DIPLOMARBEIT

Titel der Diplomarbeit

„A model of the structure of a strigolactone ABC-transporter in *Petunia hybrida*"

verfasst von

Eva Hellsberg

angestrebter akademischer Grad

Magistra der Pharmazie (Mag.pharm.)

Wien, 2015

Acknowledgements

First of all, I want to thank my supervisor Prof. Dr. Gerhard Ecker. He gave me the chance to realise this project in his group and encouraged me to believe in myself. He always supported me with his great knowledge, stunning experience and wise advise. He is an admirable group leader and scientist providing a fabulous learning and training environment for his students in which they can contribute and develop their strengths. His huge enthusiasm for science is inspiring and thrilling all the way. Thank you so much for everything, Gerhard.

Floriane Montanari, PhD student herself in our research group, was my closest advisor and caretaker to carry out this project. She is an outstanding young scientist; due to her comprehensive skills and her endless diligence, she is working very hard and is always helping others with various side tasks. Nevertheless, she had always an open ear for my questions and always scheduled time for discussions about my topic. She supported me with her ideas, but encouraged me to create my own and take decisions myself at the same time. Furthermore, she helped me to find literature and improve my writing attempts - even though this topic did not exactly count to her core expertises.
Floriane, I could never thank you enough for your time, advice and patience; I will never forget your support and care!

Last but not least, I want to thank this wonderful research group. All the people here are so helpful, friendly and take a real interest in the problems of the others that it is a real pleasure for me to work here. Moreover, the people share leisure activities, which leads to a lot of fun and a faithful atmosphere in the group.

Thank you all!

Table of contents

1 Introduction – The ABC of phytohormone translocation

Contribution of the thesis author

The ABC of Phytohormone Translocation E. Hellsberg, F. Montanari, G.F. Ecker

E. Hellsberg gathered literature about the topic of this review and wrote the parts *Introduction, Metabolite transport, Immunity, Cytokinin transport, Strigolactone transport, Elucidation of the structure of a strigolactone transporter* and *Conclusions and Outlook.*

The ABC of Phytohormone Translocation

Authors   **Eva Hellsberg, Floriane Montanari, Gerhard F. Ecker**

Affiliation   University of Vienna, Department of Pharmaceutical Chemistry, Vienna, Austria

Abstract
▼

ATP-driven transport across biological membranes is a key process to translocate solutes from the interior of the cell to the extracellular environment. In humans, ATP-binding cassette transporters are involved in absorption, distribution, metabolism, excretion, and toxicity, and also play a major role in anticancer drug resistance. Analogous transporters are also known to be involved in phytohormone translocation. These include, e. g., the transport of auxin by ABCB1/19 in *Arabidopsis thaliana*, the transport of abscisic acid by AtABCG25, and the transport of strigolactone by the *Petunia hybrida* ABC transporter PDR1. Within this article, we outline the current knowledge about plant ABC transporters with respect to their structure and function, and provide, for the first time, a protein homology model of the strigolactone transporter PDR1 from *P. hybrida*.

Supporting information available online at http://www.thieme-connect.de/products

Correspondence
Gerhard F. Ecker
University of Vienna
Department of Pharmaceutical
Chemistry
Althanstraße 14
1090 Wien
Austria
Phone: + 43 1 4277 55 110
Fax: + 43 1 4277 85 5110
gerhard.f.ecker@univie.ac.at

Introduction
▼

ATP-binding cassette transporters (ABC transporters) represent a huge superfamily of proteins expressed in all phyla from prokaryotes to humans. In humans, they attract a lot of interest due to their role in multidrug resistance in cancer therapy, their involvement in drug/drug interactions, as well as their influence on the bioavailability and toxicity of drugs. ABC transporters are characterized by a nucleotide binding domain (NBD), which hydrolyzes ATP and provides energy for the conformational changes required for solute transport across the membranes, and a transmembrane domain (TMD), which forms a pore through which the substrates pass. Some soluble members of the ABC transporters only contain NBD, while the full-size members contain two NBD modules and two TMD modules. The so-called half-size ABC transporters are built out of one NBD and one TMD, and require dimerization for the transport activity (○ **Fig. 1**). ABC transporters in humans comprise 48 proteins and are classified into six families (ABCA–ABCG), with the multidrug resistance protein 1 (MDR1, P-glycoprotein; gene ABCB1), the breast cancer resistance protein (BCRP; gene ABCG2), and the multidrug resistance-related protein 1 (MRP1; gene

ABCC1) being the most prominent ones. In plants, more than 100 genes coding for ABC transporters have been discovered to date. Plant ABC transporters cluster in eight families, from ABCA to ABCI (the ABCH subfamily has not been found in plants). While in humans the ABCG family only consists of half-size transporters, plant ABCG transporters may be half or full size [1].

The plants lack a developed and dynamic vascular system, which may explain their need for an elaborated transport system to allow a proper distribution of nutrients and signals [2]. Two types of transport happen in plants: the long-range transport, where substances flow from root to shoot and back using the xylem and phloem vessels, and the short-range transcellular transport via the apoplastic compartments. Transport proteins, carriers, and permeases are involved in the latter type, allowing substances to pass from cell to cell, and from cell to vessel [2].

In analogy to humans, also in plants, hormones are essential for vegetative growth and development. Hormones are small molecules guarding signalling processes throughout the whole plant. Several of them have been known for a long time, including auxins, abscisic acid, cytokinins, gibberellins, and ethylene. In the last years, several more were discovered, such as brassinosteroids, jas-
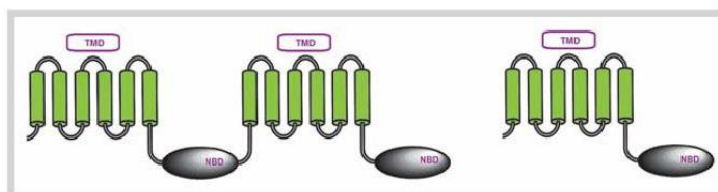
Fig. 1 Topology of a full and a half transporter. TMD = transmembrane domain, NBD = nucleotide binding domain. (Color figure available online only.)

monic acid, nitric oxide, salicylic acid, and strigolactones [3]. This group will likely grow even further in the coming years. As hormones take effect not only locally in the area of their biosynthesis but also in distant tissues, an adequate transport system is required.

In this article, we will provide an overview of the different roles of ABC transporters in plants, with a focus on hormone transport and available structural information on the ABC transport proteins involved. Finally, we will outline our protein homology modelling studies of a specific hormone transporter, PDR1 in *Petunia hybrida* (Solanaceae).

Multiple Roles of ATP-Binding Cassette Transporters in Plants
▼

Plant genomes contain more than 100 genes encoding ABC transporters [1]. Although a lot of work remains in order to identify the specific roles of each of those proteins across plant species, the current knowledge already reveals multiple critical functions. While it is not the aim of this article to extensively cover these roles (for detailed reviews, see [1] or [4]), we will briefly describe the main elucidated roles of ABC transporters in plants.

Detoxification
▼

The first role ever discovered for ABC transporters in plants was detoxification and elimination of toxic endo- and xenobiotics [5]. While the global detoxification pathway is similar to the one known for humans (oxidation by cytochrome P450, followed by a conjugation with a hydrophilic moiety like glucuronide or glutathione), the excretion step of the conjugated toxins towards the vacuolar compartment or the apoplasm is done by ABC transporters [5] such as AtABCC1, AtABCC2, and AtABCC3 [6–8]. ABCG36 in *Arabidopsis thaliana* (Brassicaceae) was shown to excrete cadmium from root hairs and root epidermal cells [9] towards the rhizosphere. Some results also point towards ABCG40 as a potential lead (Pb) extruder [10], which is surprising considering the fact that AtABCG40 is also involved in hormone import. In humans, similar metal transport roles are taken up by the solute carrier DMT1 (SLC11A2) [11].

Aluminium toxicity is a well-studied phenomenon because of its impact on crop yield. In acid soils, aluminium is present as Al^{3+}, which inhibits root growth with a subsequent decrease of shoot growth and overall crop productivity. Plants cope with Al toxicity by limiting the uptake and chelating internalized Al with malate or citrate. Recently, one Al-sensitive mutation was characterized in a previously uncharacterized half-size ABC transporter [12] expressed throughout the phloem and the epidermis of the roots in *A. thaliana*. The authors propose that this transporter helps remove Al from sensitive tissues and sequestrates it into more tol-

erant ones. Another Al-sensitive mutant in a different ABC transporter, AtABCB27, was also identified. It may facilitate vacuolar sequestration of Al in the roots [13].

Symbiosis and rhizosphere community
Via their roots, plants are in close contact with soil fungal and bacterial communities. In legumes, this contact turned into a beneficial symbiosis between specific bacterial strains that take care of nitrogen fixation while the plant is in charge of the photosynthesis of nutrients [14]. The legumes recruit their symbiosis partners via flavonoid excretion from the roots. Recent studies showed that ABC transporters, potentially of the PDR family, are involved in genistein and daidzein secretion in soybean [15,16]. Each plant species has its particular soil composition [17] that can change during the development of the plant [18]. This characteristic may be due to different abilities of roots to capture nutrients from the soil, as well as different exudations. Badri and colleagues, working on phytochemical exudations, showed that abcg30 mutants in *A. thaliana* significantly altered the amount and nature of microbial and fungal communities surrounding the roots. The phytochemicals excreted by the mutants lacking AtABCG30 were enriched in phenolic compounds such as benzoic acid or salicylic acid, and impoverished in glucose, fructose, or mannitol [19]. While no transport assays were performed in their study, the authors show that the absence of AtABCG30 resulted in pleiotropic effects in the mutants, with several transport genes overexpressed while sugar transporters were downregulated. The mutants were surrounded by a different microbiota, enriched in bacteria strains able to fixate nitrogen or to detoxify heavy metals and other pollutants. The experiment shows that lacking only one ABC transporter might completely change the expression pattern of other transporters and have consequences on the plant surroundings, including the neighbouring microbiota.

ABC transporters in roots seem to play a critical role in the establishment of an appropriate rhizobia for the plant. Controlling ABC transporter's root exudations could be a way of modifying the soil contents for beneficial purposes.

Metabolite transport
Metabolite transport is a broad subject. ○ **Table 1** provides an overview on the most prominent examples of ABC transporters involved in metabolite transport together with their specific tasks in plants.

AtABCC2 (AtMRP2) has been known for a long time to be involved in the transport of glutathione S-conjugates and chlorophyll catabolites [7]. Glutathione is necessary for the reduction of dehydro-L-ascorbic acid to ascorbic acid, and this process is crucial for hydrogen peroxide detoxification [20]. Chlorophyll is naturally catabolized during leaf senescence, and the metabolites are accumulated in the vacuole [21].

7

Transporter	Plant	Substrates	Reference	
AtABCC2	A. thaliana	gluthathione conjugates chlorophyll catabolites	Lu et al., 1998 [7]	**Table 1** Metabolite ATP-binding cassette transporters.
AtABCC5	A. thaliana	phytate (IP_6)	Nagy et al., 2009 [22]	
AtABCD1	A. thaliana	CoA esters of fatty acids indole-3-butyric acid	Zolman et al., 2001 [23]	
CjMDR1	C. japonica	berberines	Yazaki et al., 2001 [24]	
AtPDR2	A. thaliana	amino acids	Badri et al., 2008 [27]	
AtABCG34	A. thaliana	organic acids	Badri et al., 2008 [27]	

AtABCC5 transports phytate (IP_6), which is an important phosphorus supplier, into the vacuole. IP_6 is a messenger in guard cells, where it is essential for correct stomatal movement. In seeds, it helps storing metallic cations by chelating them [22].

Fatty acid β-oxidation is needed to build up acetyl-CoA for the citric acid cycle. ABCD1 (PXA1) in *A. thaliana* imports coenzyme A esters of fatty acids and indole-3-butyric acid (a precursor of auxin) into the peroxisome to enable the oxidation process [23]. The ABC transporter MDR1 uptakes berberines in *Coptis japonica*. As a consequence, the isoquinoline alkaloids accumulate in the rhizome [24]. They are synthesized in the roots, reach the rhizome through connecting xylem tissues, and are finally uptaken by CjMDR1, which is therefore working as a cell importer, probably the least common transport direction among the ABC transporters in eukaryotic cells [25]. The real use of berberines for the plant remains unclear, although there is an assumption for feedback regulation in berberine biosynthesis [26].

Badri et al. presented another extensive investigation about the connection of root exudates with ABC transporters by studying six full-length and one half transporters derived from MRP, PDR (pleiotropic drug resistance), and PGP (P-glycoprotein) subfamilies in *A. thaliana*. Their experiments were mainly based on comparisons of certain metabolites in the root exudates between transporter knockout and wild-type plants. Within their studies, they proved that ABC transporters are involved in the root secretion process and that there can be many transporters involved in the secretion of one metabolite, or a single transporter can be involved in the pathways of many metabolites [27] (○ **Table 1**).

Immunity

Several members of the PDR subfamily of ABC transporters have been found to protect distinct plant species from infection by pathogens. For example, the silencing of *Nicotiana plumbaginifolia* PDR1 increased the sensitivity of the plant to fungal and oomycete pathogens [28]. In *A. thaliana*, dysfunctional mutants of the gene ABCG36 (PDR8) become more sensitive to infectious bacterial strains [29], and ABCG40 (PDR12) is overexpressed when exposed to fungal pathogens [30].

While it is not quite clear how those full-length G family transporters help the plants to resist against bacterial and fungal infections, the study of their substrates and the phenotypes of mutant plants shed light on the transporter-associated immune system of plants. ABCG31 in *Hordeum vulgare* (barley), as well as ABCG11 and ABCG12 in *A. thaliana*, help the formation of the cuticle, which can be seen as the first protective barrier of the plant [31–33], thus indirectly contributing to the protection of the plant against stresses and infections. Another hypothesis would be the excretion of toxic secondary metabolites; the already mentioned NpPDR1 was found to secrete sclareolide, an antifungal diterpene, on the leaf surface [34]. Also, flavonoids have recognized antimicrobial and antioxidative properties [35, 36], and

it was recently shown that ABCG10 in *Medicago truncatula* (MtABCG10) modulates the isoflavonoid levels [37]. Anyway, suppression of MtABCG10 leads to a lower resistance of the plant to the root pathogen *Fusarium oxysporum*.

Hormone transport

Plant ABC transporters of the B and G family are also involved in the transport of phytohormones like auxins, abscisic acid, cytokinins, or strigolactones. The details are explained in the next section.

