembrane helices, instead into the inner P-gp cavity. And although this movement was different from the other substrates the cross-correlation pattern did not change much. The movement towards the membrane, decreased the intensity of the correlations and the anti-correlations of H5 to H1/H2 as well as the anti-correlation to H11 disappeared. This may be a result form Hoechst 33342 being sandwiched by TM5/8/9, and therefore changing the movement of H5. Comparing the results of H8/H9 to the ones of colchicine, which did not interact with these two helices (section 6.3.1), they look similar, indicating that Hoechst 33342 interaction with H8/H9 did not had influence on the movement. However, the anti-correlations of NBD2 to itself changed to a coupled movement, indicating, that the communication of ICL3 to NBD2 changed and influenced NBD2 (figure 7.12d).

7.2.3 Inhibitors

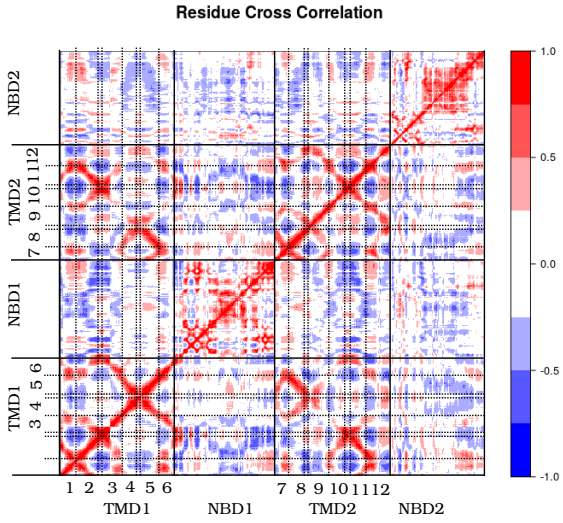
In contrast to the substrates, the inhibitors did not show a decrease of the intensity.

Although tariquidar had only interactions with H11 (section6.3.5), the inhibitor changed the correlation between H1/H2 and H2/H3 to strong anti-correlations (ICL1). These change was visible also at the corresponding H10/H11 (ICL4). An additional coupled movement to H12 appeared. However, H1 and H2 showed strong correlations to the parts of H6 and H12, to which tariquidar formed H-bonds, and therefore may additionally influenced the communication of the helices. This strong coupling did not appeared in the apo- and substrate bound P-gp. The interactions of tariquidar and H8, may also influenced H8/H9 to show a coupled movement to H10/H11. These helices, lie across the inner P-gp cavity and showed no/anti-correlations in the apo and substrate bound P-gp. This phenomenon may also be induced by the H-bonds of tariquidar to H6 and to H12, and following the connection of both P-gp sides.

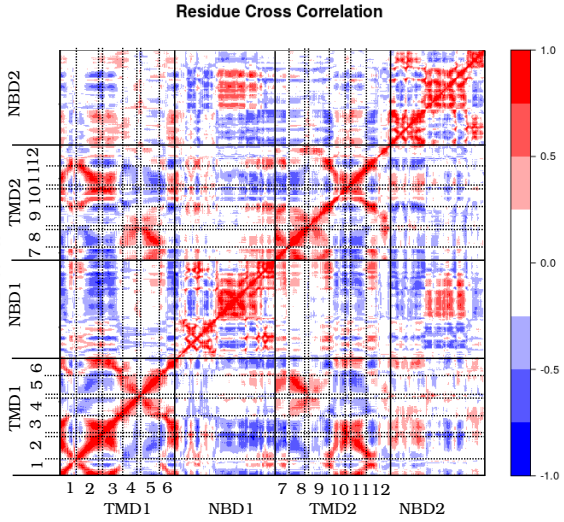
Interestingly, the coupling of the domains were stronger than in substrate bound P-gp, but not as strong as in the apo structure.



(a) Apo.



(b) Tariquidar.



(c) Cis-flupentixol.

Figure 7.13: Residue cross correlation of inhibitors.

Cis-flupentixol had the most dramatic impact of all substrates and inhibitors. The intensity of the correlations and anti-correlations was as strong in as in apo P-gp (figures 7.12 and 7.13). The coupled movement of H2 and H3 with H10/H11 was stronger, as well as the correlations with H6. Further, the anti-correlations between H4/H5 and H8/H9 increased. Also the anti-correlations between H4/H5 and H10/H11 intensified, as did H8/H9. Additionally, anti-correlations appeared between H2/H3 and H7/H8/H9, showing that this movement was not independent anymore.

Noteworthy is the drastic anti-correlation between H2/H3 and NBD1. Except for tariquidar this anti-correlations did not appear in any other P-gp conformation or ligand structure. Tariquidar changed the moving pattern of NBD1 and TMD1. The mostly not-correlating movement between NBD1 and H2/H3 changed to an anti-correlation, where as, the anti-correlated movement between H4/H5 and NBD1 became non correlating. Also within the domain strong anti-correlations appeared. Interestingly NBD2 did not show any big changes in movement.

Although the correlations and anti-correlations of cis-flupentixol were stronger, the general pattern did not change. While a correlated movement between TMD1 and NBD1 is visible, an anti-correlated movement between TMD1/NBD1 and TMD2/NBD2 exists.

The correlations and anti-correlations between helices of elacridar and trans-flupentixol did not show any differences or changes and therefore, are not explicitly mentioned (DCCMs are shown in appendix C).

Chapter 8

Conclusions

The major reason for failure of chemotherapy in cancer is multi-drug resistance. P-glycoprotein is one prominent member of the ATP-dependent membrane transport proteins, that pumps a broad range of substrates out of tumor cells in ATP-dependent mechanism. Therefore its broad specificity has been the subject of numerous attempts to inhibit the protein and restore the efficacy of anti-cancer drugs. However the crystal structure of human P-gp has still not yet been resolved.

In this thesis a homology model of human P-gp was generated using the crystal structure of murine P-gp as template, in order to proceed with looking into flexibility behavior by computational methods.

Additional P-gp conformations were created by a morphing process. Of 18 generated structures, three were taken for further proceedings. The taken P-gp conformations were all in the good quality (Ramachandran plot) comparable to the murine P-gp crystal structure. Also, transmembrane segments, which were reported by

mutagenesis studies to have distinct hydrophobic and hydrophilic faces, were successfully validated.
The generated P-gp conformations were subjected to MD simulations both in a ligand-free state and in complex with colchicine as substrate. A lipid bilayer and water molecules were added to realistically simulate the membrane environment of the receptor.

The tendency for a NBD closure of P-gp, which plays an important role in the transport mechanism, was dependent of the presence of the substrate colchicine, which promoted this closure. This results help on further refining the decisive cause of the transition mechanism of P-glycoprotein.

In addition three putative hinge regions of P-gp were characterized and the movement of helices was analyzed. In detail, H2 and H3, showed a movement, which was separated from the putative hinge level into two directions. Their "lower" part, which is connected to the ICL1, moved with the NBD1, while their "upper" part moved away from another and enabling a dislocation of ECL1 and opening of the outer P-gp pore. H4 was influenced by the substrate colchicine, as was H3. By investigating the coupled movement of H2 to H11, the independence of the substrate was detected, which led to the conclusion, that the movement of H2 is tight coupled with the movement of H3. Also, by the rearrangement of the helices, it was observed, that the conformational changes of P-gp occurred from the mid-membrane level to the intra-cellular space, while the upper region of P-gp remained with only a minimum of movement.

As P-gp is a very flexible protein it undergoes large conformational changes as part of its function, and by determining hinges, and the movement of helices details of the export cycle were deter-

mined.

Distances of residues which are located at the lateral portal formed by TM4/TM6 were analyzed to determine the existence and the influence of a substrate to his portal, which exposes the drug-binding site to the outer membrane leaflet. It was shown that colchicine stabilized the portal, while the ligand-free conformations attempted to ”close” it. This results, guide to a further understanding of the potential drug entry mechanism of P-gp and the influence by a substrate as colchicine.

Distances of residues were compared with known cross-linking studies, to investigate the reliability of the homology model and the morphing structures. Depending on the location of the measurements the different distances of the structures corresponded very well to the cross-linking studies. It could be shown, that colchicine in the binding site induced a close proximity of Walker A to LSGGQ sites, which correlated very well with the cross-linking studies.

Especially the P-gp conformation, which leads to an activation of the ATPase activity of P-gp was shown. It could be confirmed that the substrate colchicine influenced positively this ATPase switch of P-glycoprotein.

In additional simulations, substrates as Hoechst 33343 and daunorubicin, as well as inhibitors (elacridar, tariquidar and flupentixol) were docked into the putative H-binding site of the NBD-open conformation to investigate the impact of the ligands on the transporter, and if P-gp differentiates between substrates and inhibitors.

The simulations, indicated that substrates and inhibitors influenced P-gp to close the NBDs. It was not possible to make an

distinction between inhibitor and substrate.
Further, the ligand movement in P-gp was monitored to investigate the mobility within the binding pocket. A differentiation in the mobility between substrate and inhibitor was here as well not possible, but a small ligand as colchicine showed a higher fluctuation in the binding pocket, than a bigger and bulky ligand as daunorubicin.

To investigate the cause of the mobility and the movement direction of the ligands, the drug-binding site was investigated.

Colchicine moved twice towards the inner P-gp cavity, while once towards the lateral portal of TM4/TM5, where it was hold tightly. Daunorubicin, which moved also towards the inner P-gp cavity showed similar interactions as colchicne. Due to two H-bonds holding daunorubicin into position, it did not moved as far into the binding-site. On the other side, Hoechst 33342 moved towards the transmembrane helices, spanning the inner bilayer, implying it's surface interacting with TM8/TM9 to be the reason.

Elacridar moved likewise daunorubicin and colchicine towards the inner P-gp cavity, but being hold by two H-bonding to the acridine moiety. Elacridar showed in the residues interacting crossovers with Hoechst 33342 and elacridar, due to the similar docking location and the size. Cis-flupentixol as well as trans-flupentixol moved during the MD towards the outside of the cell. Correlating to its higher RMSD, trans-flupentixol covered a farther distance.

P-gp seemed not to have distinct binding pockets, and the ligands seemed to adjust themselves during the simulation. However it could be shown that some residues play a crucial role in drug-binding, since they were present in almost all ligand interactions. These residues are: *Met69* (TM1), *Phe303, Tyr310* (TM5), *Gln725*

(TM7) and *Phe983* (TM12).

The interaction were mostly hydrophobic with very few H-bonds. A differentiation of the ligands to substrate or inhibitor was not possible, however, due to the size of the ligand (e.g. Hoechst 33342 and tariquidar vs. colchicine and flupentixol) at trend of the movement and final position was observed.

In measurements of TM4 to TM6, the influence of the lingands to the lateral portal of P-gp, was investigated. As with colchicine, the ligands stabilized the protein and reduced the fluctuations of TM4 and TM6. One substrate (Hoechst 33342) opened the gate, while two inhibitors (tariquidar and trans-flupentixol) closed the portal. Although the initiation of an opening or closure could not be determined, a prediction of an existing portal by distances of *Ala230* and *Ala348* could be made. Knowing the distances of needed for the lateral portal, help to understand the potential drug entry into P-gp.

Finally, Principal Component Analysis were performed. It was shown, that P-gps dominant motions were covered by the first component. The residue-wise loading, indicated the coupled motion of the ICLs, their connecting helices and the NBDs. The conformations and the ligands showed different motions. A twisting motion of the NBDs was visible. Also the flexibility of the helices was depending on the motion of the NBDs. High residue-wise loading of the NBDs resulted mostly in low residue-wise loadings of helices.

Especially the DCCMs of P-gp showed, how complex the transporter is. A pattern of the helices and the NBDs could be shown. A correlated movement was visible in all conformations and ligands. Correlations between H1 and H2/H3 (incl. ICL1), as well

as to H6 and H10/H11 (incl. ICL4) were visible. Additionally a correlation between H4/H5 (incl. ICL2) and H8/H9 (incl. ICL3) as well anti-correlations of these helices to H1/H2. The conformations and the ligands differed in the intensity of these correlations/anti-correlations to each other, but it was not possible to distinguish the reasons/causes of the different movements.

This thesis represents the first computational design attempt for different P-gp conformations, generated by a morphing procedure. The conformations were analyzed and results indicating possible motions for the transition of P-gp from one conformation to the other were presented. Additionally, ligand interactions and influences to P-gp high affinity structure were shown.

All results reported here could be achieved only by involving a combination of different techniques. This proves that there is no universal method valid for studying and understanding the complex mechanism of an ABC-transporter like P-glycoprotein and that there is still potential of investigating this fascinating protein. However, it also shows the limitation of computational capacity by this large system size required for MD of P-gp, which is a critical limitation in the use of computationally intensive sampling simulations and the dependence of computational power.

8.1 Outlook

Future work will focus on

- The influence of the substrate colchicine in an atomic level
- The influence of substrates and inhibitors on the other P-gp conformations, with focus on helical rearrangements and lateral portal of TM4/TM6
- The prolongation of the molecular dynamics simulations in to a 100 ns range
- Principal component analysis of single domains (TMD1, TMD2 etc.)

Chapter 9

Crystal structure of human P-gp

Shortly after completion of this thesis, within one month a high resolution electron microscopy structure of a human-mouse chimeric P-gp with the third generation inhibitor zosuquidar (PDB ID: 6FN1, 6FN4) by Alam et. al. [126] and a human P-gp in the NBD-open conformation with bound ATP by Kum. et al. [127] were published.

For this thesis a murine crystal structure had been used to generate a homology model of P-gp. To check the validity of the model the 3D structures were compared. The RMSD between the model and the human-mouse chimeric resolved structure were calculated to be 5.34 Å.

This is a relatively good fit suggesting a high degree of similarity between the two structures. The largest differences are the flexible ECL1 and the NBDs, which appear to be closer to another than in the homology model (figure 9.1).

TM4 and TM10 showed also large differences. These two transmembrane helices, were modeled as straight helices, correspond-

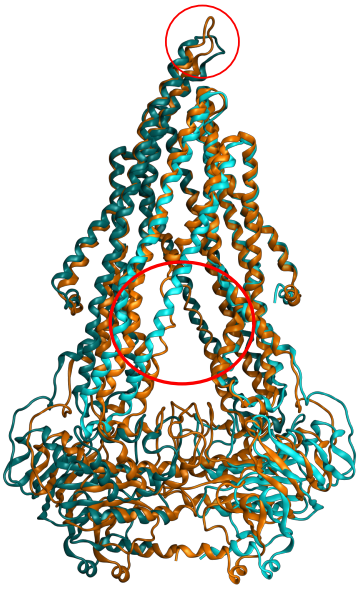
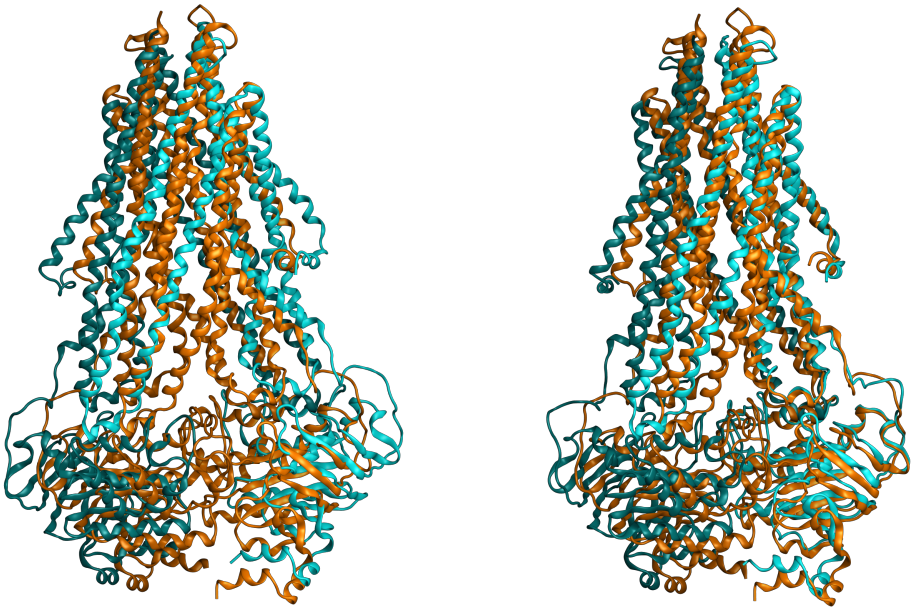


Figure 9.1: Superposition of PDB ID: 6FN1 (orange) and the homology model (blue/cyan). The red circle shows the biggest differences.

ing to the murine crystal structure. In the human-mouse chimeric P-gp they appear to have a kink into the inner P-gp cavity. A similar kink of TM4 was already seen in PDB-ID: 4Q9I (section 5.4.3). TM4 appeared in molecular dynamic simulation to show no evidence of a kink, however TM10 was the only helix, that showed in all MDs a loss of helical structure. In this sequence the human-mouse chimeric P-gp shows a kink.

The RMSD of the homology model and the human crystal structure was calculated to be relatively high at 8.4 Å. The main cause are the NBDs, showing a closer distance. Therefore the PDB-ID 6C0V was compared to the NBD-$^1/_2$-closed conformation, which had a similar NBD distance. Here the RMSD value was 4.7 Å, showing a high similarity to the solved crystal structure (figure 9.2).

The differences occur in ICL1 and ICL4 and their helical connections lying in the intracellular space. NBD1 has a slightly different orientation, resulting to a different conformation of the ICLs and the helices.



(a) NBD-open **(b)** NBD-$^1/_2$-closed

Figure 9.2: Superposition of PDB ID: 6C0V (orange) and the homology model (blue/cyan).