Phytohormone ATP-Binding Cassette Transporters
▼
Auxin transport

The phytohormone auxin (whose main natural form is indole-3-acetic acid, IAA) is involved in root development [38], vascular differentiation [39], cell wall construction [40], plant growth [41], as well as in response to light (reviewed in [42]) or gravity [43]. Auxin transport has been extensively studied in the model plant *A. thaliana*. While the plant vascularization by xyleme and phloeme can carry phytohormones from the root to the stem apex or the reverse, lateral transport from biosynthesis loci (shoot apex, leaves, and roots [44]) to these vessels requires what is called the polar auxin transport [2]. The current model of polar auxin transport includes a family of import proteins, the permease-like AUX carriers [43]. In the acidic apoplastic compartment, around 20% of auxin is in its neutral, undissociated form and can therefore cross the plasmic membrane by passive diffusion [45]. On the other side, since a long time auxin efflux is thought to involve protein complexes [46]. More recent discoveries identified the plant-specific pin-formed (PIN) family [47] as an auxin export carrier. In addition, several members of the B subfamily of ABC transporters are also able to transport auxin: Pgp1 (ABCB1) and Pgp19 (ABCB19) [48], Pgp4 (ABCB4) in the roots [49], and Pgp14 (ABCB14) and Pgp15 (ABCB15) during the stem lignification [50].

The most studied ABC transporters, Pgp1 and Pgp19, seem to play a crucial role in polar auxin transport; co-expressed in a polarized manner with the PIN members PIN1 or PIN2, they stabilize their membrane trafficking and localization [51]. Pgps, in association with the PIN proteins, synergistically participate in auxin efflux after receiving regulatory signals from their protein partners of the immunophilin TWD family [52, 53]. Protein-protein binding occurs between Pgp and TWD, which increases the efflux activity of the entire system (○ **Fig. 2**). Such protein-protein binding concerning an ABC transporter and a soluble immunophilin is not well studied in humans, and the potential impact of ABC transporters in membrane trafficking of other membrane proteins probably deserves some attention in humans, too.
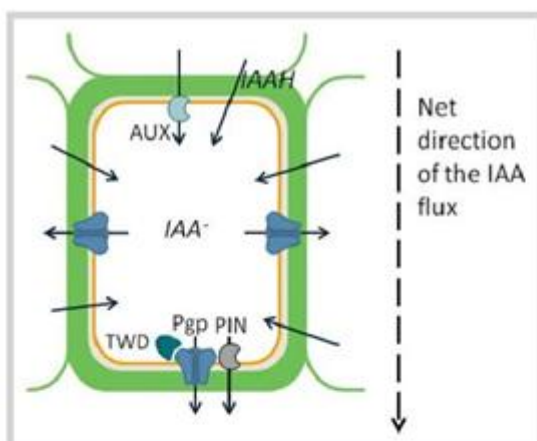
Fig. 2 Schema of directional auxin flux and role of the different transporters. Adapted from Blakeslee et al. [44]. (Color figure available online only.)

However, Pgps also have a nonpolar localization in small meristematic cells where they oppose the passive rediffusion of protonated auxin from the membrane, which otherwise could disrupt the polar transport. In those lateral membranes, Pgps would therefore not be linked to PIN proteins, but function as individual auxin efflux transporters (○ Fig. 2) [44].

Recently, a specific role of Pgp19 in phototropism was discovered by Christie et al. [54]. Phototropism is the asymmetric stem growth phenomenon occurring when the plant is only partially enlightened. In this case, an auxin gradient is created that makes the distal side of the stem grow faster than the enlightened side, resulting in a bending of the stem. The protein kinase phot1 is sensitive to light, and, upon illumination, will phosphorylate the cytosolic C terminal region of Pgp19, resulting in an inhibition of auxin efflux. The subsequent auxin accumulation will then serve for the lateral flux towards the shadowy side of the stem. The authors hypothesize that the phosphorylation actually inhibits the interaction between Pgp19 and its activator TWD1.

A parallel phenomenon has been reported for Pgp1. The serine-threonine protein kinase PINOID (PID) is a partner of the TWD1-Pgp efflux complex. In the absence of TWD1, phosphorylation of Pgp1 by PID leads to a higher activity of auxin efflux. But when TWD1 is bound to its ABC transporter partner, the result of PID action is a complete abolition of auxin transport [55].

Polar transport of the plant hormone auxin is controlled by PIN- and ABCB/PGP efflux catalysts. PIN polarity is regulated by the AGC protein kinase, PINOID (PID), while ABCB activity was shown to be dependent on the interaction with the FKBP42, TWISTED DWARF1 (TWD1). Using co-immunoprecipitation (co-IP) and shotgun LC-MS/MS analysis, PID was identified as a valid partner in the interaction with TWD1. *In vitro* and yeast expression analyses indicated that PID specifically modulates ABCB1-mediated auxin efflux in an action that is dependent on its kinase activity and that is reverted by quercetin binding and thus the inhibition of PID autophosphorylation. Triple ABCB1/PID/TWD1 cotransfection in tobacco (*Nicotiana benthamiana*) revealed that PID enhances ABCB1-mediated auxin efflux, but blocks ABCB1 in the presence of TWD1. Phosphoproteomic analyses identified S634 as a key residue of the regulatory ABCB1 linker and a very likely target of PID phosphorylation that determines both transporter drug binding and activity. In summary, this study provides evidence that PID phosphorylation has a dual, counteractive impact on ABCB1 activity that is coordinated by the TWD1-PID interaction.

Auxin transport is not exclusively regulated by a protein-protein interaction and phosphorylation of Pgps. Endogenous and exogenous small molecules are also known to inhibit the auxin transporters. The family of phytotropins was described in the 70s as auxin transport inhibitors. They share a 2-carboxyphenyl group linked to another aromatic system [56]. Selected structures of phytotropins are shown in ○ Fig. 3, notably the 1-N-naphthyl-phtalamic acid (NPA), the cyclopropyl propane dione (CPD), the 2-(1-pyrenoyl)benzoic acid (BPA), and DPX1840. The pharmacophoric features elucidated by Katekar comprise a carboxylic acid moiety (or equivalent), coplanarity of the two aromatic rings, and a distance of 7.3 A between the two aromatic rings. The common effect of the members of this family is a blockage of the polar auxin transport, a biphasic root growth inhibition, and a loss of gravitropism [57]. Katekar and Geissler were able to show that NPA has a weak auxin-like behavior at a low concentration, and that NPA and CPD share the same target. More recently, it was shown
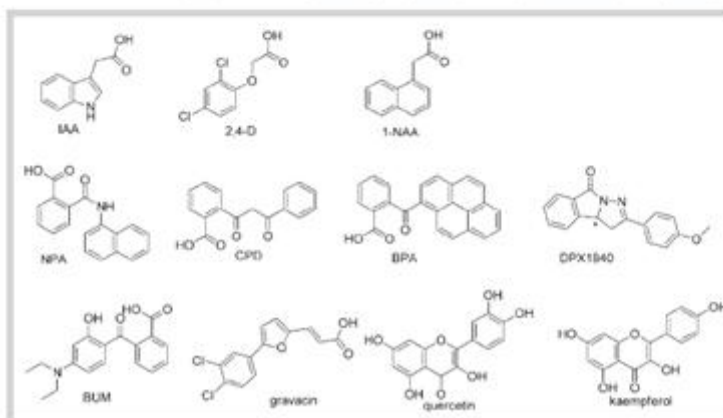
Fig. 3 Chemical structures of auxin and auxin transport inhibitors. IAA = indole-3-acetic acid, 2,4-D = 2,4-dichlorophenoxyacetic acid, 1-NAA = naphthalene-1-acetic acid, NPA = 1-N-naphthylphtalamic acid, CPD = cyclopropyl propane dione, BPA = 2-(1-pyrenoyl)benzoic acid, BUM = 2-[4-(diethylamino)-2-hydroxybenzoyl]benzoic acid.
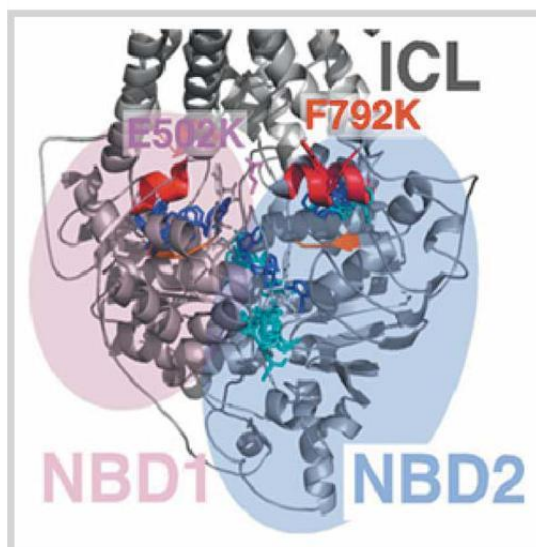
Fig. 4 *In silico* drug binding to the N- and C-terminal ABCB1 nucleotide-binding folds (NBD1 and − 2) suggest overlapping and distinct inhibitor binding pockets for BUM (cyan) and NPA (blue). Note that NPA docks to pockets flanked by coupling helices (red) and the Q loop (orange) of NBD1 and NBD2, whereas BUM docks only to the pocket corresponding to NBD2. Relevant residues Glu502 and Phe792 mutagenized under C are represented as pink and red sticks. Research as originally published in Kim et al. [60]; © American Society for Biochemistry and Molecular Biology. (Color figure available online only.)

that NPA acts by disrupting the protein-protein interaction between Pgp1 and TWD1 [53], that the loss of photo- and gravitropism in *A. thaliana* hypocotyls is due to an inhibition of Pgp19 [58], and that it is also an inhibitor of Pgp4 [59] and Pgp1 [52].

Docking into an homology model of Pgp1 revealed that NPA binds in the nucleotide binding domain (NBD) region at the interface with the transmembrane domain (○ **Fig. 4**) [60]. This is further supported by site-directed mutagenesis, which showed that

a mutation in this region (E502K) abolishes NPA binding. Kim and colleagues also presented a new Pgp inhibitor, 2-[4-(diethylamino)-2-hydroxybenzoyl]benzoic acid (BUM, ○ **Fig. 3**), which is 30 times more potent than NPA. BUM also binds to the NBDs and disrupts the interaction with the TWD partner.

Gravacin, the 3-(5-[3,4-dichlorophenyl]-2-furyl)-acrylic acid (○ **Fig. 3**), is an inhibitor of root and shoot gravitropism, inhibits the response to auxin [61,62], and specifically Pgp19 (and not Pgp1). Gravacin displaces 60% of NPA bound to Pgp19, which may indicate a partial overlap of their respective binding sites. However, gravacin does not disrupt the interaction between TWD and Pgp19.

Finally, flavonoids have also been shown to interact with plant ABC transporters. Those endogenous metabolites are synthesized widely among the plant kingdom and inhibit ABC transporters in mammals [63] via binding to the NBDs. In *A. thaliana* and *Cucurbita pepo* hypocotyls, quercetin, and kaempferol (○ **Fig. 3**) decrease auxin efflux in an NPA-competitive fashion [64].

Strikingly, although ABC transporters in mammalians are able to transport a wide range of substrates, the substrate specificity of auxin efflux Pgps in plants is quite high. They are only able to transport indole-3-acetic acid (IAA, the main form of auxin), 2,4-dichlorophenoxyacetic acid (2,4-D) and naphthalene-1-acetic acid (NAA; ○ **Fig. 3**; [48]). Bailly and colleagues [65] compared the properties of the translocation chambers of Pgp1 and Pgp19 obtained by homology modelling with the crystal structure of mouse Pgp (ABCB1, [66]). The authors observed very little conservation between plant and animal Pgps in terms of the residues exposed to the cavity. Plant auxin exporters share a common surface electrostatic pattern in the translocation chamber; near the entrance, the potentials are negative, then evolve towards neutral potentials close to the binding zones. The mammalian electrostatic surfaces are much more diverse, which may explain the diversity of substrates seen in mammalian ABC transporters (○ **Fig. 5**).

To summarize, several members of the B family of ABC transporters in plants have been shown to play an important role in polar auxin efflux, both by direct transport and by facilitating and regulating the export by PIN proteins. Phosphorylation of the nucleotide binding domain and protein-protein interactions may also modulate their activity. This mechanism is not well studied in human ABC transporters, but could be highly relevant

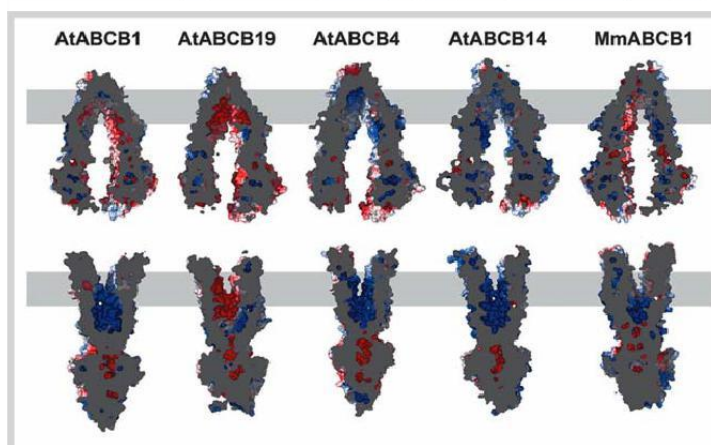**Fig. 5** Surface electrostatic potential of the ABCB translocation chamber reflects its substrate specificity. Cut views of *Arabidopsis* (At) and mouse (Mm) ABCB proteins' surface electrostatics in both inward- (up) and outward-facing conformers (down). The gray box represents the estimated position of the lipid bilayer. Figure as originally published in Bailly et al. [65]. (Color figure available online only.)
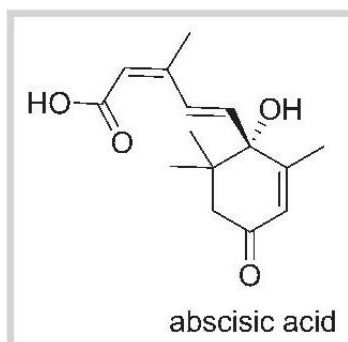
Fig. 6 Structure of abscisic acid.

there, too. The substrate specificity of plant ABC transporters is quite opposite to the promiscuity reported for animal ABC transporters, which might be due to specific electrostatic patterns in the translocation chambers.