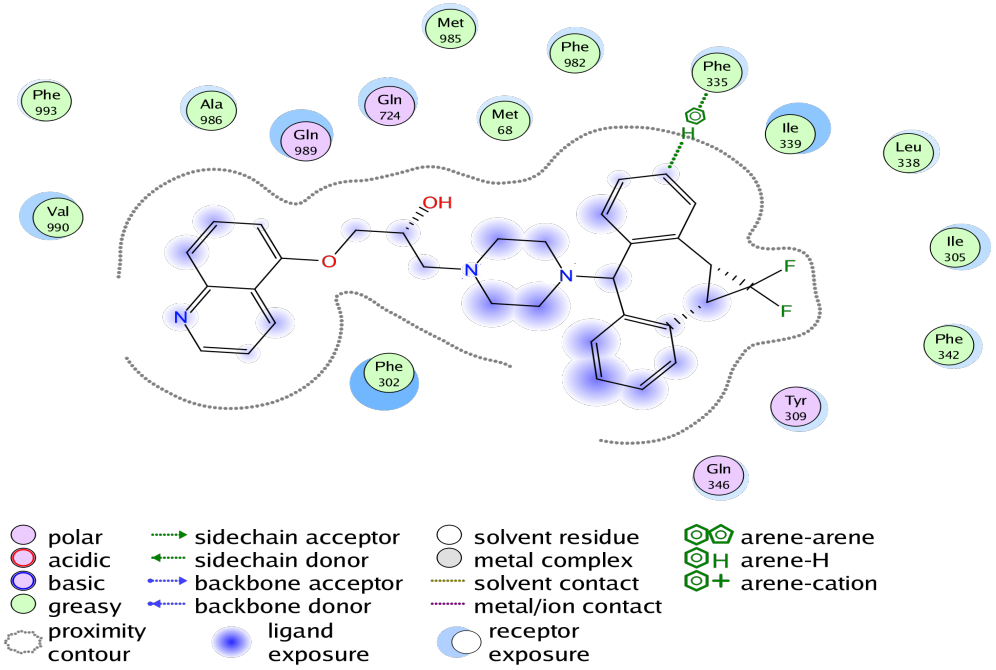


Figure 9.3: Ligand interaction in binding cavity of the inhibitor zosuquinidar and the mouse-human chimeric P-gp.

The co-crystallized inhibitor zosuquinidar showed interactions with 16 residues (figure 9.3), of which all were identified in this thesis to have interactions with a ligand, which supports the identified drug-binding cavity.

Appendix A

List of abbreviations

3D Three dimensional

ATP Adenosine triphosphate

BLAST Basic Local Alignment Search Tool [46]

COM Center of mass

dBBn Dibromodimane (thiol reactive compound)

DCCM Dynamic cross-correlation map

DSSP define secondary structure to protein (Analysis tool)

ECL Extracellular loop

ENM Elastic network model

H Helix

H-bond Hydrogen bond

ICL Intracellular Loop

MD Molecular dynamics

MDR Multidrug resistance

MOE Molecular Operation Environment [51]

mP-gp Mouse P-glycoprotein

MTS Methanethiosulfonate

NBD Nucleotid Binding Domain

nm Nano meter

NMA Normal mode analysis

NMR Nuclear magnetic resonance

ns Nano second

PC Principal components

PDB Protein Data Bank [45]

P-gp P-glycoprotein

Pi Inorganic phosphate

POPC 1-Palmitoyl-2-oleoylphosphatidylchinoline

ps Pico second

RCNMA Rigid cluster normal mode analysis

RMSD Root mean square deviation

RMSF Root mean square fluctuation

TM Transmembrane

TMD Transmembrane Domain

UniProt Universal Protein Resource [52]

VMD Visual molecular dynamics [110]

List of Figures

1.1 Topology of P-glycoprotein. 6
1.2 Vacuum cleaner and flippase model modified from [25] and [26]. 7
1.3 ATP-switsch model 9
1.4 Nucleotide occlusion model 11

2.1 Alignment of human P-gp sequence to the 4M1M template sequence 21
2.2 *phi* and *psi* rotations. 24
2.3 Relation of sequence identity and the backbone atom RMSD value (modified from C. Chothial and A.M. Lesk [55]). 27
2.4 Overall model quality for the template. 28
2.5 Overall model quality for the homology model. 28
2.6 Residues of the homology model facing the drug-binding site 33
2.7 Location of the drug-binding sites 1 and 2 35
2.8 Drug-binding site1 view from outside of cell 36
2.9 Drug-binding site 2 view from outside of cell 40

3.1 GROMACS morphing procedure 49
3.2 UCSF Chimera cluster 52
3.3 Morphing structures 54

3.4 Morphing structure NBD-1/4-closed 58
3.5 morphing structure NBD-1/2-closed 60
3.6 Morphing structure NBD-3/4-closed 61
3.7 Normal mode approach 64
3.8 Inserted linker 65
3.9 Loss of structure 72

4.1 NBD-open P-gp conformation embedded in POPC bilayer membrane. 81
4.2 Example of data merging 93

5.1 Superposing of the NBD-open conformation 96
5.2 COM calculations of the NBD distances of ligand-free P-gp conformations. 98
5.3 COM calculations of the NBD distances of P-gp conformations including colchicine. 100
5.4 RMSD of the whole Protein, and the separate domains. 103
5.5 DSSP of the TMD1 of NBD-3/4-closed conformation of P-gp 105
5.6 Unwinding of helix 9 107
5.7 RMSF of NBD-open 108
5.8 Marked areas of possible P-gp hinges 110
5.9 Number of H-bonds of helix 3 and helix 4 114
5.10 Trend of number of H-bonds of helix 3 115
5.11 Trend of number of H-bonds of helix 4 116
5.12 Location and movement of the investigated helices . . 118
5.13 Rearrangement of helix 3 and helix 3. 121
5.14 Rearrangement of helix 3 and helix 4. 124
5.15 Rearrangement of helix 2 and helix 4. 127
5.16 Rearrangement of helix 4 and helix 11. 130
5.17 Length of helices 2, 3, 4 and 11 131

5.18 Location of TM4/TM6 gate 133
5.19 Location of residues picked for TM4/TM6 measurement 133
5.20 Grid points calculated by Volmap of VMD [110] 134
5.21 Occupancy of backbone atoms of TM4 and TM6. . . . 136
5.22 TM4 kink observed by *P. Szewczyk et al.* [111]. 137
5.23 Distances TM4 to TM6 NBD-open conforamtions . . . 138
5.24 Distances TM4 to TM6 NBD-1/4-closed conformations 139
5.25 Distances TM4 to TM6 NBD-1/2-closed, NBD-3/4-closed and NBD-closed conformations 141
5.26 Kink of crystal structure PDB-ID 4Q9I 142
5.27 MTS (methanethiosulfonate) compounds and their spans in Å. For the distance of the $C_\alpha atoms of P-gp, 5.5\,\mathring{A}\ has to be added to the linker span$
5.28 Correlation of ATPase activity 145
5.29 Residues which were compaied to cross-linking data . 146
5.30 Distances of L175 to N820 of ligand-free P-gp conformations 148
5.31 Distances of L175 to N820 of P-gp conformations with colchicine 149
5.32 Distances P517 to I1050 ligand-free conformations . . 150
5.33 Distances of P517 to I1050 conformations with colchicine as ligand 151
5.34 distances of A80 to R741 of the ligand-free P-gp conformations. 153
5.35 Distances of A80 to R741 of the colchicine included conformations. 155
5.36 Residues that line the drug-binding site 156
5.37 Residues which were taken to measure the dimensions of P-gp drug-binding cavity 159

6.1 NBD-distance measurement of the NBD-open conformation with docked substrates 165

6.2 NBD-distance measurement of NBD-open conformation with docked inhibitors 166
6.3 RMSD of colchicine 169
6.4 RMSD of colchicine 169
6.5 RMSD of daunorubicin and Hoechst 33342 170
6.6 RMSD of elacridar and tariquidar 171
6.7 RMSD of cis-flupentixol and trans-flupentixol 172
6.8 Ligand interactions of colchicine 173
6.9 Colchicine position after 35 ns 174
6.10 Orientation of Gly346 and Ala343. 175
6.11 Binding pocket of colchicine 175
6.12 Ligand interactions of colchicine 177
6.13 Daunorubicin in binding pocket 178
6.14 Ligand interactions of daunorubicin 179
6.15 Ligand interactions of Hoechst 33342 180
6.17 Movement of elacridar 182
6.18 Ligand interactions of elacridar 184
6.19 Binding pocket of tariquidar 185
6.20 Ligand interactions of tariquidar 187
6.21 Residues interacting with H-bonds with tariquidar . . 188
6.22 Docking position of cis- and trans-flupentixol 189
6.23 Ligand interactions cis- and trans-flupentixol 190
6.24 Binding pocket cist-flupentixol and trans-flupentixol . 192
6.25 Ligandinteractions with P-gp after 35 ns. 193
6.26 Distances of portal TM4/TM6 residues for P-gp including daunorubicin and Hoechst 33342 197
6.27 Distances of portal TM4/TM6 residues for P-gp including elacridar and tariquidar 198
6.28 Distances of portal TM4/TM6 residues for P-gp including cis-flupentixol and trans-flupentixol 200

6.29 Dimensions of initial portal of TM4/TM6, Cα atom measurement 201
6.30 Dimensions of TM4/TM6 portal, protein surface . . . 202
6.31 Dimensions of TM4/TM6 portal of P-gp including cis-flupentixol 202
6.32 TM4/TM6 portal of P-gp including cis-flupentixol and trans-flupentixol, respectively 203
6.33 TM4/TM6 portal of P-gp including Hoechst 33342 and tariquidar, respectively 204

7.1 PCA NBD-open. The movement through space is shown from blue to white to red. Each dot corresponds to a different time frame of the MD simulation. Different cluster of dots correspond to different conformations. 208
7.2 Residue-wise loading of first three PC of NBD-open . . 210
7.3 P-gp movement 211
7.4 Residue-wise loading of first three PC of NBD-open . . 213
7.5 Residue-wise loadings of morphing structures. 215
7.6 NBD-$^1/_4$-closed movement. 216
7.7 Residue-wise loadings of NBD-closed. 217
7.8 Residue-wise loadings of PC1 of the substrates. 220
7.9 Residue-wise loadings of PC1 of elacridar and tariquidar 221
7.10 Residue-wise loadings of PC1 of cis- and trans-flupentixol 223
7.11 Cross-correlation maps of P-gp and morphing structures. 229
7.12 Cross-correlation maps of P-gp and substrates 232
7.13 Residue cross correlation of inhibitors. 235

9.1 Superposition of PDB ID: 6FN1 (orange) and the homology model (blue/cyan). The red circle shows the biggest differences. 246

9.2 Superposition of PDB ID: 6C0V (orange) and the homology model (blue/cyan). 247
9.3 Ligand interaction in binding cavity of the inhibitor zosuquinidar and the mouse-human chimeric P-gp. . 248

B.1 NBD-1/4-closed movement. 261
B.2 NBD-1/2-closed movement. 262
B.3 NBD-3/4-closed movement. 263
B.4 NBD-closed movement. 264

List of Tables

2.1 Listings of the BLAST results. 19
2.2 Distribution of residues by means of ramachandran plots analysed by Procheck. 24
2.3 Distribution of residues by means of ramachandran plots analysed by MOE. 25
2.4 Comparison of the transmembrane sequences of the homology model and the UniProt listings. 30
2.5 Table of residues (TMD1) located in drug-binding site 1. 38
2.6 Table of residues (TMD2) located in drug-binding site 1. 39
2.7 Table of residues (TMD1) located in the drug-binding site 2. 42
2.8 Table of residues (TMD2) located in the drug-binding site 2. 43

3.1 RMSD [Å] of the five conformations 55
3.2 Ramachandran plot results of the morphing transition states generated by GROMACS 56
3.3 Ramachandran plot results of the morphing transition states generated by UCSF Chimera 56
3.4 Rigid cluster decomposition of the first RCNMA run . 68
3.5 Rigid cluster decomposition of the second RCNMA run 69
3.6 RMSD of the NMSim Server conformations 71

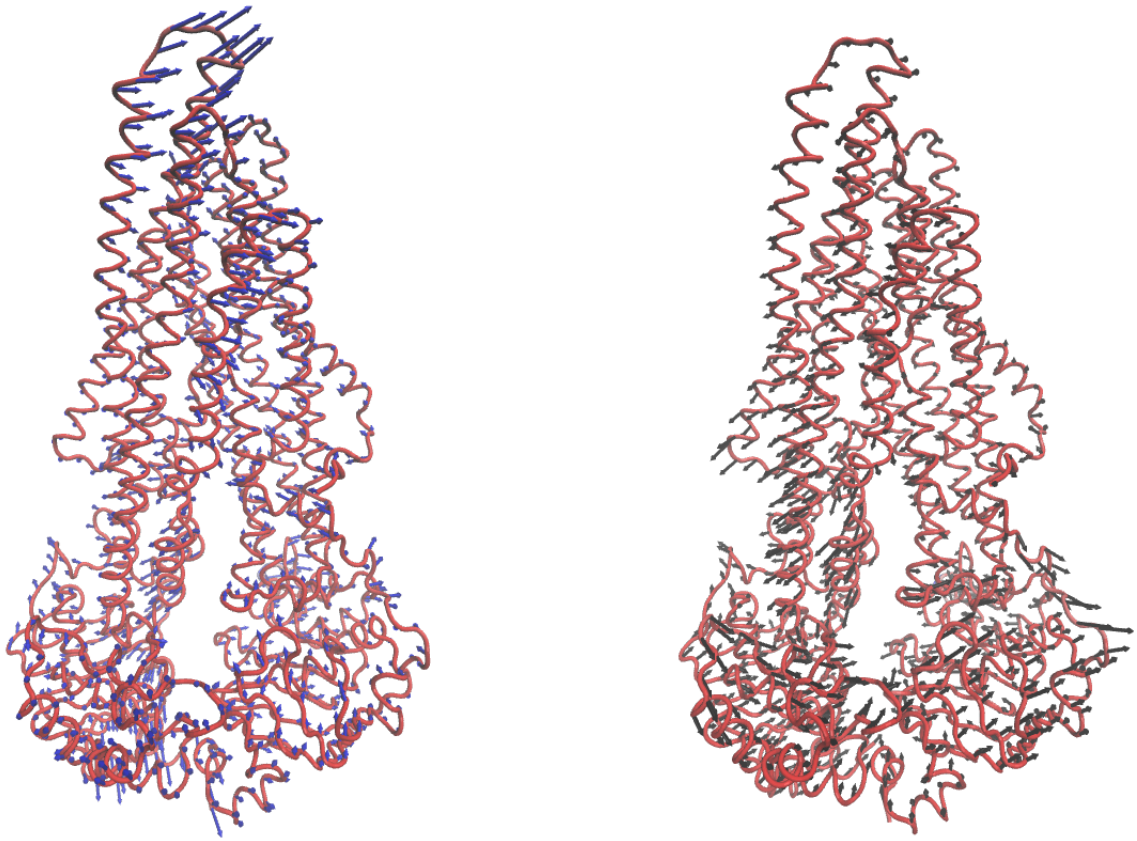
4.1 Amount of molecules of each simulation. 83

5.1 center of mass calculation of the NBD distances before the molecular dynamics simulation of the colchicine docked conformations. 101
5.2 Regions of helical unwinding in all conformational P-gp structures. 104
5.3 Different helical structures of P-gp conformations, identified by DSSP analysis 106
5.4 Dimensions of drug-binding domain 157

7.1 Production time. *(The equilibration and production time may be seen in section 6.2)*. 206
7.2 Cross-correlation pattern 227