Abscisic acid transport

Abscisic acid (ABA) is a plant hormone that regulates developmental aspects like germination [67], and modulates resistance to drought [68], high salinity [69], or even pathogen infection [70].

During drought, the levels of ABA increase in the plants, leading to a closure of the stomatal pore triggered by a change of shape of the guard cells. The closure of the stomata reduces water loss by transpiration. On the contrary, when the humidity levels are high, CYP707 enzymes are expressed that oxidize and degrade the stock of ABA. CYP707A3 reduces the amount of ABA in vascular tissues, while CYP707A1 acts in the guard cells [71].

For obvious agricultural applications, it is essentially the link between ABA and water stress resistance that has attracted the interest of researchers over the last decades. It is known that, upon water stress, ABA is synthesized in vascular parenchymal cells [72,73]. Given that the essential action site of ABA under drought conditions is in the guard cells, the question of the transport of ABA from the vascular companion cells to the stomata arises [74]. ABA is a weak acid (○ **Fig. 6**), thus, like auxin, it can diffuse passively into cells when it is in its undissociated form [75]. However, passive diffusion and local biosynthesis of ABA in the guard cell [73] can explain neither the amount of ABA found in guard cells nor the fast response of the plant upon stress signalling. In 2009, Kang and colleagues [75] identified the ABC transporter PDR12 (ABCG40) in *A. thaliana* as necessary for a timely response to drought in guard cells and also for normal seed germination. Localized in the plasma membrane, this protein is expressed in the guard cells, in the seedlings, and in the roots. PDR12 expression is ABA sensitive; when the plants are treated with ABA, the expression of PDR12 increases. *A. thaliana* mutants lacking PDR12 become more sensitive to drought than the wild-type plants.

The substrate specificity of PDR12 is high, as only the natural stereoisomer S-ABA can be transported, not the synthetic R-ABA or ABA-glucose-ester, auxin, or benzoic acid. However, ABA import to the guard cells is inhibited by the classical ABC transporter inhibitors glibenclamide and verapamil.

But how does ABA leave the phloem companion cells and xylem parenchyma cells where it is synthesized? In 2010, Kuromori and colleagues screened various *A. thaliana* mutant lines for germination phenotypes and identified ABCG25 as an ABA exporter [72,

76]. Upon overexpression of this ABC transporter, the germination growth is inhibited because of an accumulation of ABA in the seeds. The authors found that, in adult plants, ABCG25 is coexpressed in the vascular system with the enzymes that catalyze ABA biosynthesis. Also, ABCG25 seems highly specific for S-ABA [76].

Yet another ABC transporter seems to be involved in drought sensitivity, maybe via the transport of ABA: ABCG22, another half-transporter from the ABCG subfamily that was recently discovered in *A. thaliana* because of its link with water transpiration and drought resistance [77]. The gene is expressed at the plasma membrane of stem, fruit, flower, and leaf cells. When knocked out, the mutants lack drought resistance, but no direct proof of ABA transport has been shown yet for this transporter.

Cytokinin transport

Cytokinins (CK) are, as their name suggests, responsible for cytokinesis, which is the final stage in cell division. They are also involved in many other developmental processes of plants: delay of leaf senescence [78], control of root/shoot balance [79], transduction of nutritional signals [80], differentiation of plant cells [81], chloroplast development and chlorophyll retention [82], and stress responses [79]. Thus, the whole plant morphology is influenced by the endogenous level of CKs.

Kinetin (○ **Fig. 7**) was the first CK discovered in 1955. Until today there is no consensus about whether it occurs naturally or not [83,84]. There are two classes of CKs [85]: purines and phenylurea derivatives (○ **Fig. 7**). The former can be divided into isoprenoid and aromatic CKs (○ **Fig. 7**).

Depending on the nature of the side chains, the individual CKs have different activities, functions, or tissue localizations. In *A. thaliana*, for example, the main CKs are trans-zeatin and isopentenyl adenin (○ **Fig. 7**) [81]. Trans-zeatin is transported through the xylem [86], whereas isopentenyl adenin is the main CK in the phloem [87]. This different distribution of the individual agents in different tissues suggests that CKs act not only as local, but also as long-distance signals [88]. This finding suggests the presence of a transport system for CKs in plants. The involvement of the *A. thaliana* purine permease family (AtPUP) and most likely of a representative of the equilibrative nucleoside transporters (ENT) to the translocation of cytokinins has been known for quite some time [88,89].

The most actual findings report that an ABC transporter of the G family is involved in CK transport from root to shoot. A recent publication on this topic by Ko et al. describes that AtABCG14, mainly expressed in the root, is essentially involved in the CK transport to the shoot by allowing for the loading of CKs to the xylem sap, thus enabling all further translocations [90]. In their study, candidate CK transporters were selected by the expression of colocalization with CK biosynthesis genes in the root and other genes induced by CK treatment. Among the candidates, the seedlings with atabcg14 mutants showed a different shoot to root ratio, smaller leaves, and longer roots. The mature plants exhibited smaller rosette leaves, shorter and thinner stems, a lower number and size of xylem and phloem cells, smaller lignin level, and fewer seeds. All of these phenotypes could be reverted by external application of CKs. The atabcg14 mutants had reduced the root-to-shoot translocation of CKs, and the tZ-type CKs concentration in the xylem sap was reduced by more than 90%. Finally, grafting a mutant atabcg14 shoot onto a wild-type root caused recovery of the shoot growth. These experimental results show that root-synthesized CKs are essential for shoot growth.

11
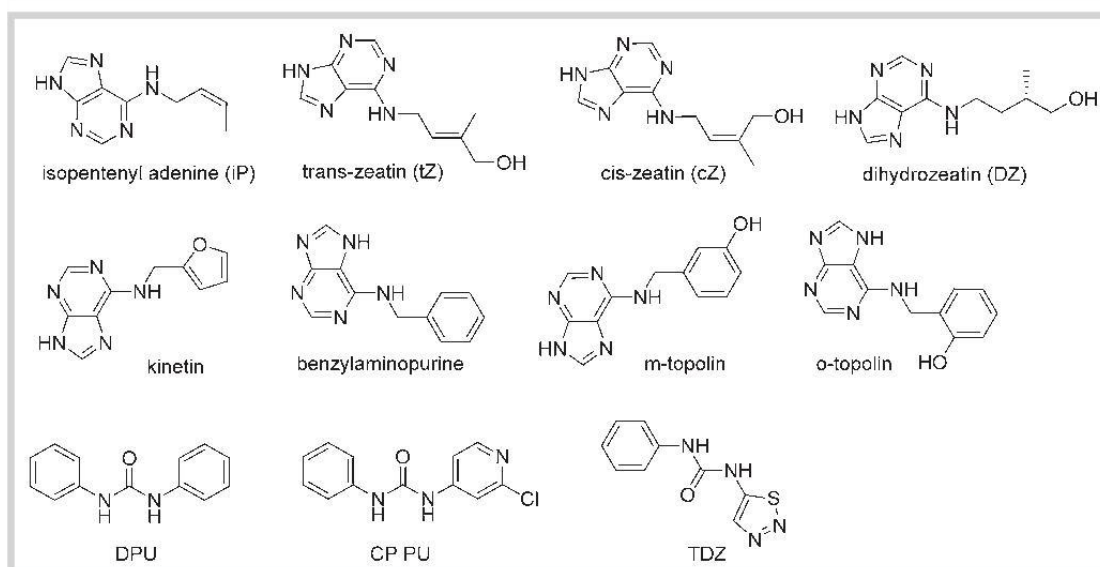
Fig. 7 Chemical structures of the most known cytokinins. DPU = 1,3-diphenylurea, CP PU = N-(2-chloro-4-pyridyl)-N'- phenylurea, TDZ = thidiazuron.
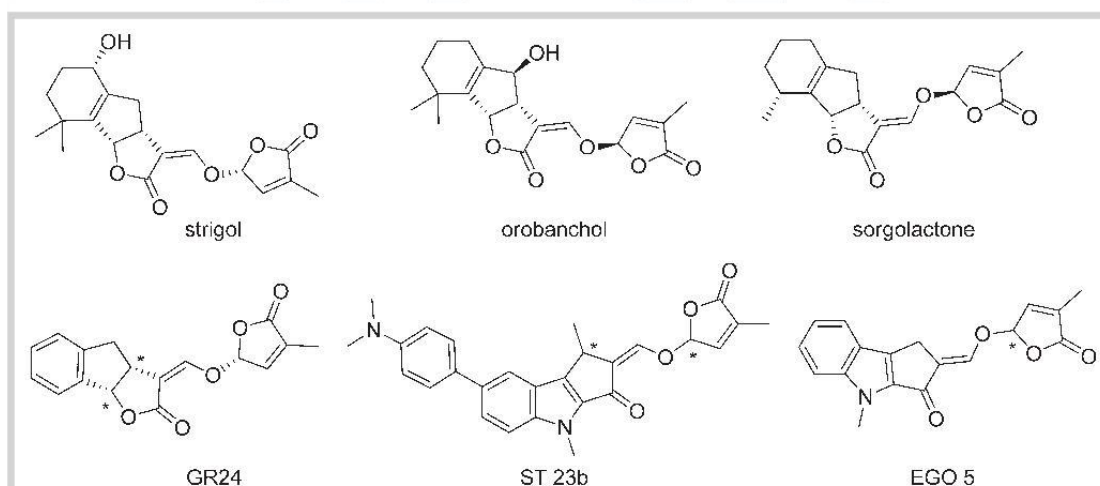
Fig. 8 Chemical structures of selected strigolactones. Upper row: naturally occurring SLs; lower row: synthetic analogues.

Strigolactone transport

Strigolactones (SLs) are carotenoid-derived plant hormones that play a central role in the regulation of shoot branching by suppressing the bud outgrowth activity [91,92].

The naturally occurring SLs [93] are divided into two major families, one with the BCD rings of strigol as a leading scaffold and the other one with the BCD rings of the natural orobanchol stereoisomer (○ **Fig. 8**). Furthermore, additional SLs have been described which do not belong to these two families [94]. In addition, numerous synthetic analogs of SLs are known, with GR24 (○ **Fig. 8**) having evolved as a reference compound. Finally, fluorescent analogs have also been developed, such as EGO 5 and ST 23b (○ **Fig. 8**) [95].

Besides their function in branching inhibition, SLs are growth activators of arbuscular mycorrhizal fungi (AMF) [96], which live in a symbiosis with most of the land plants. The greatest benefit for a plant from this symbiosis is made under low phosphate availability, and exactly under these conditions, SLs are exudated in an increased amount [91,97].

These findings were the basic foundation for Kretzschmar and colleagues to search for involved ABC transporters, more exactly the group of PDR transporters. PDRs are often found in roots [98],

Hellsberg E et al. The ABC of… Planta Med 2015; 81: 474–487

12

Fig. 9 PDR1-mediated strigolactones exudation from hypodermal passage cells (HPC) into the rhizosphere to enable an interaction with arbuscular mycorrhizal fungi (AMF). Adapted from Kretzschmar et al. [99] (see also Supporting Information). (Color figure available online only.)

Fig. 10 MmABCB1 open state, Sav1866 closed state. (Color figure available online only.)

and the activity of one PDR member, AtABCG30 (PDR2 in *A. thaliana*), is known to affect the soil microbial community [19]. There is also evidence of their affinity for compounds structurally related to SLs [99].
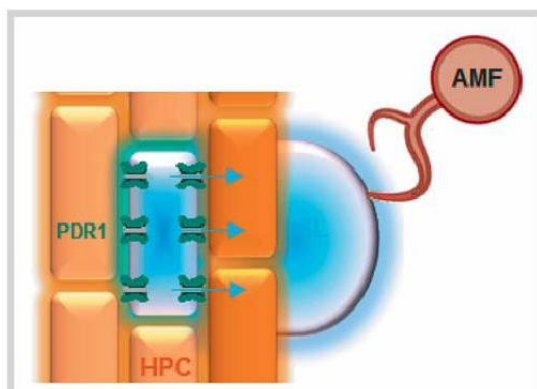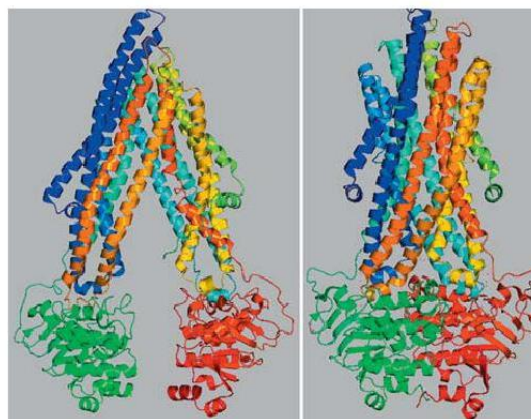
Six primary candidates were chosen and investigated for expression levels in roots under different phosphate concentrations or the amount of AMF colonization. This approach led to the discovery of an ABCG transporter, PDR1 in *P. hybrida*, the first protein associated with SL cell export [99]. The plasma membrane localization of PDR1 was confirmed through the fusion with GFP. The need for transport of a signalling molecule relevant for arbuscular mycorrhizal symbiosis was proven by the comparison of mutant and wild-type plants. The final conclusion is that PDR1 mediates the SL export from hypodermal passage cells into the rhizosphere (Fig. 9).

Apart from these findings, there is also evidence of SL transport through the xylem from the root, where SLs are mainly synthesized, to the shoot [100]. The required transport system remains unclear, which raises the question whether other ABC transporters could be involved in loading and unloading the SLs to and from the xylem.

Elucidation of the structure of a strigolactone transporter

The field of plant hormones and their transport is steadily expanding, and there is an urgent need for an understanding of the molecular basis of substrate and inhibitor interactions.

Yang and Murphy developed structural models of AtABCB4 and AtABCB19 in a comparative study of the auxin transporters PIN, ABCB (PGP), and AUX/LAX [101]. The bacterial ABC transporter Sav1866 served as a template for the models. Their results indicate that AtABCB4 has three binding sites for IAA, whereas AtABCB19 has only two. This might be due to the different transport directions, as AtABCB4 can act as an importer and exporter, whereas AtABCB19 is only an exporter.

As already mentioned in the auxin transport section, a more extensive modelling approach was published in 2012. The structures of AtABCB1, AtABCB4, AtABCB19, and AtABCB14 (in this study considered as a non-auxin transporter) were explored by

the development of homology models. Analysis of the models led to the hypothesis that the mammalian and plant ABCB transporters separated very early during evolution according to the structures of their binding sites [65]. MmABCB1 (*Mus musculus* P-glycoprotein, 3G5 U [102]) was used as a template for the open state and Sav1866 (*Staphylococcus aureus*, 2HYD [102]) for the closed state models (Fig. 10).

The models indicate that ABCB family members have characteristic translocation chambers in mammals and in plants. Docking studies improved the understanding of the relationships between the transport processes and the binding sites, and shed light into the substrate specificity and translocation mechanisms. This work could provide a basis for future research aimed at a structure-based design of inhibitors. However, it is worth mentioning that Kaneda et al. found AtABCB14 to be an auxin transporter, too [50].

A very similar approach was used in our group to elucidate the structure of the already mentioned PDR1 transporter in *P. hybrida*, which acts as an SL transporter. While all the previously modelled plant ABC transporters belong to the ABCB subfamily, PhPDR1 belongs to the ABCG subfamily, which exhibits a reverse topology to the other ABC transporters (Fig. 11). For this kind of protein, not a single crystal structure has been solved up to now, so finding a proper template is a challenging task.

Recently, Rutledge et al. published an exhaustive modelling approach for exploring the 3D structure of PDR5 in *Saccharomyces cerevisiae*. The TMD of Sav1866 and the NBD of hemolysin B were used as templates for the model in its open-to-out conformation. For the open to in state, they used the mouse Pgp structure [103]. The workflow we chose for our studies is strongly based on this work.

At first, the TMDs had to be defined. The knowledge about helix localizations in plant transporters is scant. Thus, the Uniprot database [104] indicates different numbers of transmembrane helices (between 12 for AtABCG30 to 14 for NtPDR1), and no information at all for the protein of interest, PhPDR1. The previously modelled plant transporters had the topologies of their particular templates, which means 6 helices – NBD – 6 helices – NBD for the ABCB subfamily. In our case the topology is not yet investi-
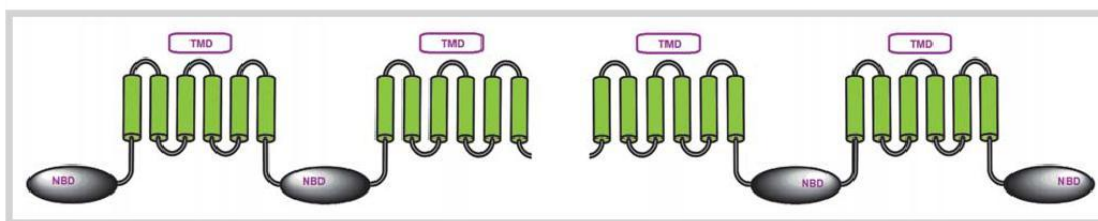
Fig. 11   Reverse topology of pleiotropic drug resistance transporters (left) in comparison to a regular topology (right). (Color figure available online only.)

Program	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	Reference
TopPred	527-547	555-575	610-630	640-660	668-688	694-714	755-775		1203-1223	1232-1252	1288-1308	1315-1335	1345-1365	1376-1396	1425-1445	Claros and von Heijne, 1994
DAS	525-542	556-572	606-628	640-648	667-686		756-776	1204-1219	1236-1249	1288-1302	1315-1336	1347-1364	1382-1392	1423-1445	Cserzo et al., 1997
TMPred	523-543	556-572	610-628		666-687	695-714	754-777	1202-1221	1232-1252		1312-1332	1345-1365	1377-1395	1422-1445	Hofman and Stoffel, 1993
SOSUI	524-546	559-581	607-629	635-657	666-688		757-779	1201-1223	1237-1259	1284-1306	1316-1338	1346-1368	1376-1398	1420-1442	Hirokawa et al., 1998
PRED-TMR	522-543	555-571	610-630		665-687	695-714	753-774	1205-1220	1232-1248	1290-1308	1315-1335	1345-1365	1376-1393	1426-1443	Pasquier et al., 1999
SPLIT-SERVER	522-543	555-581	596-630		663-868		752-777	1200-1220	1234-1249	1280-1310	1316-1339	1345-1366	1377-1392	1423-1445	Juretic et al., 2002
TMHMM	521-543	558-580	593-615		665-687		751-773	1201-1223	1236-1258	1286-1308	1315-1334	1344-1366	1373-1395	1423-1445	Krogh et al., 2001
SACS MEMSAT	522-543	553-569			671-687		753-777	1205-1223	1232-1248	1290-1308	1315-1335	1345-1365	1376-1392	1422-1445	Jones et al., 1994
MEMSAT-SVM	526-543	558-579	599-629	638-663	667-688		749-776	1203-1220	1236-1257	1276-1306	1314-1341	1345-1363	1376-1391	1424-1447	Nugent and Jones, 2009
PredictProtein	524-541	560-578	600-624	641-659	664-684	693-710	753-774	1202-1219	1237-1254	1286-1308	1313-1334	1344-1363	1344-1363	1425-1444	Rost et al., 2004

Fig. 12   Overview of tools and the predicted helix sections. (Color figure available online only.)

gated, and there is no template with a reverse topology available. Hence, we used several different software packages for predicting transmembrane helices and compared the predictions (Fig. 12) [105–114].

The different packages for predicting the secondary structure disagreed for a few helices, some predicting 13 or only 11 transmembrane helices. However, an odd number of helices implies that the N- and C-terminal ends are on different sides of the membrane, which seems quite unlikely for ABC transporters. In a next step, a multiple sequence alignment (MSA) of 11 sequences from plant, bacterial, and mammalian ABC transporters combining the current knowledge of transporter topologies was used to assign the missing helices, following the hypothesis of the common 6 + 6 topology.

The final choice of the template was based on an MSA with the potential templates (Fig. 13), including crystallized ABC transporters reported in the Protein Data Bank (PDB) [102] and the high-quality, reliably validated homology model of Pdr5 in *S. cerevisiae*. The latter was already used as a template to model the 3D structure of Cdr1 in *Candida albicans* (Rawal et al., 2013). As can be seen in Fig. 13, PDR5 shows the highest sequence identity percentage and was thus chosen as a template for further comparative modelling steps.

PDR5 and PDR1 were aligned pairwise in MOE [115], and the result was manually edited to align the respective predicted trans-

membrane helices. Gaps were allowed in the least conserved loop regions to accommodate for those changes. The final alignment (Figs. 1S and 2S, Supporting Information) with 24.9% overall identity of the whole sequence (TMD1 26.4%/NBD1 22.2%/TMD2 13.4%/NBD2 34.1%) was used as an input for Modeller 9.12 [116] to create ten models of the open-to-in conformation.

Modeller provides three validation scores: molpdf, the DOPE score (discrete optimized protein energy), and the GA341 score. According to these, the best three models were chosen for further validation. Using the Ramachandran plots and the G-factors computed using PROCHECK [117] on PDBsum [118], a final model with 91.2% and 88.9% of the amino acids in the most favored regions, respectively, for TMD1 and TMD2, as well as a G-factor of − 0.01 and − 0.16 for TMD1 and TMD2, was selected (Fig. 14) [119].

We further validated the model by comparing the respective positions of the proline residues in both the template and the model. Proline can cause transmembrane helix breaks or kinks. Therefore, proper alignment of these residues is required for a coherent helical geometry in the model. In the case of our model, no mismatches of prolines between template and target structure occurred.

Finally, the localizations of charged residues were investigated to assure that they do not point towards the membrane (Fig. 15).
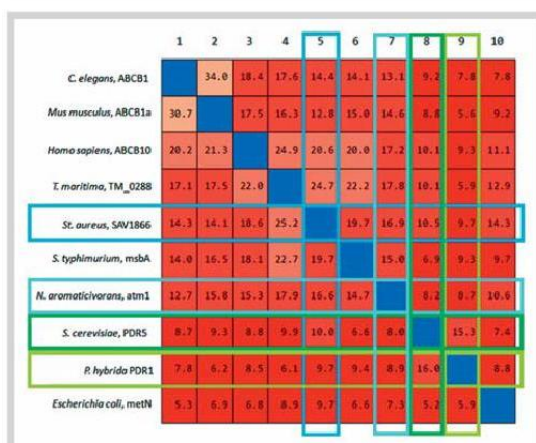
Fig. 13 Identity percentage matrix derived from the multiple sequence alignment of the possible templates with PhPDR1. PDR5 has the highest identity percentage [115]. (Color figure available online only.)
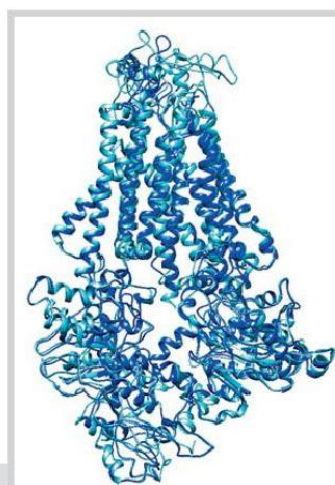
Fig. 14 The final model of PhPDR1 (dark blue) aligned to its template PDR5 (light blue) [119]. (Color figure available online only.)
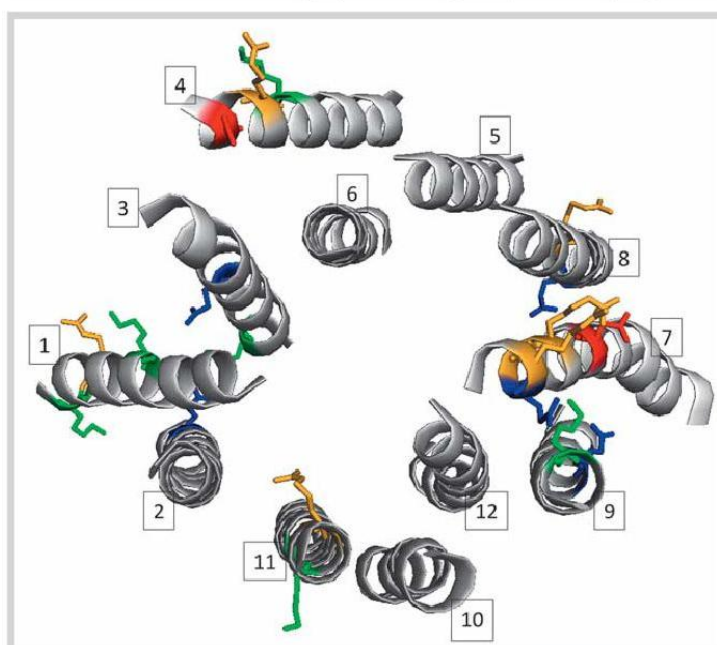
Fig. 15 Sight on top of the helices (1–12), examples for charged residues. Aspartate = red, glutamate = blue, lysine = green, arginine = yellow-orange. (Color figure available online only.)

As ○ Fig. 15 shows, the charged residues in helices 2, 5, 6, 9, 10, and 12 fit perfectly to their environments. To show how the others are embedded, they are presented from different views in ○ Fig. 16.

Lysine 517 and arginine 518 are on the very bottom of helix 1, so they might interact with the hydrophilic parts of the phospholipids. Lysine 526 and glutamate 615 probably form a hydrogen bond (○ Fig. 16 A). Aspartate 628 and arginine 635 might also interact with phospholipid head groups, whereas lysine 638 is directly pointing towards the membrane (○ Fig. 16 B), which is quite unlikely. In helix 7, glutamate 1225 and arginine 1226 are again near the top of the membrane, while the orientation of aspartate 1219 is also pointing towards the lipophilic part of the membrane. Arginine 1260 is close to the bottom of helix 8 (○ Fig. 16 C) and might also exhibit a different rotamer. Lysine 1365 is localized on the outer surface of the protein, a hydrophilic interaction on top here is conceivable, too (○ Fig. 16 D). Thus, this analysis shows that a few charged amino acids need in-depth analysis and that the model requires some modifications in these regions.
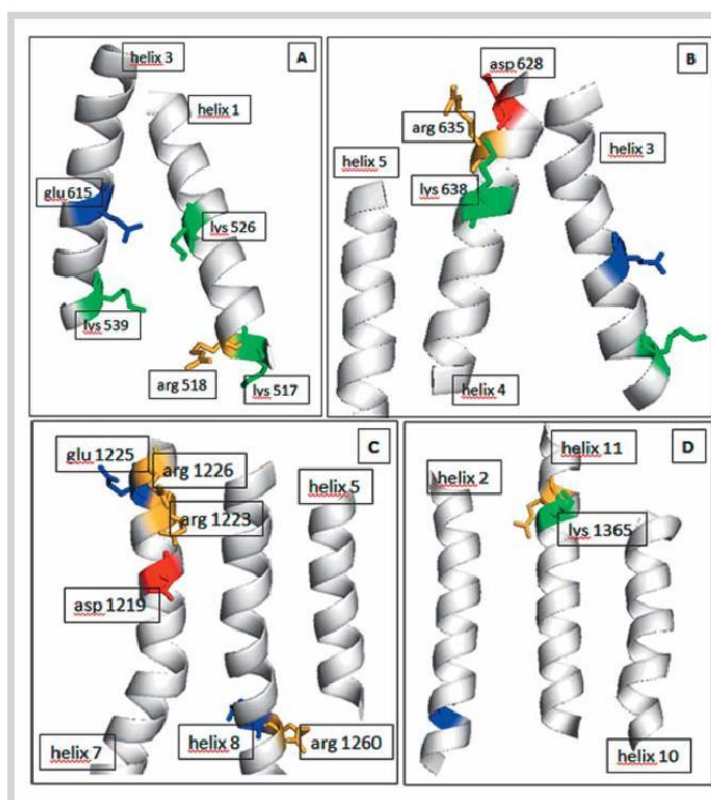
Fig. 16 A more detailed depiction of the charged residues in the single helices. (Color figure available online only.)
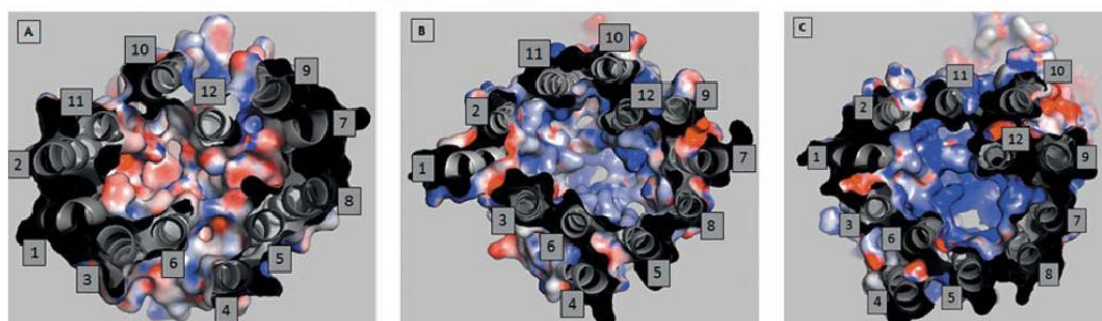
Fig. 17 Insight into the translocation chambers of AtABCB1 (A), PhPDR1 (B), and PDR5 in *Saccharomyces cervisiae* from the bottom view of the trans- membrane domains (numbered from 1–12). Isosurface colors: red = negative, white = neutral, blue = positive [120]. (Color figure available online only.)