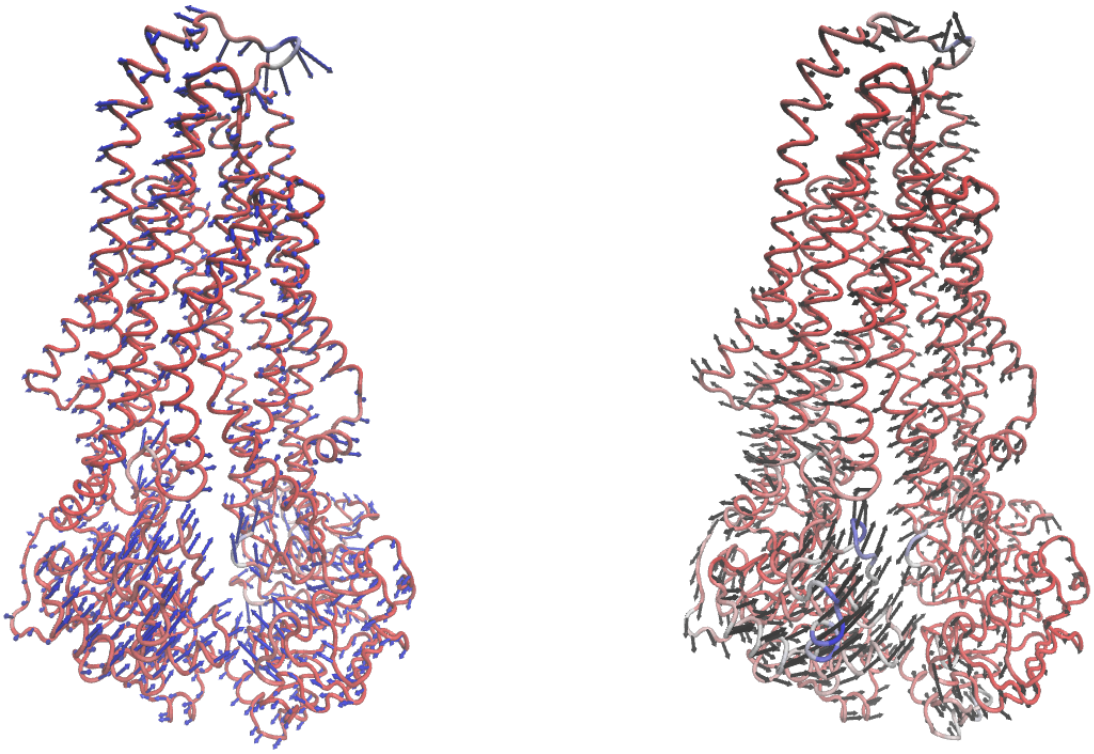
Appendix B

PCA movement



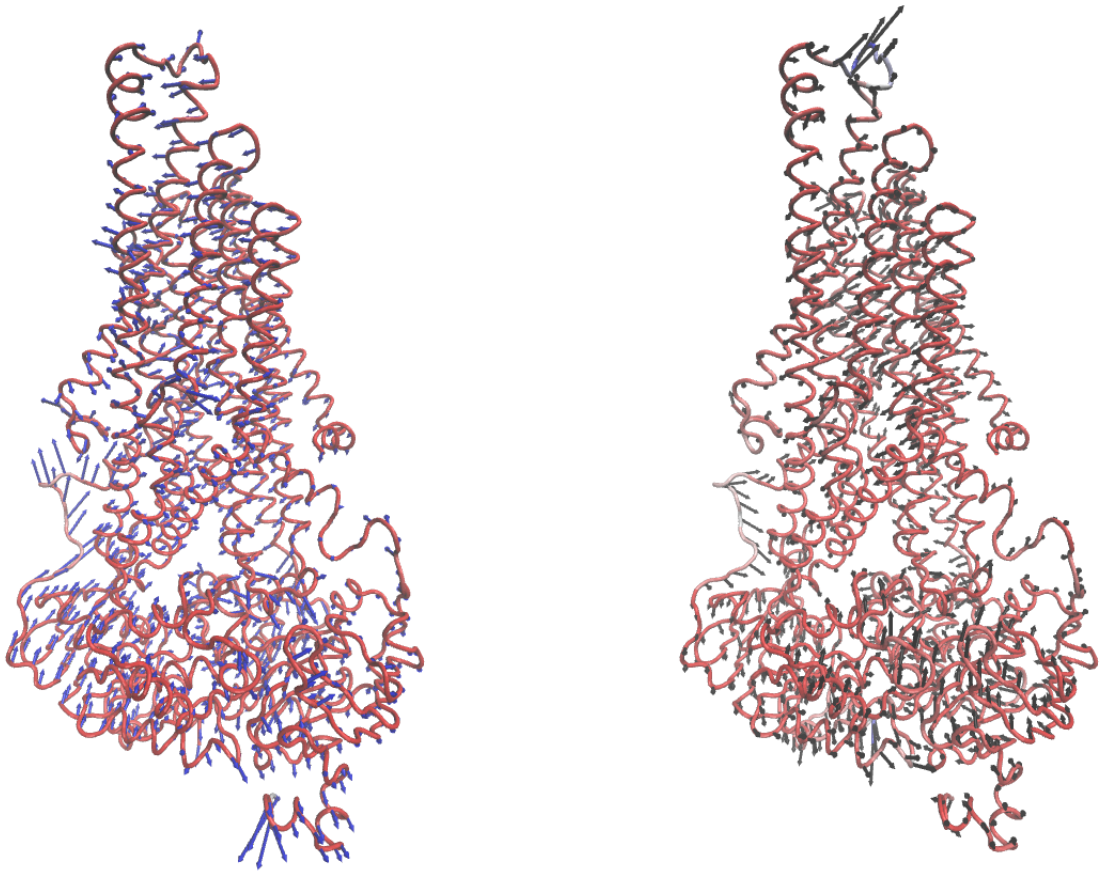
(a) PC1. **(b)** PC2.

Figure B.1: NBD-1/4-closed movement.



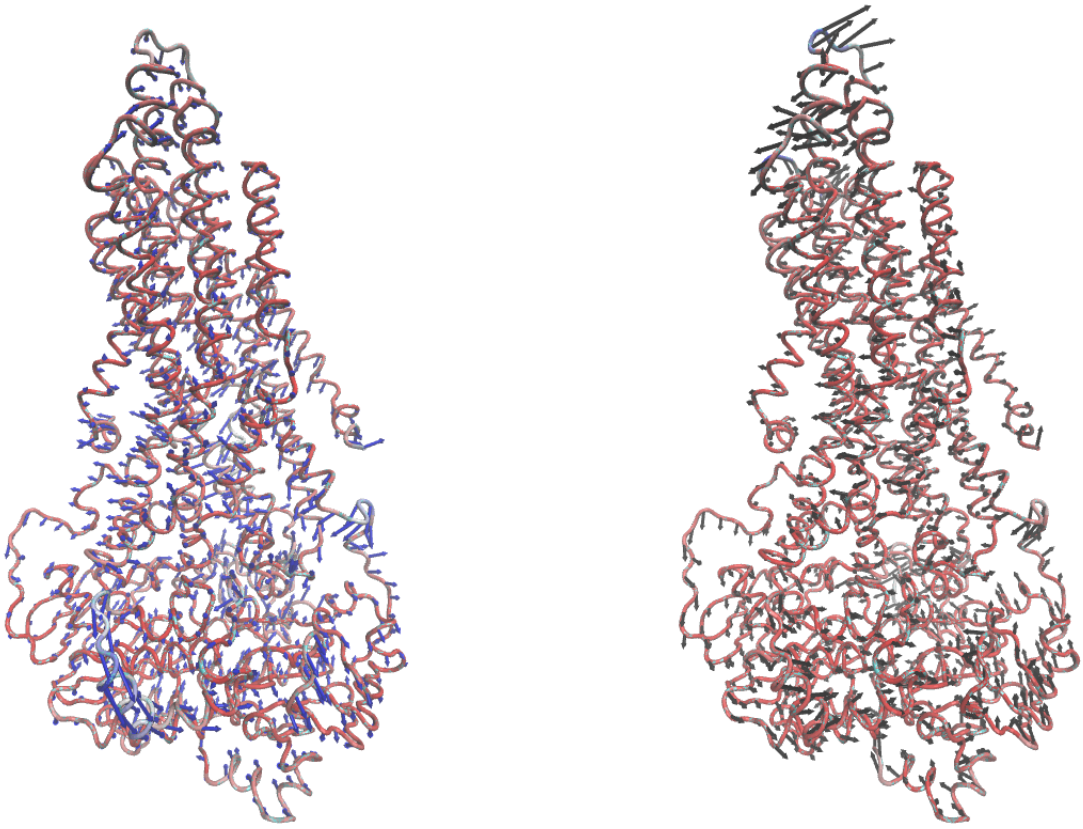
(a) PC1. **(b)** PC2.

Figure B.2: NBD-1/2-closed movement.



(a) PC1. **(b)** PC2.

Figure B.3: NBD-3/4-closed movement.

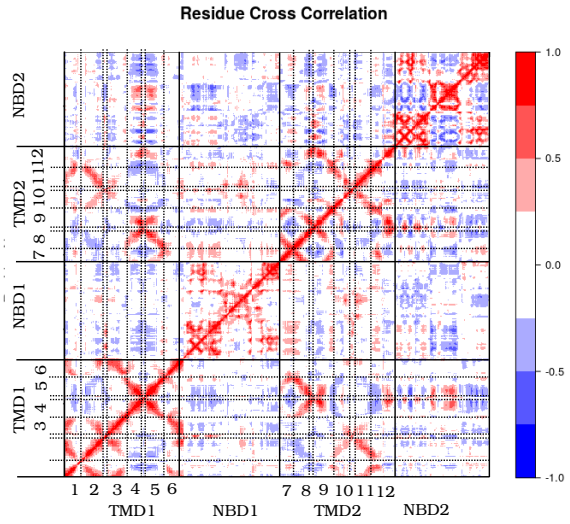


(a) PC1. **(b)** PC2.

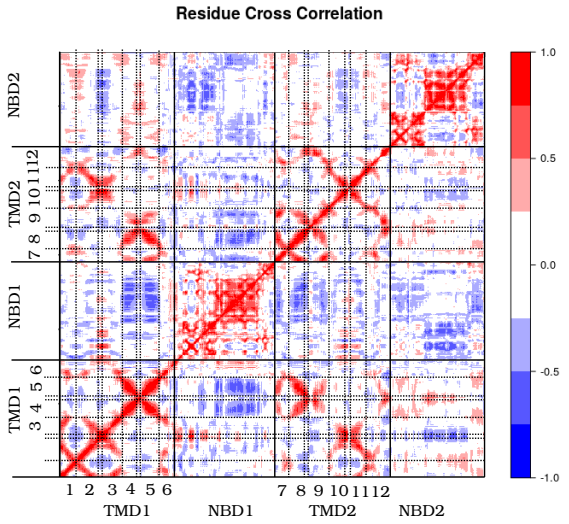
Figure B.4: NBD-closed movement.