Finally, the electrostatic potentials of the model were investigated in PyMOL [120] to characterize the translocation chamber. The surface was compared with those of the template PDR5 and AtABCB1. As mentioned before, AtABCB1 is an auxin transporter and shows a very mixed electrostatic potential surface in the translocation chamber (○ Fig. 17 A), whereas PDR5 forms a positive isosurface (○ Fig. 17 C). PhPDR1 represents a rather neutral, leaning to positive translocation chamber (○ Fig. 17 B).

After these first investigations, this model could now serve as starting point for further studies, such as molecular dynamics simulations and docking of SL analogs.

Conclusions and Outlook
▼
The understanding of the molecular basis of transport across cellular membranes and its modulation by small molecules becomes increasingly important. In this article, we present an overview

16

about the currently known ABC transporters in plants including their hitherto discovered functions such as detoxification, symbiosis and rhizosphere communication, metabolite transport, and protection against exogene dangers. Special attention was put on the translocation of phytohormones, including auxins, abscisic acid, cytokinins, and strigolactones. Although most of the hormones have been known for a long time and some of the transport processes are well defined, a molecular understanding of the transport process is still missing. Although our knowledge of transmembrane transport proteins grows constantly, the complexity of these sophisticated systems should not be underestimated. Nevertheless, the recently published X-ray structures of bacterial and mammalian ABC transporters allow for the development of comparative protein models of their plant analogs, which aid in the understanding of their structure and function. To demonstrate the usefulness of this approach, we built a homology model of PhPDR1, a strigolactone transporter from *P. hybrida*. These 3D models are a promising basis for future research as they offer the possibility to explore binding sites, discover potential substrates and inhibitors, and to understand their interactions with the protein. This definitely will also pave the way for the discovery of new agrochemicals.

Supporting information

The pairwise alignment of PDR5 and PDR1, which was used as input for Modeller [116] to calculate the models, is available as Supporting Information.

Conflict of Interest
▼

The authors declare no conflict of interest.

References

1 *Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, Martinoia E.* Plant ABC Transporters. Arabidopsis Book 2011; 9: e0153
2 *Robert HS, Friml J.* Auxin and other signals on the move in plants. Nat Chem Biol 2009; 5: 325–332
3 *Santner A, Estelle M.* Recent advances and emerging trends in plant hormone signalling. Nature 2009; 459: 1071–1078
4 *Kretzschmar T, Burla B, Lee Y, Martinoia E, Nagy R.* Functions of ABC transporters in plants. Essays Biochem 2011; 50: 145–160
5 *Martinoia E, Grill E, Tommasini R, Kreuz K, Amrhein N.* ATP-dependent glutathione S-conjugate "export" pump in the vacuolar membrane of plants. Nature 1993; 364: 247–249
6 *Lu YP, Li ZS, Rea PA.* AtMRP1 gene of *Arabidopsis* encodes a glutathione S-conjugate pump: isolation and functional definition of a plant ATP-binding cassette transporter gene. Proc Natl Acad Sci U S A 1997; 94: 8243–8248
7 *Lu YP, Li ZS, Drozdowicz YM, Hortensteiner S, Martinoia E, Rea PA.* AtMRP2, an *Arabidopsis* ATP binding cassette transporter able to transport glutathione S-conjugates and chlorophyll catabolites: functional comparisons with Atmrp1. Plant Cell 1998; 10: 267–282
8 *Tommasini R, Vogt E, Fromenteau M, Hörtensteiner S, Matile P, Amrhein N, Martinoia E.* An ABC-transporter of *Arabidopsis thaliana* has both glutathione-conjugate and chlorophyll catabolite transport activity. Plant J 1998; 13: 773–780
9 *Kim DY, Bovet L, Maeshima M, Martinoia E, Lee Y.* The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance. Plant J 2007; 50: 207–218
10 *Lee M, Lee K, Lee J, Noh EW, Lee Y.* AtPDR12 contributes to lead resistance in *Arabidopsis*. Plant Physiol 2005; 138: 827–836
11 *Mackenzie B, Takanaga H, Hubert N, Rolfs A, Hediger MA.* Functional properties of multiple isoforms of human divalent metal-ion transporter 1 (DMT1). Biochem J 2007; 403: 59–69
12 *Larsen PB, Geisler MJB, Jones CA, Williams KM, Cancel JD.* ALS3 encodes a phloem-localized ABC transporter-like protein that is required for aluminum tolerance in *Arabidopsis*. Plant J 2005; 41: 353–363
13 *Larsen PB, Cancel J, Rounds M, Ochoa V.* *Arabidopsis* ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminum toxic environment. Planta 2007; 225: 1447–1458
14 *Prell J, Poole P.* Metabolic changes of *Rhizobia* in legume nodules. Trends Microbiol 2006; 14: 161–168
15 *Sugiyama A, Shitan N, Yazaki K.* Involvement of a soybean ATP-binding cassette-type transporter in the secretion of genistein, a signal flavonoid in legume-*Rhizobium* symbiosis. Plant Physiol 2007; 144: 2000–2008
16 *Sugiyama A, Shitan N, Yazaki K.* Signaling from soybean roots to *Rhizobium*: An ATP-binding cassette-type transporter mediates genistein secretion. Plant Signal Behav 2008; 3: 38–40
17 *Priha O, Grayston SJ, Pennanen T, Smolander A.* Microbial activities related to C and N cycling and microbial community structure in the rhizospheres of *Pinus sylvestris*, *Picea abies* and *Betula pendula* seedlings in an organic and mineral soil. FEMS Microbiol Ecol 1999; 30: 187–199
18 *Mougel C, Offre P, Ranjard L, Corberand T, Gamalero E, Robin C, Lemanceau P.* Dynamic of the genetic structure of bacterial and fungal communities at different developmental stages of *Medicago truncatula* Gaertn. cv. Jemalong line J5. New Phytol 2006; 170: 165–175
19 *Badri DV, Quintana N, El Kassis EG, Kim HK, Choi YH, Sugiyama A, Verpoorte R, Martinoia E, Manter DK, Vivanco JM.* An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. Plant Physiol 2009; 151: 2006–2017
20 *Noctor G, Foyer CH.* Ascorbate and Glutathione: Keeping Active Oxygen Under Control. Annu Rev Plant Physiol Plant Mol Biol 1998; 49: 249–279
21 *Hinder B, Schellenberg M, Rodoni S, Ginsburg S, Vogt E, Martinoia E, Matile P, Hörtensteiner S.* How plants dispose of chlorophyll catabolites. Directly energized uptake of tetrapyrrolic breakdown products into isolated vacuoles. J Biol Chem 1996; 271: 27233–27236
22 *Nagy R, Grob H, Weder B, Green P, Klein M, Frelet-Barrand A, Schjoerring JK, Brearley C, Martinoia E.* The *Arabidopsis* ATP-binding cassette protein AtMRP5/AtABCC5 is a high affinity inositol hexakisphosphate transporter involved in guard cell signaling and phytate storage. J Biol Chem 2009; 284: 33614–33622
23 *Zolman BK, Silva ID, Bartel B.* The *Arabidopsis* pxa1 mutant is defective in an ATP-binding cassette transporter-like protein required for peroxisomal fatty acid beta-oxidation. Plant Physiol 2001; 127: 1266–1278
24 *Yazaki K, Shitan N, Takamatsu H, Ueda K, Sato F.* A novel *Coptis japonica* multidrug-resistant protein preferentially expressed in the alkaloid-accumulating rhizome. J Exp Bot 2001; 52: 877–879
25 *Shitan N, Bazin I, Dan K, Obata K, Kigawa K, Ueda K, Sato F, Forestier C, Yazaki K.* Involvement of CjMDR1, a plant multidrug-resistance-type ATP-binding cassette protein, in alkaloid transport in *Coptis japonica*. Proc Natl Acad Sci U S A 2003; 100: 751–756
26 *Sato F, Takeshita N, Fitchen JH, Fujiwara H, Yamada Y.* S-adenosyl-l-methionine: scoulerine-9-O-methyltransferase from cultured *Coptis japonica* cells. Phytochemistry 1993; 32: 659–664
27 *Badri DV, Loyola-Vargas VM, Broeckling CD, De-la-Peña C, Jasinski M, Santelia D, Martinoia E, Sumner LW, Banta LM, Stermitz F, Vivanco JM.* Altered profile of secondary metabolites in the root exudates of *Arabidopsis* ATP-binding cassette transporter mutants. Plant Physiol 2008; 146: 762–771
28 *Bultreys A, Trombik T, Drozak A, Boutry M.* *Nicotiana plumbaginifolia* plants silenced for the ATP-binding cassette transporter gene NpPDR1 show increased susceptibility to a group of fungal and oomycete pathogens. Mol Plant Pathol 2009; 10: 651–663
29 *Kobae Y, Sekino T, Yoshioka H, Nakagawa T, Martinoia E, Maeshima M.* Loss of AtPDR8, a plasma membrane ABC transporter of *Arabidopsis thaliana*, causes hypersensitive cell death upon pathogen infection. Plant Cell Physiol 2006; 47: 309–318

Electronic reprint for personal use

30 Campbell EJ, Schenk PM, Kazan K, Penninckx IAMA, Anderson JP, Maclean DJ, Cammue BPA, Ebert PR, Manners JM. Pathogen-responsive expression of a putative ATP-binding cassette transporter gene conferring resistance to the diterpenoid sclareol is regulated by multiple defense signaling pathways in Arabidopsis. Plant Physiol 2003; 133: 1272–1284

31 Bessire M, Borel S, Fabre G, Carraça L, Efremova N, Yephremov A, Cao Y, Jetter R, Jacquat AC, Métraux JP, Nawrath C. A member of the Pleiotropic Drug Resistance family of ATP binding cassette transporters is required for the formation of a functional cuticle in Arabidopsis. Plant Cell 2011; 23: 1958–1970

32 Panikashvili D, Savaldi-Goldstein S, Mandel T, Yifhar T, Franke RB, Höfer R, Schreiber L, Chory J, Aharoni A. The Arabidopsis DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. Plant Physiol 2007; 145: 1345–1360

33 Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, Kunst L, Samuels AL. Plant cuticular lipid export requires an ABC transporter. Science 2004; 306: 702–704

34 Jasiński M, Stukkens Y, Degand H, Purnelle B, Marchand-Brynaert J, Boutry M. A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion. Plant Cell 2001; 13: 1095–1107

35 Cushnie TPT, Lamb AJ. Recent advances in understanding the antibacterial properties of flavonoids. Int J Antimicrob Agents 2011; 38: 99–107

36 Jia Z, Zou B, Wang X, Qiu J, Ma H, Gou Z, Song S, Dong H. Quercetin-induced H(2)O(2) mediates the pathogen resistance against Pseudomonas syringae pv. Tomato DC3000 in Arabidopsis thaliana. Biochem Biophys Res Commun 2010; 396: 522–527

37 Banasiak J, Biala W, Staszków A, Swarcewicz B, Kepczynska E, Figlerowicz M, Jasinski M. A Medicago truncatula ABC transporter belonging to subfamily G modulates the level of isoflavonoids. J Exp Bot 2013; 64: 1005–1015

38 Okada K, Shimura Y. Modulation of root growth by physical stimuli. In: Meyerowitz EM, Somerville CR. Arabidopsis. Cold Spring Harbor: Laboratory Press; 1994: 665–684

39 Scarpella E, Marcos D, Friml J, Berleth T. Control of leaf vascular patterning by polar auxin transport. Genes Dev 2006; 20: 1015–1027

40 Cosgrove DJ. Growth of the plant cell wall. Nat Rev Mol Cell Biol 2005; 6: 850–861

41 Bonner J, Bandurski RS. Studies of the physiology, pharmacology, and biochemistry of the auxins. Annu Rev Plant Physiol 1952; 3: 59–86

42 Whippo CW, Hangarter RP. Phototropism: bending towards enlightenment. Plant Cell 2006; 18: 1110–1119

43 Marchant A, Kargul J, May ST, Muller P, Delbarre A, Perrot-Rechenmann C, Bennett MJ. AUX1 regulates root gravitropism in Arabidopsis by facilitating auxin uptake within root apical tissues. EMBO J 1999; 18: 2066–2073

44 Blakeslee JJ, Peer WA, Murphy AS. Auxin transport. Curr Opin Plant Biol 2005; 8: 494–500

45 Kramer EM, Bennett MJ. Auxin transport: a field in flux. Trends Plant Sci 2006; 11: 382–386

46 Morris DA, Rubery PH, Jarman J, Sabater M. Effects of inhibitors of protein synthesis on transmembrane auxin transport in Cucurbita pepo L. hypocotyl segments. J Exp Bot 1991; 42: 773–783

47 Petrásek J, Mravec J, Bouchard R, Blakeslee JJ, Abas M, Seifertová D, Wisniewska J, Tadele Z, Kubes M, Covanová M, Dhonukshe P, Skupa P, Benková E, Perry L, Krecek P, Lee OR, Fink GR, Geisler M, Murphy AS, Luschnig C, Zazímalová E, Friml J. PIN proteins perform a rate-limiting function in cellular auxin efflux. Science 2006; 312: 914–918

48 Geisler M, Blakeslee JJ, Bouchard R, Lee OR, Vincenzetti V, Bandyopadhyay A, Titapiwatanakun B, Peer WA, Bailly A, Richards EL, Ejendal KFK, Smith AP, Baroux C, Grossniklaus U, Müller A, Hrycyna CA, Dudler R, Murphy AS, Martinoia E. Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1. Plant J 2005; 44: 179–194

49 Cho M, Lee SH, Cho HT. P-glycoprotein4 displays auxin efflux transporter-like action in Arabidopsis root hair cells and tobacco cells. Plant Cell 2007; 19: 3930–3943

50 Kaneda M, Schuetz M, Lin BSP, Chanis C, Hamberger B, Western TL, Ehlting J, Samuels AL. ABC transporters coordinately expressed during lignification of Arabidopsis stems include a set of ABCBs associated with auxin transport. J Exp Bot 2011; 62: 2063–2077

51 Titapiwatanakun B, Blakeslee JJ, Bandyopadhyay A, Yang H, Mravec J, Sauer M, Cheng Y, Adamec J, Nagashima A, Geisler M, Sakai T, Friml J, Peer WA, Murphy AS. ABCB19/PGP19 stabilises PIN1 in membrane microdomains in Arabidopsis. Plant J 2009; 57: 27–44

52 Bouchard R, Bailly A, Blakeslee JJ, Oehring SC, Vincenzetti V, Lee OR, Paponov I, Palme K, Mancuso S, Murphy AS, Schulz B, Geisler M. Immunophilin-like TWISTED DWARF1 modulates auxin efflux activities of Arabidopsis P-glycoproteins. J Biol Chem 2006; 281: 30603–30612

53 Bailly A, Sovero V, Vincenzetti V, Santelia D, Bartnik D, Koenig BW, Mancuso S, Martinoia E, Geisler M. Modulation of P-glycoproteins by auxin transport inhibitors is mediated by interaction with immunophilins. J Biol Chem 2008; 283: 21817–21826

54 Christie JM, Yang H, Richter GL, Sullivan S, Thomson CE, Lin J, Titapiwatanakun B, Ennis M, Kaiserli E, Lee OR, Adamec J, Peer WA, Murphy AS. phot1 inhibition of ABCB19 primes lateral auxin fluxes in the shoot apex required for phototropism. PLoS Biol 2011; 9: e1001076

55 Henrichs S, Wang B, Fukao Y, Zhu J, Charrier L, Bailly A, Oehring SC, Linnert M, Weiwad M, Endler A, Nanni P, Pollmann S, Mancuso S, Schulz A, Geisler M. Regulation of ABCB1/PGP1-catalysed auxin transport by linker phosphorylation. EMBO J 2012; 31: 2965–2980

56 Katekar GF. Inhibitors of the geotropic response in plants: a correlation of molecular structures. Phytochemistry 1976; 15: 1421–1424

57 Katekar GF, Geissler AE. Auxin Transport Inhibitors: IV. Evidence of a common mode of action for a proposed class of auxin transport inhibitors: the phytotropins. Plant Physiol 1980; 66: 1190–1195

58 Nagashima A, Uehara Y, Sakai T. The ABC subfamily B auxin transporter AtABCB19 is involved in the inhibitory effects of N-1-naphthyphthalamic acid on the phototropic and gravitropic responses of Arabidopsis hypocotyls. Plant Cell Physiol 2008; 49: 1250–1255

59 Lewis DR, Miller ND, Splitt BL, Wu G, Spalding EP. Separating the roles of acropetal and basipetal auxin transport on gravitropism with mutations in two Arabidopsis multidrug resistance-like ABC transporter genes. Plant Cell 2007; 19: 1838–1850

60 Kim JY, Henrichs S, Bailly A, Vincenzetti V, Sovero V, Mancuso S, Pollmann S, Kim D, Geisler M, Nam HG. Identification of an ABCB/P-glycoprotein-specific inhibitor of auxin transport by chemical genomics. J Biol Chem 2010; 285: 23309–23317

61 Surpin M, Rojas-Pierce M, Carter C, Hicks GR, Vasquez J, Raikhel NV. The power of chemical genomics to study the link between endomembrane system components and the gravitropic response. Proc Natl Acad Sci U S A 2005; 102: 4902–4907

62 Rojas-Pierce M, Titapiwatanakun B, Sohn EJ, Fang F, Larive CK, Blakeslee J, Cheng Y, Cutler SR, Cuttler S, Peer WA, Murphy AS, Raikhel NV. Arabidopsis P-glycoprotein19 participates in the inhibition of gravitropism by gravacin. Chem Biol 2007; 14: 1366–1376

63 Conseil G, Baubichon-Cortay H, Dayan G, Jault JM, Barron D, Di Pietro A. Flavonoids: a class of modulators with bifunctional interactions at vicinal ATP- and steroid-binding sites on mouse P-glycoprotein. Proc Natl Acad Sci U S A 1998; 95: 9831–9836

64 Jacobs M, Rubery PH. Naturally occurring auxin transport regulators. Science 1988; 241: 346–349

65 Bailly A, Yang H, Martinoia E, Geisler M, Murphy AS. Plant lessons: exploring ABCB functionality through structural modeling. Front Plant Sci 2012; 2: 108

66 Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL, Chang G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science 2009; 323: 1718–1722

67 Groot SP, Karssen CM. Dormancy and germination of abscisic acid-deficient tomato seeds: studies with the sitiens mutant. Plant Physiol 1992; 99: 952–958

68 Marshall JG, Scarratt JB, Dumbroff EB. Induction of drought resistance by abscisic acid and paclobutrazol in jack pine. Tree Physiol 1991; 8: 415–421

69 Saeedipour S. Salinity tolerance of rice lines related to endogenous abscisic acid (ABA) level synthesis under stress. AJPS 2011; 5: 628–633

70 Finkelstein RR, Gampala SSL, Rock CD. Abscisic acid signaling in seeds and seedlings. Plant Cell 2002; 14 (Suppl. 1): S15–S45

71 Okamoto M, Hanada A, Kamiya Y, Yamaguchi S, Nambara E. Measurement of abscisic acid and gibberellins by gas chromatography/mass spectrometry. Methods Mol Biol 2009; 495: 53–60

72 Kuromori T, Miyaji T, Yabuuchi H, Shimizu H, Sugimoto E, Kamiya A, Moriyama Y, Shinozaki K. ABC transporter AtABCG25 is involved in abscisic acid transport and responses. Proc Natl Acad Sci U S A 2010; 107: 2361–2366

73 Koiwai H, Nakaminami K, Seo M, Mitsuhashi W, Toyomasu T, Koshiba T. Tissue-specific localization of an abscisic acid biosynthetic enzyme, AAO3, in Arabidopsis. Plant Physiol 2004; 134: 1697–1707

18

74 Schachtman DP, Goodger JQD. Chemical root to shoot signaling under drought. Trends Plant Sci 2008; 13: 281–287

75 Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, Lee Y. PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. Proc Natl Acad Sci U S A 2010; 107: 2355–2360

76 Kuromori T, Sugimoto E, Shinozaki K. Intertissue signal transfer of abscisic acid from vascular cells to guard cells. Plant Physiol 2014; 164: 1587–1592

77 Kuromori T, Sugimoto E, Shinozaki K. Arabidopsis mutants of AtABCG22, an ABC transporter gene, increase water transpiration and drought susceptibility. Plant J 2011; 67: 885–894

78 Gan S, Amasino RM. Inhibition of leaf senescence by autoregulated production of cytokinin. Science 1995; 270: 1986–1988

79 Werner T, Motyka V, Strnad M, Schmülling T. Regulation of plant growth by cytokinin. PNAS 2001; 98: 10487–10492

80 Samuelson ME, Larsson CM. Nitrate regulation of zeatin riboside levels in barley roots: effects of inhibitors of N assimilation and comparison with ammonium. Plant Science 1993; 93: 77–84

81 Sakakibara H. Cytokinins: activity, biosynthesis, and translocation. Annu Rev Plant Biol 2006; 57: 431–449

82 Chory J, Reinecke D, Sim S, Washburn T, Brenner M. A role for cytokinins in de-etiolation in Arabidopsis (det mutants have an altered response to cytokinins). Plant Physiol 1994; 104: 339–347

83 George EF, Hall MA, Klerk GJD. Plant growth regulators II: Cytokinins, their analogues and antagonists. In: George EF, Hall MA, Klerk GJD, editors. Plant propagation by tissue culture. Amsterdam: Springer; 2008: 205–226

84 Barciszewski J, Siboska GE, Pedersen BO, Clark BFC, Rattan SIS. Evidence for the presence of kinetin in DNA and cell extracts. FEBS Letters 1996; 393: 197–200

85 Skoog F, Armstrong DJ. Cytokinins. Annu Rev Plant Physiol 1970; 21: 359–384

86 Kiba T, Takei K, Kojima M, Sakakibara H. Side-chain modification of cytokinins controls shoot growth in Arabidopsis. Dev Cell 2013; 27: 452–461

87 Corbesier L, Prinsen E, Jacqmard A, Lejeune P, Onckelen HV, Périlleux C, Bernier G. Cytokinin levels in leaves, leaf exudate and shoot apical meristem of Arabidopsis thaliana during floral transition. J Exp Bot 2003; 54: 2511–2517

88 Hirose N, Takei K, Kuroha T, Kamada-Nobusada T, Hayashi H, Sakakibara H. Regulation of cytokinin biosynthesis, compartmentalization and translocation. J Exp Bot 2008; 59: 75–83

89 Gillissen B, Bürkle L, André B, Kühn C, Rentsch D, Brandl B, Frommer WB. A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in Arabidopsis. Plant Cell 2000; 12: 291–300

90 Ko D, Kang J, Kiba T, Park J, Kojima M, Do J, Kim KY, Kwon M, Endler A, Song WY, Martinoia E, Sakakibara H, Lee Y. Arabidopsis ABCG14 is essential for the root-to-shoot translocation of cytokinin. PNAS 2014; 111: 7150–7155

91 Umehara M, Hanada A, Yoshida S, Akiyama K, Arite T, Takeda-Kamiya N, Magome H, Kamiya Y, Shirasu K, Yoneyama K, Kyozuka J, Yamaguchi S. Inhibition of shoot branching by new terpenoid plant hormones. Nature 2008; 455: 195–200

92 Gomez-Roldan V, Fermas S, Brewer PB, Puech-Pagès V, Dun EA, Pillot JP, Letisse F, Matusova R, Danoun S, Portais JC, Bouwmeester H, Bécard G, Beveridge CA, Rameau C, Rochange SF. Strigolactone inhibition of shoot branching. Nature 2008; 455: 189–194

93 Boyer FD, de Saint Germain A, Pillot JP, Pouvreau JB, Chen VX, Ramos S, Stévenin A, Simier P, Delavault P, Beau JM, Rameau C. Structure-activity relationship studies of strigolactone-related molecules for branching inhibition in garden pea: molecule design for shoot branching. Plant Physiol 2012; 159: 1524–1544

94 Zwanenburg B, Pospíšil T. Structure and activity of strigolactones: new plant hormones with a rich future. Mol Plant 2013; 6: 38–62

95 Prandi C, Occhiato EG, Tabasso S, Bonfante P, Novero M, Scarpi D, Bova ME, Miletto I. New potent fluorescent analogues of strigolactones: synthesis and biological activity in parasitic weed germination and fungal branching. Eur J Org Chem 2011; 2011: 3781–3793

96 Akiyama K, Matsuzaki K, Hayashi H. Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. Nature 2005; 435: 824–827

97 López-Ráez JA, Charnikhova T, Gómez-Roldán V, Matusova R, Kohlen W, De Vos R, Verstappen F, Puech-Pages V, Bécard G, Mulder P, Bouwmeester H. Tomato strigolactones are derived from carotenoids and their biosynthesis is promoted by phosphate starvation. New Phytol 2008; 178: 863–874

98 Moons A. Transcriptional profiling of the PDR gene family in rice roots in response to plant growth regulators, redox perturbations and weak organic acid stresses. Planta 2008; 229: 53–71

99 Kretzschmar T, Kohlen W, Sasse J, Borghi L, Schlegel M, Bachelier JB, Reinhardt D, Bours R, Bouwmeester HJ, Martinoia E. A petunia ABC protein controls strigolactone-dependent symbiotic signalling and branching. Nature 2012; 483: 341–344

100 Kohlen W, Charnikhova T, Liu Q, Bours R, Domagalska MA, Beguerie S, Verstappen F, Leyser O, Bouwmeester H, Ruyter-Spira C. Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonarbuscular mycorrhizal host Arabidopsis. Plant Physiol 2011; 155: 974–987

101 Yang H, Murphy AS. Functional expression and characterization of Arabidopsis ABCB, AUX 1 and PIN auxin transporters in Schizosaccharomyces pombe. Plant J 2009; 59: 179–191

102 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucl Acids Res 2000; 28: 235–242

103 Rutledge RM, Esser L, Ma J, Xia D. Toward understanding the mechanism of action of the yeast multidrug resistance transporter Pdr5 p: a molecular modeling study. J Struct Biol 2011; 173: 333–344

104 The UniProt Consortium. Activities at the Universal Protein Resource (UniProt). nucleic acids res. 42: D191–D198 (2014). Available at http://uniprot.org. Accessed July 1, 2014

105 Claros MG, von Heijne G. TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci 1994; 10: 685–686

106 Cserzö M, Wallin E, Simon I, von Heijne G, Elofsson A. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 1997; 10: 673–676

107 Hofman K, Stoffel W. TMbase – a database of membrane spanning proteins segments. Biol Chem Hoppe Seyler 1993; 347: 166

108 Hirokawa T, Boon-Chieng S, Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 1998; 14: 378–379

109 Pasquier C, Promponas VJ, Palaios GA, Hamodrakas JS, Hamodrakas SJ. A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. Protein Eng 1999; 12: 381–385

110 Juretić D, Zoranić L, Zucić D. Basic charge clusters and predictions of membrane protein topology. J Chem Inf Comput Sci 2002; 42: 620–632

111 Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001; 305: 567–580

112 Jones DT, Taylor WR, Thornton JM. A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry 1994; 33: 3038–3049

113 Nugent T, Jones DT. Transmembrane protein topology prediction using support vector machines. BMC Bioinformatics 2009; 10: 159

114 Rost B, Yachdav G, Liu J. The PredictProtein server. Nucleic Acids Res 2004; 32: W321–W326

115 Molecular Operating Environment (MOE), 2013.08. Montreal, Canada: Chemical Computing Group, Inc.; 2013

116 Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics 2006; Chapter 5: Unit 5.6

117 Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26: 283–291

118 Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM. PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends Biochem Sci 1997; 22: 488–490

119 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004; 25: 1605–1612

120 The PyMOL Molecular Graphics System, Version 1.6. Schrödinger, LLC. Available at http://www.pymol.org. Accessed January 1, 2014

Hellsberg E et al. The ABC of … Planta Med 2015; 81: 474–487

19

2 Aim of the thesis

As outlined above, ABC transporters play manifold roles in plants. Thus, the aim of this thesis was to establish a protein model of the PDR1 strigolactone transporter in *Petunia hybrida* by comparative modelling. After thorough validation, the model should be used to identify the locations of possible binding sites for strigolactone type substrates.