Appendix C

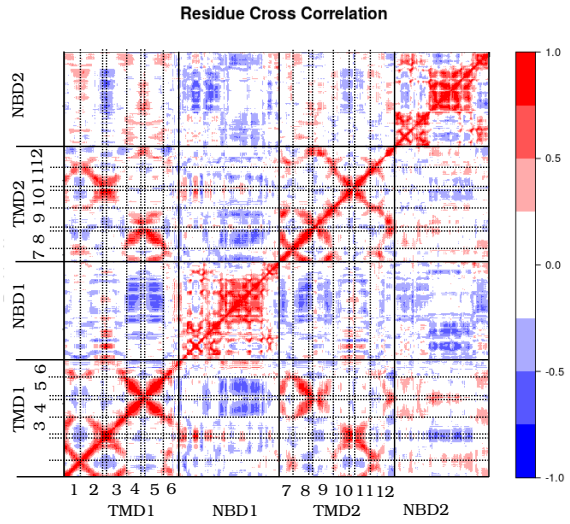
DCCM plots



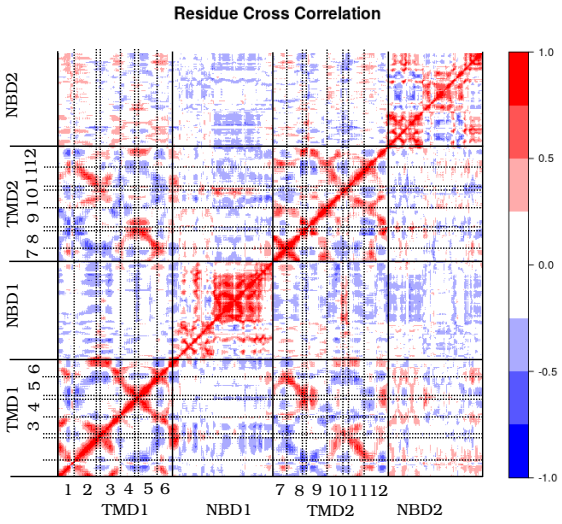
(a) Elacridar



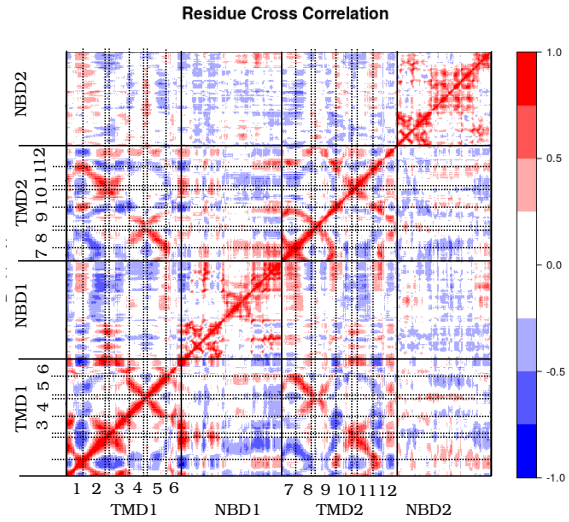
(b) Trans-flupentixol.



(c) Trans-flupentixol.



(a) NBD-1/2-closed.



(b) NBD-1/4-closed.

Bibliography

[1] Fojo, A. T.; Ueda, K.; Slamon, D. J.; Poplack, D. G.; Gottesman, M. M.; Pastan, I. Expression of a multidrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci U S A.* **1987**, *84*, 265–9.

[2] Fletcher, J. I.; Williams, R. T.; Henderson, M. J.; MD, M. D. N.; Haber, M. ABC transporters as mediators of drug resistance and contributors to cancer cell biology. *Drug Resist Updat.* **2016**, *26*, 1–9.

[3] Gottesman, M. M.; Fojo, T.; Bates, S. E. Multidrug Resistance in Cancer: Role of ATP-Dependent transporters. *Nat Rev Cancer* **2002**, *2*, 48–58.

[4] Higgins, C. F. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **1992**, *8*, 67–113.

[5] Dean, M.; Annilo, T. Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annu Rev Genomics Hum Genet.* **2005**, *6*, 123–42.

[6] Rees, D. C.; Johnson, E.; Lewinson, O. ABC transporters: the power to change. *Nat Rev Mol Cell Biol.* **2009**, *10*, 218–27.

[7] Szöllősi, D.; Rose-Sperling, D.; Hellmich, U. A.; Stockner, T. Comparison of mechanistic transport cycle models of ABC

exporters. *Biochim Biophys Acta.* **2017**, *S0005-2736*, 30346–2.

[8] Huls, M.; Russel, F. G. M.; Masereeuw, R. The role of ATP binding cassette transporters in tissue defense and organ regeneration. *J Pharmacol Exp Ther.* **2009**, *328*, 3–9.

[9] Beis, K. Structural basis for the mechanism of ABC transporters. *Biochem Soc Trans.* **2015**, *43*, 889–93.

[10] Seeger, M. A.; van Veen, H. W. Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta.* **2009**, *1794*, 725–37.

[11] ter Beek, J.; Guskov, A.; Slotboom, D. J. Structural diversity of ABC transporters. *J Gen Physiol.* **2014**, *143*, 419–435.

[12] E, E. B.-O.; Doeven, M. K.; Poolman, B. ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett* **2006**, *580*, 1023–35.

[13] Dawson, R.; Locher, K. Structure of a bacterial multidrug ABC transporter. *Nature* **2006**, *443*, 180–185.

[14] Saier, M. H. J.; Paulsen, I. T. Phylogeny of multidrug transporters. *Semin Cell Dev Biol.* **2001**, *12*, 205–13.

[15] Juliano, R. L.; Ling, V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochem Biophys Acta* **1976**, *455*, 152–62.

[16] Beaulieu, E.; Demeule, M.; Ghitescu, L.; Beliveau, R. P-glycoprotein is strongly expressed in the luminal membranes of the endothelium of blood vessels in the brain. *Biochem J.* **1997**, *326*, 539–44.

[17] Sharom, F. J. The P-glycoprotein Multidrug Transporter. *Essays Biochem.* **2011**, *50*, 161–78.

[18] Zoghbi, M. E.; Mok, L.; Swartz, D. J.; Singh, A.; Fendley, G. A.; Urbatsch, I. L.; Altenberg, G. A. Substrate-induced conformational changes in the nucleotide-binding domains of lipid bilayer-associated P-glycoprotein during ATP hydrolysis. *J Biol Chem.* **2017**, *292*, 20412–24.

[19] Doshi, R.; van Veen, H. W. Substrate binding stabilizes a pre-translocation intermediate in the ATP-binding cassette transport protein MsbA. *J Biol Chem* **2013**, *288*, 21638–47.

[20] Ishikawa, T.; Hirano, H.; Onishi, Y.; Sakurai, A.; Tarui, S. Functional evaluation of ABCB1 (P-glycoprotein) polymorphisms: high-speed screening and structureactivity relationship analyses. *Drug Metab Pharmacokinet.* **2004**, *19*, 1–4.

[21] Jin, M. S.; Oldham, M. L.; Zhang, Q.; Chen, J. Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* **2012**, *490*, 566–69.

[22] Aller, S.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P.; Trinh, Y.; Zhang, Q.; Urbatsch, I.; Chang, G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Sience* **2009**, *323*, 1718–22.

[23] Li, J.; Jaimes, K.; Aller, S. Refined structures of mouse P-glycoprotein. *Protein Science* **2014**, *23*, 34–46.

[24] Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **2007**, *446*, 749–57.

[25] Siarheyeva, A.; Lopez, J. J.; Glaubitz, C. Localization of multidrug transporter substrates within model membranes. *Biochemistry* **2006**, *45*, 6203–11.

[26] Higgins, C. F.; Gottesman, M. M. Is the multidrug transporter a flippase? *Trands Biochem Sci.* **1992**, *17*, 18–21.

[27] Eytan, G. D. Mechanism of multidrug resistance in relation to passive membrane permeation. *Biomed Pharmacother.* **2005**, *59*, 90–7.

[28] Smith, P. C.; Karpowich, N.; Millen, L.; Moody, J. E.; Rosen, J.; Thomas, P. J.; Hunt, J. F. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell* **2002**, *10*, 139–49.

[29] Higgins, C. F.; Linton, K. J. The ATP switch model for ABC transporters. *Nat Struct Mol Biol* **2004**, *11*, 918–26.

[30] Linton, K. J.; Higgins, C. F. Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch.* **2007**, *453*, 555–67.

[31] Liu, R.; Sharom, F. J. Site-directed fluorescence labeling of P-glycoprotein on cysteine residues in the nucleotide binding domains. *Biochemistry* **1996**, *35*, 11865–73.

[32] Siarheyeva, A.; Liu, R.; Sharom, F. J. Characterization of an Asymmetric Occluded State of P-glycoprotein with Two Bound Nucleotides: Implication For Catalysis. *J Biol Chem.* **2010**, *285*, 7575–86.

[33] Loo, T. W.; Clarke, D. M. Mutational analysis of ABC proteins. *Arch Biochem Biophys.* **2008**, *476*, 51–64.