3 Methods

3.1 Transmembrane domain (TMD) prediction

As PhPDR1 belongs to the Pleiotropic Drug Resistance subfamily (within the ABCG family), it shares the common topology of these proteins. They are comprised of 2 nucleotide binding domains (NBDs) and two TMDs containing 6 transmembrane helices (TMHs) each (Fig. 1).
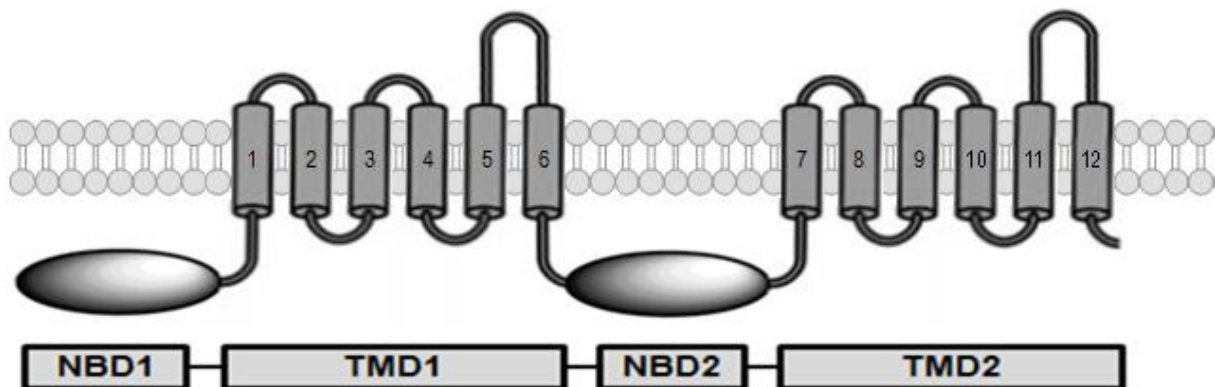
Fig. 1: The reverse topology of PDR and ABCG transporters (schematic illustration).

Due to the lack of information about the localisation of the transmembrane helices (TMHs) for the protein of interest, we used 10 prediction tools to discriminate TMHs from other structural elements in the protein and compared the results.

Prediction tools: The TMHs show different properties than other structural elements of a protein. For example, as they are spanning through the cell membrane, they are surrounded by a very lipophilic environment. Accordingly, the attributes of the amino acids in this area will be rather hydrophobic. The helices themselves are longer than just the membrane; they can extend into the intra- or extracellular compartments. Therefore, it is very important to use tools which do not only predict the secondary

structure of the protein but are also able to determine the amino acids (AA) in the membrane. The structural differences and much more acquired knowledge about proteins are exploited by the prediction tools in all possible ways (table 1).

All of the tools were used with their default options.

Prediction tool	website	Basis of the method
TopPred [1]	http://mobyle.pasteur.fr/	charge of residues, compositional distance method [2]
DAS [3]	http://www.sbc.su.se/	dot plots (query sequence against a database of membrane proteins)
TMPred [4]	http://www.ch.embnet.org/	statistical analysis of a database containing known proteins
SOSUI [5]	http://harrier.nagahama-i-bio.ac.jp/sosui/	hydropathy index [6], amphiphilicity index, index of AA charges
PRED-TMR [7]	http://athina.biol.uoa.gr/PRED-TMR/	hydrophobicity analysis
SPLIT-SERVER [8]	http://split.pmfst.hr/split/4/	basic charge motifs and positive residues
TMHMM [9]	http://www.cbs.dtu.dk/	hidden Markov Model
SACS MEMSAT [10]	http://www.sacs.ucsf.edu/cgi-bin/memsat.py	statistical tables (log likelihoods) of membrane protein data
MEMSAT-SVM [11]	http://bioinf.cs.ucl.ac.uk/psipred/?memsatsvm=1	support vector machine classifier
PredictProtein [12]	https://www.predictprotein.org/	neural network systems

table 1: list of the used programs, web presence and basis of functionality.

3.2 Sequence Alignments

Sequence alignments answer the purpose to analyze the similarity of strings of all kinds (here: protein sequences). A pairwise alignment compares two sequences by stringing them under each other and arranging them in a way to match as many identical amino acids as possible. The multiple sequence alignment (MSA) operates on the same basis, but compares 3 or more sequences with each other.

There are countless tools providing these kinds of algorithms. We tried a lot of them for experimental reasons; the ones which were relevant for our results are described below.

3.2.1 BLAST (Basic Local Alignment Search Tool)

We used the sequence of PhPDR1 as an input for a BLAST (Basic Local Alignment search tool) [13] search via Uniprot [14]. BLAST compares the input sequence with parts of the sequences in an existing database (here: UniProtKB) in form of a pairwise alignment. As an output, it delivers the most similar sequences according to the identity percentage.

BLAST is available via http://www.uniprot.org/blast/ and was used with its default options.

3.2.2 MOE (Molecular Operating Environment)

MOE [15] is a software for chemical computations and molecular modelling providing a huge set of different applications. The one for protein sequence alignments is called "Protein Align". The input format is fasta. The input sequences can be cut, changed in their order, renamed, labelled and colored. For the alignments, there are several settings to choose which enable possibilities for individual solutions. The most important one of these for our investigations were the gap penalties, which define the penalties for creating/extending a gap in the alignment. We increased the gap start penalty several times (for details, see chapter 4 Results and discussion). The software was used to cut several input sequences to the needed sections and to create pairwise alignments and multiple sequence alignments as well.

3.2.3 MAFFT

MAFFT [16,17] is a multiple alignment program for amino acid or nucleotide sequences based on fast Fourier transform (FFT). The AA sequence is translated into a string comprising the polarity and volume values of every AA.

MAFFT is available via http://mafft.cbrc.jp/alignment/server/

The input is a plain text file; it can be pasted or uploaded. The result is displayed on the website (Fig. 2) but is also available in fasta format.

Fig. 2: cutout of a MAFFT output of a query to align PhPDR1 and ScPDR5 (PDR5 in *Saccharomyces cerevisiae*). The sequences are stringed under each other. "-" in the sequence means a gap. Under the sequence alignment, the matches are characterized with "*" for identical AA and ":" for AA with similar properties.

Here again, several options can be chosen to match with individual problems. The parameter of particular importance was the gap opening penalty as well, which we adapted to our case (4 Results and discussion).

3.2.4 Jalview

Jalview [18] is a sequence alignment viewer. The alignments can be depicted in various colors and also manually edited, which is a great help for the analyses.
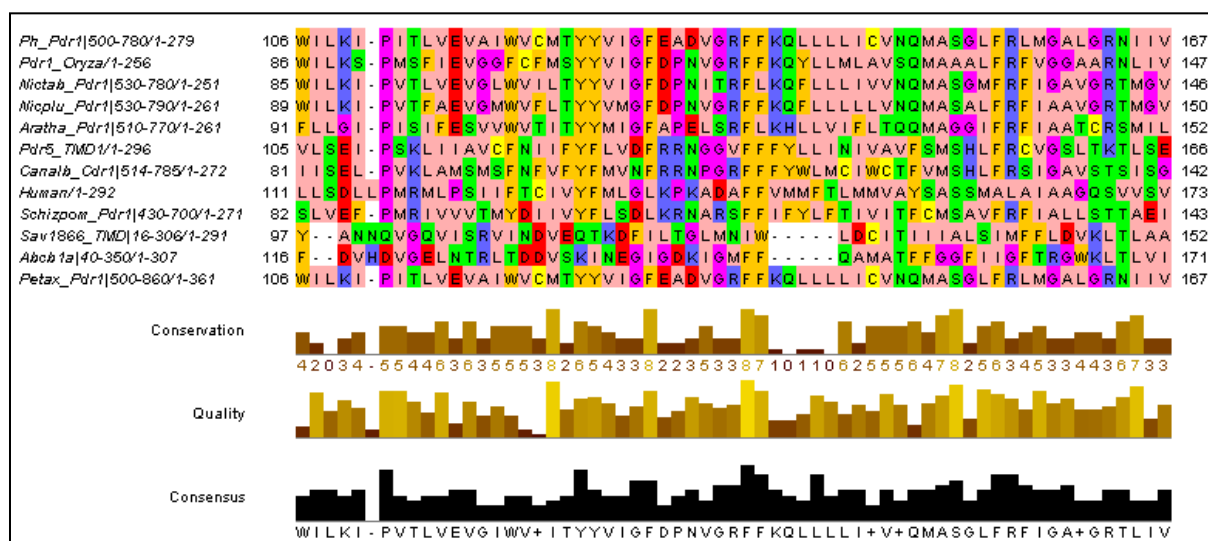
Fig. 3: cutout of the depiction of an MOE alignment in Jalview. In this color scheme, the AA are colored according to their properties (e.g. aromatic AA in yellow). Under the alignment, the conservation of the residues and the quality of the alignment are shown, as well as a consensus suggestion.

For this project, the software was used to analyze the different sequence alignments, to edit them and to improve the results according to the current issue.

3.3 Homology modelling - modeller software

The Modeller software [19] is used to create 3D models. As an input two files are needed: the alignment of template and target in ali format and the pdb file of the template. Finally, a python file containing the appropriate commands is necessary (Fig. 4).

```python
from modeller import *
from modeller.automodel import *

env = environ()

#env.io_data.atom_files_directory = ['.', 'C:\Users\Eva\Desktop\Thesis\model\']
a = automodel(env, alnfile='C:\Users\Eva\Desktop\Thesis\model\fifth_alignment/aln5_mod3_modi_140821.ali',
              knowns='Pdr5', sequence='Pdr1', assess_methods=(assess.DOPE, assess.GA341))
a.starting_model = 1
a.ending_model = 10
a.make()
```

Fig. 4: Input python file for Modeller.

We used Modeller version 9.12 in the default options. As an output we set 10 models (pdb files) including 3 assessment scores (already implemented in Modeller): molpdf, DOPE score and GA341.

Molpdf is an objective function and is automatically calculated. The "best" model has the lowest value. The GA341 score [20] is based on the sequence identity of the two proteins. It ranges from 0.0 to 1.0, the higher the value, the better; all in all it is the least recommended out of these three to choose the model. The DOPE score (Discrete Optimized Protein Energy [21]) is using a statistical potential which is optimized for model evaluation. It is the most reliable one of these three scores; the more negative the score, the better.

Modeller is available at https://salilab.org/modeller/about_modeller.html, information about documentation, tutorials, the manuals and a lot of further information can be found there.

3.4 Validation of the model

3.4.1 Ramachandran plots

The Ramachandran plot [22] is a crucial method to judge the quality of a protein 3D structure. The backbone of a protein is characterized by two dihedral angles (φ and ψ); the Ramachandran plot represents the distribution of their combinations in the protein backbone.

We used this method via an online service named PDBsum [23] provided by EMBL-EBI (The European Bioinformatics Institute, http://www.ebi.ac.uk/). PDBsum is available at http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html. On this page, a PDB-file can be uploaded and the Ramachandran plot, together with the related G-factors will be sent per email.

The Ramachandran plot statistics provided are divided into residues in most favoured, additional allowed, generously allowed and disallowed regions. According to the percentage distribution in these areas, the quality of the structure can be judged. A high quality model obtains more than 90% in the most favoured regions.

The G-factors measure the quality of given stereochemical properties; in this case certain properties of the dihedral angles. The values should be above -0,5.

In our approach, we used the Ramachandran plots at first to select the best model from the modeller output and later to control the validation steps to improve the model according these certain aspects (see 4.5 Validation of the model).

3.4.2 Protein Preparation Wizard

The Protein Preparation Wizard in the Schrödinger software [24] is a tool to improve the overall structure of a protein. It takes care of the H-bond assignment and performs a restrained energy minimization of the structure. We used the Protein Preparation Wizard to improve the overall structure of the chosen homology model and as an in-process control after the implementation of each new rotamer to the protein (see chapter 3.4.4.1 Rotamers).

The success of the method can be evaluated by Ramachandran plots, which provides a section with reported "bad contacts" between residues that are located too close to each other.

3.4.3 Proline and glycine residues

Prolines and glycines have to be treated with a particular attention in homology modelling as they are able to cause kinks or even breaks in helices. That means, if there is a proline or glycine in the transmembrane helix of the model but not of the template, the affected helix of the model will be probably buckled whereas the helix of the template is not and vice versa.

Therefore, we superposed the template and the model in the software PyMOL [25] and checked each single glycine and proline in the transmembrane domains to see how they match (for further details see 4.5.1 Proline and glycine residues).

3.4.4 Charged and polar residues

As a membrane protein is framed by a very lipophilic environment, it is quite obvious that charged residues will rather point to the pore of the protein. Therefore, checking

their orientations is a good method to validate the model. We describe this approach in depth in our paper (see 1 Introduction – The ABC of phytohormone translocation, section "Elucidation of the structure of a strigolactone transporter"). The same approach was done later for the polar residues as well.

3.4.4.1 Rotamers

For charged or polar residues that pointed directly into the membrane, various rotamers were calculated in Schrödinger Maestro [26]. The residue of interest has to be selected and by using the menu option "Rotamers" the programme delivers its suggestions of possible rotamers. In the default options (as it was used), 25 rotamers maximum are provided. Each rotamer was analyzed manually to determine which one fits best in every single case. The final chosen rotamer was examined in MOE [15] for possible clashes or H-bonds (for further details see 4.5.2 Charged and polar residues).

3.4.5 Electrostatic potentials

The aim of electrostatic potentials is to depict how the charges of the residues are spread over the protein and thereby to reflect negatively charged, neutral and positively charged areas throughout the structure. The derived information is very useful to characterize the translocation chamber of transporters because it enables the possibility to draw conclusions about the transported compounds.

In this study, the electrostatic potentials were calculated using the PyMOL [25] plugin APBS (Adaptive Poisson-Boltzmann Solver) [27]. Before the calculation, the structure has to be prepared with PDB2PQR [28,29] to add missing atoms, optimize the structure for favourable hydrogen bonding and more. The APBS itself calculates the solvation energies and electrostatic properties for visualization afterwards.

Both plugins were used in their default options.