[34] Pleban, K.; Kopp, S.; Csaszar, E.; Peer, M.; Hrebicek, T.; Rizzi, A.; Ecker, G. F.; Chiba, P. P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. *Mol Pharmacol.* **2005**, *67*, 365–74.

[35] Shapiro, A. B.; Ling, V. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur J Biochem.* **1997**, *250*, 130–7.

[36] Shapiro, A. B.; Fox, K.; Lam, P.; Ling, V. Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. Evidence for a third drug-binding site. *Eur J Biochem.* **1999**, *259*, 841–50.

[37] Martin, C.; Berridge, G.; Higgins, C. F.; Mistry, P.; Charlton, P.; Callaghan, R. Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol.* **2000**, *58*, 624–32.

[38] Brinkmann, U.; Eichelbaum, M. Polymorphisms in the ABC drug transporter gene MDR1. *Pharmacogenomics J.* **2001**, *1*, 59–64.

[39] Binkhathlan, Z.; A., A. L. P-glycoprotein inhibition as a therapeutic approach for overcoming multidrug resistance in cancer: current status and future perspectives. *Curr Cancer Drug Targets* **2013**, *13*, 326–46.

[40] Krishna, R.; Mayer, L. D. Multidrug resistance (MDR) in cancer. Mechanisms, reversal using modulators of MDR and the role of MDR modulators in influencing the pharmacokinetics of anticancer drugs. *Eur J Pharm Sci.* **2000**, *11*, 265–83.

[41] Thomas, H.; Coley, H. M. Overcoming multidrug resistance in cancer: an update on the clinical strategy of inhibiting p-glycoprotein. *Cancer Control.* **2003**, *10*, 159–65.

[42] Baguley, B. C. Multiple drug resistance mechanisms in cancer. *Mol Bioltechnol.* **2010**, *46*, 308–16.

[43] Pusztai, L.; Wagner, P.; Ibrahim, N.; Rivera, E.; Theriault, R.; Booser, D.; Symmans, F. W.; Wong, F.; Blumenschein, G.; Fleming, D. R.; Rouzier, R.; Boniface, G.; Hortobagyi, G. N. Phase II study of tariquidar, a selective P-glycoprotein inhibitor, in patients with chemotherapy-resistant, advanced breast carcinoma. *Cacer* **2005**, *104*, 682–91.

[44] Pokharel, D.; Roseblade, A.; Oenarto, V.; Lu, J. F.; Bebawy, M. Proteins regulating the intercellular transfer and function of P-glycoprotein in multidrug-resistant cancer. *Ecancermedicalscience.* **2017**, *11*, 786.

[45] Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. The Protein Data Bank. *Nucleic Acids Research* **2000**, *28*, 235–242.

[46] Altschul, S.; Gish, W.; abd E.W. Myers, W. M.; Lipman, D. Basic Local Alignment Tool. *J. Mol. Biol.* **1990**, *215*, 403–10.

[47] Jin, M.; Oldham, M.; Zhang, Q.; Chen, J. Crystal structure of the multidrug transporter P-glycoprotein from C. elegans. *Nature* **2012**, *490*, 566–69.

[48] Shintre, C. et al. Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9710–15.

[49] Hohl, M.; C.Briad,; Gruetter, M.; Seeger, M. Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat. Struct. Mol. Biol.* **2012**, *19*, 395–402.

[50] Styczynski, M.; Jensen, K.; Rigoutsos, I.; Stephanopoulos, G. BLOSUM62 miscalculations improve search performance. *Nat. Biotech* **2008**, *26*, 274–275.

[51] Inc., C. C. G. Molecular Operating Environment (MOE) 2013.8. 2016; 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7.

[52] Consortium, T. U. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2008**, *36 (Database Issue)*, D190–D195.

[53] Globisch, C. Molecular Modelling Untersuchungen des ABC-Transporters P-Glykoprotein. 2008; Dissertation, University of Bonn, Germany.

[54] Laskowski, R.; MacArthur, M.; Moss, D.; Thorton, J. PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **1993**, *26*.

[55] Chothial, C.; Lesk, A. The relation between the divergence of sequence and structure in proteins. *The EMBO Journal* **1986**, *5*, 823–826.

[56] Sippl, M. Recognition of errors in three-dimensional structures of proteins. *Proteins* **1993**, *17*, 355–362.

[57] Wiederstein, M.; Sippl, M. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35 (Web Server issue)*, W407–10.

[58] Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Identification of residues in the drug translocation pathway of the human multidrug resistance P-glycoprotein by arginine mutagenesis. *J Biol Chem* **2009**, *284*, 24074–87.

[59] Loo, T.; Clarke, D. Drug Rescue Distinguishes between Different Structural Models of Human P-Glycoprotein. *Biochemistry* **2013**, *52*, 7167–69.

[60] Loo, T.; Clarke, D. Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. *J Biol Chem.* **2001**, *276*, 14972–9.

[61] Shapiro, A.; Ling, V. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur J Biochem* **1997**, *250*, 130–7.

[62] Chufan, E.; Kapoor, K.; H.M.Sim,; Singh, S.; Talele, T.; Durell, S.; Ambudkar, S. Multiple transport-active binding sites are available for a single substrate on human P-glycoprotein (ABCB1). *PLoS One* **2013**, *8*, e82463.

[63] McCormick, J.; Vogel, P.; Wise, J. Multiple Drug Transport Pathways through Human P-Glycoprotein. *Biochemistry* **2015**, *54*, 4374–4390.

[64] Loo, T.; Clarke, D. Mapping the Binding Site of the Inhibitor Tariquidar That Stabilizes the First Transmembrane Domain of P-glycoprotein. *J Biol Chem.* **2015**, *290*, 29389–29401.

[65] Loo, T. W.; Clarke, D. M. Determining the dimensions of the drug-binding domain of human P-glycoprotein using thiol

cross-linking compounds as molecular rulers. *J. Biol. Chem.* **2001**, *276*, 36877–80.

[66] Loo, T.; Bartlett, M.; Clarke, D. Permanent activation of the human P-glycoprotein by covalent modification of a residue in the drug-binding site. *J Biol Chem.* **2003**, *278*, 20449–52.

[67] Loo, T.; Bartlett, M.; Clarke, D. Suppressor mutations in the transmembrane segments of P-glycoprotein promote maturation of processing mutants and disrupt a subset of drug-binding sites. *J Biol Chem.* **2007**, *282*, 32043–52.

[68] Chufan, E. E.; Kapoor, K.; Sim, H. M.; Singh, S.; Talele, T. T.; Durell, S. R.; Ambudkar, S. V. Multiple transport-active binding sites are available for a single substrate on human P-glycoprotein (ABCB1). *PLoS One* **2013**, *8*, e82463.

[69] Loo, T.; Clarke, D. Cross-linking of human multidrug resistance P-glycoprotein by the substrate, tris-(2-maleimidoethyl)amine, is altered by ATP hydrolysis. Evidence for rotation of a transmembrane heli. *J Biol Chem.* **2001**, *276*, 31800–5.

[70] Song, J.; Melera, P. Transmembrane domain (TM) 9 represents a novel site in P-glycoprotein that affects drug resistance and cooperates with TM6 to mediate [125I]iodoarylazidoprazosin labeling. *Mol Pharmacol.* **2001**, *60*, 254–61.

[71] Loo, T.; Clarke, D. Identification of residues in the drug-binding site of human P-glycoprotein using a thiol-reactive substrate. *J Biol Chem* **1997**, *272*, 31945–8.

[72] Loo, T.; Clarke, D. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem.* **2002**, *277*, 44332–8.

[73] Loo, T.; Clarke, D. Mutations to amino acids located in predicted transmembrane segment 6 (TM6) modulate the activity and substrate specificity of human P-glycoprotein. *Biochemistry* **1996**, *33*, 14049–57.

[74] Dey, S.; Hafkemeyer, P.; Pastan, I.; Gottesman, M. A single amino acid residue contributes to distinct mechanisms of inhibition of the human multidrug transporter by stereoisomers of the dopamine receptor antagonist flupentixol. *Biochemistry* **1999**, *38*, 6630–9.

[75] Gruol, D.; Bernd, J.; Phippard, A.; I.Ojima,; Bernacki, R. The use of a novel taxane-based P-glycoprotein inhibitor to identify mutations that alter the interaction of the protein with paclitaxel. *Mol Pharmacol.* **2001**, *60*, 104–13.

[76] Doenmez, C.; Khunweeraphong, N.; Parveen, Z.; Schmid, D.; Artaker, M.; Ecker, G.; Sitte, H.; Pusch, O.; Stockner, T.; Chiba, P. Pore-exposed tyrosine residues of P-glycoprotein are important hydrogen-bonding partners for drugs. *Mol Pharmacol.* **2014**, *85*, 420–8.

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

[78] Loo, T.; Clarke, D. Identification of residues in the drug-binding domain of human P-glycoprotein. Analysis of transmembrane segment 11 by cysteine-scanning mutagenesis and inhibition by dibromobimane. *J Biol Chem.* **1999**, *274*, 35388–92.

[79] Loo, T.; Clarke, D. Functional consequences of glycine mutations in the predicted cytoplasmatic loops of P-glycoprotein. *J.Biol.Chem* **1994**, *269*, 7243–8.

[80] Schlitter, J.; Engels, M.; Krüger, P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *J Mol Grpah* **1994**, *2*, 84–9.

[81] Ma, J. Unsefulness and limitations of normal mode analysis inmodeling dynamics of biomolecular complexes. *Structure* **2005**, *13*, 373–80.

[82] Lopez-Blanco, J.; Miyashita, O.; Tama, F.; Chacon, P. Normal Mode Analysis Techniques in Structural Biology. 2014; eLS. John Wiley & Sons, Ltd: Chichester.

[83] Berendsen, H.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **1995**, *91*, 43–56.

[84] Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **2004**, *25*, 1605–12.

[85] Kelley, L.; Gardner, S.; Sutcliffe, M. An automated approach for clustering an ensemble of NMR-derived protein structures

into conformationally related subfamilies. *Protein Eng.* **1996**, *9*, 1063–5.

[86] Loo, T.; Bartlett, M.; Clarke, D. Val133 and Cys137 in transmembrane segment 2 are close to Arg935 and Gly939 in transmembrane segment 11 of human P-glycoprotein. *J Biol Chem.* **2004**, *279*, 18232–8.

[87] Krüger, D. M.; Ahmed, A.; Gohlke, H. NMSim Web Server: integrated approach for normal mode-based geometric simulations of biologically relevant conformational transitions in proteins. *Nucleic Acids Res* **2012**, *40 (Web Server issue*, W310–W316.

[88] Best, R.; Zhu, X.; Shim, J.; Lopes, P.; Mittal, J.; Feig, M.; Mackerell, A. J. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–73.

[89] MacKerell, A. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* **1998**, *102*, 3586–616.

[90] Oostenbrink, C.; Villa, A.; Mark, A.; van Gunsteren, W. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–76.

[91] Marrink, S.; Risselada, H.; Yefimov, S.; Tieleman, D.; Vries, A. D. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B.* **2007**, *111*, 7812–24.

[92] Schmidt, T.; Kandt, C. LAMBADA and InflateGRO2: Efficient Membrane ALignment and Insertion of Membrane Proteins for Molecular Dynamics Simulations. *J. Chem: Inf. Model.* **2012**, *52*, 2657–69.

[93] Wolf, M.; Hoefling, M.; Aponte-Santamaria, C.; Grubmueller, H.; Groenhof, G. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J Comput Chem.* **2010**, *31*, 2169–74.

[94] Berendsen, H.; Postma, J.; van Gunsteren, W.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. *Intermolecular Forces* **1981**, 331–342, Reidel and Dordrecht, Holland.

[95] Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, 926–35.

[96] Mahoney, M.; Jorgensen, W. A five-site model for liquid water and the reproduction of the density anomoly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, 8910–22.

[97] Leach, A. Molecular Modelling. 2001; Pearson Prentice Hall, Harlow.

[98] Essmann, U.; Oerera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh Ewald method. *J. CHem. Phys.* **1995**, *103*, 8577–93.

[99] CHeng, A.; K.M. Merz, J. Application of the Nosé-Hoover Chain Algorithm to the Study of Protein Dynamics. *J. Phys. Chem.* **1996**, *100*, 1927–37.

[100] Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys.* **1977**, *23*, 327–41.

[101] Jones, G.; Willet, P.; Glen, R.; an R. Taylor, A. L. Development and validation of a genetic algorithm for flexible docking. *J MolBiol* **1997**, *267*, 727–48.

[102] Jones, G.; Willett, P.; Glen, R. Molecular recognition of receptorsites using a genetic algorithm with a description of desolvation. *JMol Biol* **1995**, *245*, 43–53.

[103] Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical coring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425–45.

[104] Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. Flexible docking using Tabu search and an empirical steimate of binding affinity. *Proteins* **1998**, *33*, 367–82.

[105] Schüttelkopf, A.; van Aalten, D. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr.* **2004**, *60*, 1355–63.

[106] Kabsch, W.; Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.

[107] Touw, W. G.; Baakman, C.; Black, J.; te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G. A series of PDB re-

lated databases for everyday needs. *Nucleic Acids Research* **2015**, *43 (Database issue)*, D364–68.

[108] Oldham, M. L.; Davidson, L. A.; Chen, J. Structural insights into ABC transporter mechanism. *Curr. Opin. Struct. Biol.* **2008**, *18*, 726–33.

[109] Kumar, P.; Bansal, M. Dissecting π-helices: sequence, structure and function. *FEBS J.* **2015**, *282*, 4415–32.

[110] Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.

[111] Szewczyk, P.; Tao, H.; McGrath, A. P.; Villaluz, M.; Rees, S. D.; Lee, S. C.; Doshi, R.; Urbatsch, I. L.; Zhang, Q.; Chang, G. Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. *Acta Crytallogr. D. Biol.* **2015**, *71*, 732–41.

[112] Loo, T. W.; Clarke, D. M. Membrane topology of a cysteine-less mutant of human P-glycoprotein. *J. Biol. Chem.* **1995**, *270*, 843–8.

[113] Loo, T. W.; Bartlett, M. C.; Detty, M. R.; Clarke, D. M. The ATPase Activity of the P-glycoprotein Drug Pump Is Highly Activated When the N-terminal and Central Regions of the Nucleotide-binding Domains Are Linked Closely Together. *J. Biol. Chem.* **2012**, *287*, 26806–16.

[114] Loo, T. W.; Clarke, D. M. Identification of the distance between the homologous halves of P-glycoprotein that triggers the High/Low ATPase activity switch. *J. Biol.Chem.* **2014**, *289*, 8484–92.

[115] Esser, L.; Zhou, F.; Pluchino, K. M.; Shiloach, J.; Ma, J.; Tang, W.; Gutierrez, C.; Zhang, A.; Shukla, S.; Madigan, J. P.; Zhou, T.; Kwong, P. D.; Ambudkar, S. V.; Gottesman, M. M.; Xia, D. Structures of the Multidrug Transporter P-glycoprotein Reveal Asymmetric ATP Binding and the Mechanism of Polyspecificity. *J. Biol. Chem.* **2017**, *292*, 446–61.

[116] Loo, T. W.; Clarke, D. M. Cysteines introduced into extracellular loops 1 and 4 of human P-glycoprotein that are close only in the open conformation spontaneously form a disulfide bond that inhibits drug efflux and ATPase activity. *J. Biol. Chem.* **2014**, *289*, 24749–58.

[117] Jones, P. M.; George, A. M. Mechanism of the ABC transporter ATPase domains: catalytic models and the biochemical and biophysical record. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 39–50.

[118] Loo, T. W.; Clarke, D. M. The packing of the transmembrane segments of human multidrug resistance P-glycoprotein is revealed by disulfide cross-linking analysis. *J. Biol. Chem.* **2000**, *275*, 5253–6.

[119] Shapiro, A. B.; Corder, A. B.; Ling, V. P-glycoprotein-mediated Hoechst 33342 transport out of the lipid bilayer. *Eur J Biochem.* **1997**, *250*, 115–21.

[120] Maki, N.; Dey, S. Biochemical and pharmacological properties of an allosteric modulator site of the human P-glycoprotein (ABCB1). *Biochem Pharmacol.* **2006**, *72*, 145–55.

[121] Mandal, D.; Moitra, K.; Ghosh, D.; Xia, D.; Dey, S. Evidence for modulatory sites at the lipid-protein interface of the human multidrug transporter P-glycoprotein. *Biochemistry* **2012**, *51*, 2852–66.

[122] Dey, S.; Hafkemeyer, P.; Pastan, I.; Gottesman, M. M. A single amino acid residue contributes to distinct mechanisms of inhibition of the human multidrug transporter by stereoisomers of the dopamine receptor antagonist flupentixol. *Biochemistry* **1999**, *38*, 6630–9.

[123] Hayward, S.; Groot, B. L. Normal modes and essential dynamics. *Methods Mol Biol.* **2008**, *443*, 89–106.

[124] Amadei, A.; Linssen, A. B.; Berendsen, H. J. Essential dynamics of proteins. *Proteins* **1993**, *17*, 412–25.

[125] Amadei, A.; Linssen, A. B.; de Groot, B. L.; van Aalten, D. M.; Berendsen, H. J. An efficient method for sampling the essential subspace of proteins. *J Biol Struct. Dyn.* **1996**, *13*, 615–25.

[126] Alam, A.; Küng, R.; Kowal, J.; McLeod, R. A.; Tremp, N.; Broude, E. V.; Roninson, I. B.; Stahlberg, H.; Locher, K. P. Structure of a zosuquidar and UIC2-bound human-mouse chimeric ABCB1. *PNAS* **2018**,

[127] Kim, Y.; Chen, J. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science* **2018**, *359*, 915–9.

Verfassererklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die den verwendeten Werken wörtlich oder inhaltlich entnommenen Stellen als solche gekennzeichnet.

Bonn, den 30.08.2019

Melanie S. Hafner