3.5 Docking study

After completion of the homology model and the validation process, we wanted to take a first look on PhPDR1s binding mechanisms. Therefore, a rudimentary docking study was conducted with orobanchol, a plant hormone in the class of the strigolactones (see 1 Introduction – The ABC of phytohormone translocation, section "Strigolactone transport") which is transported by PhPDR1 [30].

The study was performed with Schrödinger Glide software [31,32], the details are explained below.

3.5.1 Ligand Preparation

The structure of orobanchol was prepared in Schrödinger LigPrep [33], which generates possible ionization states at a target pH of 7,0 ± 2,0, generates tautomers, stereoisomers and a low energy ring conformation. The SMILES code of the structure was extracted from chemspider [34] (www.chemspider.com) with specified chiral centers (see 2 Aim of the thesis, section "Strigolactone transport").

LigPrep was used in its default options.

3.5.2 Receptor grids

Before the docking approach, a grid for the calculation areas has to be set. In this case, the whole transmembrane region should be considered, as there is no information about any binding sites yet. According to the size restriction of the receptor grid, 2 boxes were placed to cover the whole section of interest. This was done by marking the involved amino acids in MOE [15] and inserting the extracted information to Glide's "Receptor Grid Generation" where the box residues can be specified.

3.5.3 Orobanchol docking poses

Orobanchol was docked into both receptor grids in two separated runs. The default options were used, except that the number of poses per ligand to include was set up to 100.

3.5.3.1 Clustering of the poses

The clustering of the poses was performed with the support of two in-house scripts. At first, an in-house implementation of a MOE [15] script was used to calculate the RMSD (root-mean-square deviation) matrix of the poses, according to which they were clustered afterwards at a defined niveau (n = 3) which corresponds to the maximal distance within a cluster in angstrom with the help of another in-house implementation of a script executed in the R software [35]. The underlying algorithm is a hierarchical cluster analysis on a set of dissimilarities and methods to analyze it.

3.5.3.2 involved amino acids

The poses in the considered clusters were analyzed one by one manually in MOE [15] to detect any interactions between ligand and protein.

4 Results and discussion

4.1 TMD definition

One main characterization of PDR proteins is their topology: they contain two nucleotide binding domains (NBDs) and 2 transmembrane domains (TMDs) (Fig. 1). If a homology model is created, it is very important that these structural elements are put in the right place - the NBDs into the intracellular side of the membrane and the TMDs spanning through the membrane. Therefore, the TMDs have to be clearly defined and their locations determined to be aligned in a favourable way. And as there is no information yet provided about PhPDR1s transmembrane domains, we used several prediction tools to identify TMDs. These tools are able to differentiate a TMD from other structural elements based on various algorithms (see 3.1, Transmembrane domain (TMD) prediction).

In Fig. 5, the results of the 10 prediction tools used for the protein topology determination are depicted. For the helices 1, 2, 3, 5 and 7 in TMD1, the results are

very consistent. This is not the case for helix 4, where only 5 out of 10 prediction tools identify a transmembrane region and similarly for helix 6, where only 4 tools define a transmembrane region. Apart from these findings, an odd number of helices is not likely for ABC transporters in plants. According to the current knowledge, their N- and C-termini as well as their nucleotide binding domains (NBDs) are located in the cytosol [36].

For TMD2, the predicted helices are even more compliant (Fig. 5). But again, we face the same problem as for TMD1: the prediction tools (except TMPred) define 7 helices, a number that is ruled out by the common ABC transporter topology.

For the whole sequence, the number of predicted TMHs varies from 11 to 14.

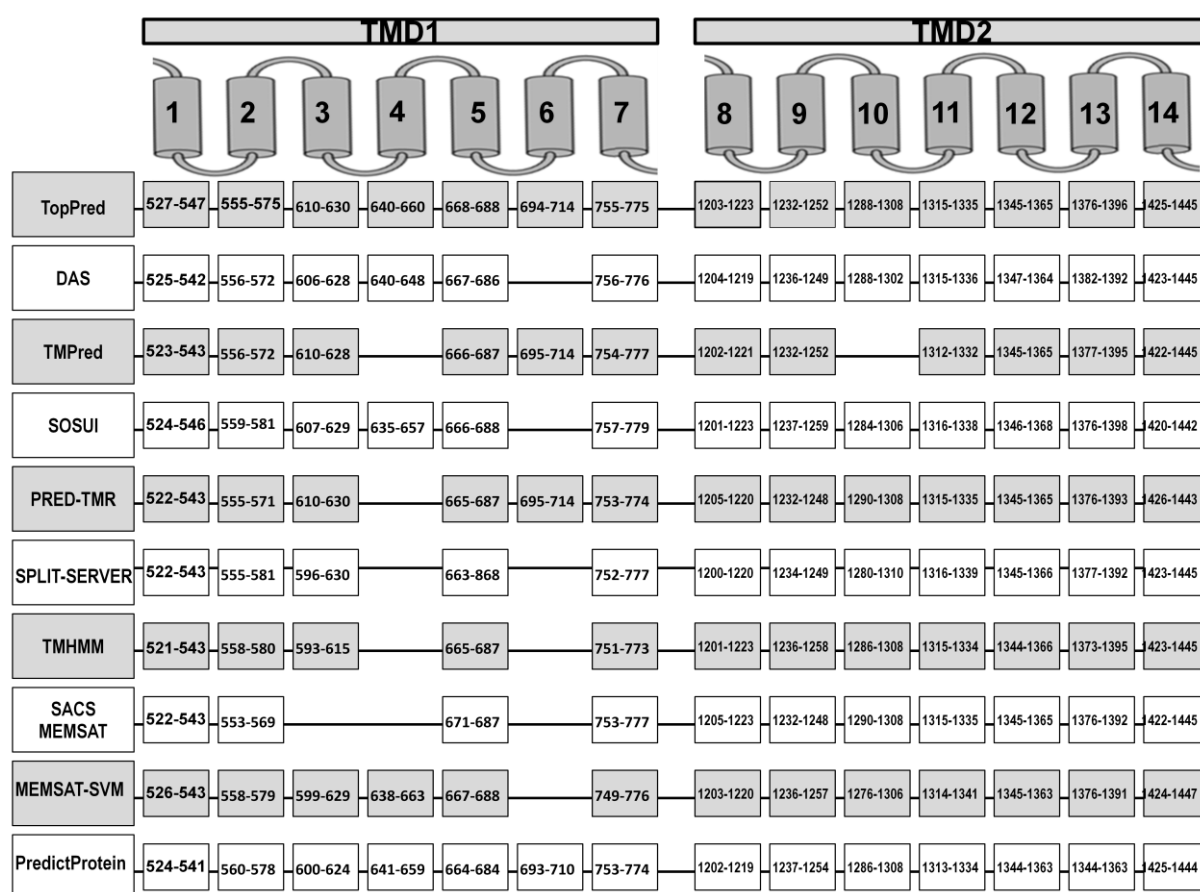|	TMD1							TMD2						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
TopPred	527-547	555-575	610-630	640-660	668-688	694-714	755-775	1203-1223	1232-1252	1288-1308	1315-1335	1345-1365	1376-1396	425-1445
DAS	525-542	556-572	606-628	640-648	667-686		756-776	1204-1219	1236-1249	1288-1302	1315-1336	1347-1364	1382-1392	423-1445
TMPred	523-543	556-572	610-628		666-687	695-714	754-777	1202-1221	1232-1252		1312-1332	1345-1365	1377-1395	422-1445
SOSUI	524-546	559-581	607-629	635-657	666-688		757-779	1201-1223	1237-1259	1284-1306	1316-1338	1346-1368	1376-1398	420-1442
PRED-TMR	522-543	555-571	610-630		665-687	695-714	753-774	1205-1220	1232-1248	1290-1308	1315-1335	1345-1365	1376-1393	426-1443
SPLIT-SERVER	522-543	555-581	596-630		663-868		752-777	1200-1220	1234-1249	1280-1310	1316-1339	1345-1366	1377-1392	423-1445
TMHMM	521-543	558-580	593-615		665-687		751-773	1201-1223	1236-1258	1286-1308	1315-1334	1344-1366	1373-1395	423-1445
SACS MEMSAT	522-543	553-569			671-687		753-777	1205-1223	1232-1248	1290-1308	1315-1335	1345-1365	1376-1392	422-1445
MEMSAT-SVM	526-543	558-579	599-629	638-663	667-688		749-776	1203-1220	1236-1257	1276-1306	1314-1341	1345-1363	1376-1391	424-1447
PredictProtein	524-541	560-578	600-624	641-659	664-684	693-710	753-774	1202-1219	1237-1254	1286-1308	1313-1334	1344-1363	1344-1363	425-1444

Fig. 5: results of the prediction tools: each box represents a predicted transmembrane region. The rows show the results of the single prediction tools. The columns present the different results for every certain transmembrane region (1-7 for TMD1, 8-14 for TMD2). If a box is missing in the scheme, there was no transmembrane region predicted by the corresponding software.

To summarize, the results of the 10 prediction tools deliver a first idea to find potential helix locations for PhPDR1.

Multiple Sequence Alignment (MSA) to solve the locations of the TMDs in PhPDR1

At first, we had to decide which proteins should be part of the MSA. Therefore, we analyzed the result of our BLAST query (see 3.2.1, BLAST (Basic Local Alignment Search Tool) and Fig. 6). The BLAST analysis delivered the 250 most similar proteins to PhPDR1, according to sequence identity. Unfortunately, some problems occurred:

1. 175 out of 250 entries were uncharacterized or putative proteins or both.
2. Many entries describe the same protein with minor differences mainly due to automated annotations [37] (remaining only 29 different proteins).
3. For most PDR proteins there is no information provided about their topology (remaining 6 entries with topology information in Uniprot).

Entry	Protein names	Match hit	Identity
H6WS94	ABCG/PDR subfamily ABC protein (Petunia hybrida)		100.0%
H6WS93	ABCG/PDR subfamily ABC transporter (Petunia axillaris)		99.7%
K4CLY8	Uncharacterized protein (Solanum lycopersicum)		91.5%
K4CLY7	Uncharacterized protein (Solanum lycopersicum)		89.3%
A0A022Q2U6	Uncharacterized protein (Erythranthe guttata)		81.0%
A0A061GNG0	Pleiotropic drug resistance 12 (Theobroma cacao)		80.2%
D7SUM9	Putative uncharacterized protein (Vitis vinifera)		80.4%
B9RJZ6	ATP-binding cassette transporter, putative (Ricinus communis)		78.8%
A0A068U8L9	Coffea canephora DH200=94 genomic scaffold, scaffold_14 (Coffea canephora)		79.4%
A0A059CSN0	Uncharacterized protein (Eucalyptus grandis)		79.3%
A5BAG5	Putative uncharacterized protein (Vitis vinifera)		79.4%
V4UHT1	Uncharacterized protein (Citrus clementina)		78.1%
B9IKS8	Uncharacterized protein (Populus trichocarpa)		77.9%
A0A0A0LNI2	Uncharacterized protein (Cucumis sativus)		78.1%
C8CA13	Pleiotropic drug resistance protein (Cucumis sativus)		78.0%
I1K8Y7	Uncharacterized protein (Glycine max)		77.0%
A0A0B2PCQ8	Pleiotropic drug resistance protein 1 (Glycine soja)		76.9%
A0A067GW95	Uncharacterized protein (Citrus sinensis)		77.7%
K7KIL7	Uncharacterized protein (Glycine max)		76.8%
M5Y1X8	Uncharacterized protein (Prunus persica)		78.8%

Fig. 6: Twenty top ranked hits of the BLAST query; 14 entries are uncharacterized proteins.

Amongst these 6 remaining proteins were 4 PDR1 proteins, which were chosen for the MSA (table 2); for our first investigations we wanted to have look only at the PDR1 proteins out of this BLAST result. Furthermore, we included PDR1 in Petunia axillaris although there is no topology information but 99% sequence identity, which improves the alignment (table 2).

The other 6 included proteins are listed in table 2 as well, with a short explanation why they were chosen.

Protein	Organism	Reasoning
PaPDR1	*Petunia axillaris*	BLAST result 99% sequence identity Improvement of the alignment; the TM regions are not known
NtPDR1	*Nicotiana tabacum*	BLAST result 73% sequence identity TMDs provided by Uniprot
NpPDR1	*Nicotiana plumbaginifolia*	BLAST result 69% sequence identity TMDs provided by Uniprot
OsPDR1	*Oryza sativa ssp. japonica*	BLAST result 68% sequence identity TMDs provided by Uniprot
AtPDR1	*Arabidopsis thaliana*	MOE alignment 59% sequence identity model organism for plant experiments TMDs provided by Uniprot
SpPDR1	*Schizosaccharomyces pombe*	MOE alignment 20% sequence identity representing PDR1 for fungi TMDs provided by Uniprot
ScPDR5	*Saccharomyces cerevisiae*	MOE alignment 25% sequence identity existing 3D-model of a PDR-transporter TMDs provided by Rutledge et al. [38]
CaCDR1	*Candida albicans*	MOE alignment 29% sequence identity existing 3D-model using ScPDR5 as a template TMDs provided by Rawal et al. [39]
hBCRP	*Homo sapiens*	MOE alignment TMD1 20%, TMD2 21% representing ABCG topology for humans TMDs provided by Wang et al. [40]
MmAbcb1a	*Mus musculus*	MOE alignment 18% sequence identity TMDs extracted from crystal structure (4M1M)

table 2: Proteins used for the MSA to solve the locations of the TMDs in PhPDR1.

The main aim of this approach was the analysis of the transmembrane parts of other proteins to be able to draw conclusions regarding PhPDR1. Multiple sequence alignments (MSAs) of PhPDR1 with 10 carefully selected proteins were performed separately for TMD1 and TMD2 and the helices were colored for visualization Fig. 7.

The alignment for TMD1 was created in MOE with the default options (Fig. 7). The result is very consistent, as most of the helices are aligned very well and in a reasonable manner. For TMD2, we used MOE (gap start: 23) and MAFFT (gap start: 3). The gap start parameters had to be increased to these high values because the alignments contained a lot of gaps in the beginning. To avoid the introduction of gaps, e.g. in the middle of a transmembrane helix, the gap start penalty can be raised. This approach improved the alignment in MOE for TMD2 (Fig. 7). We used 2 different alignments for TMD2 because the MOE alignment gave a satisfying result, but not for the very last TMH. The very same helix was aligned properly in MAFFT, although the rest of the alignment is not as good as the result of MOE. This observation was used in the decision process for PhPDR1s TMHs.
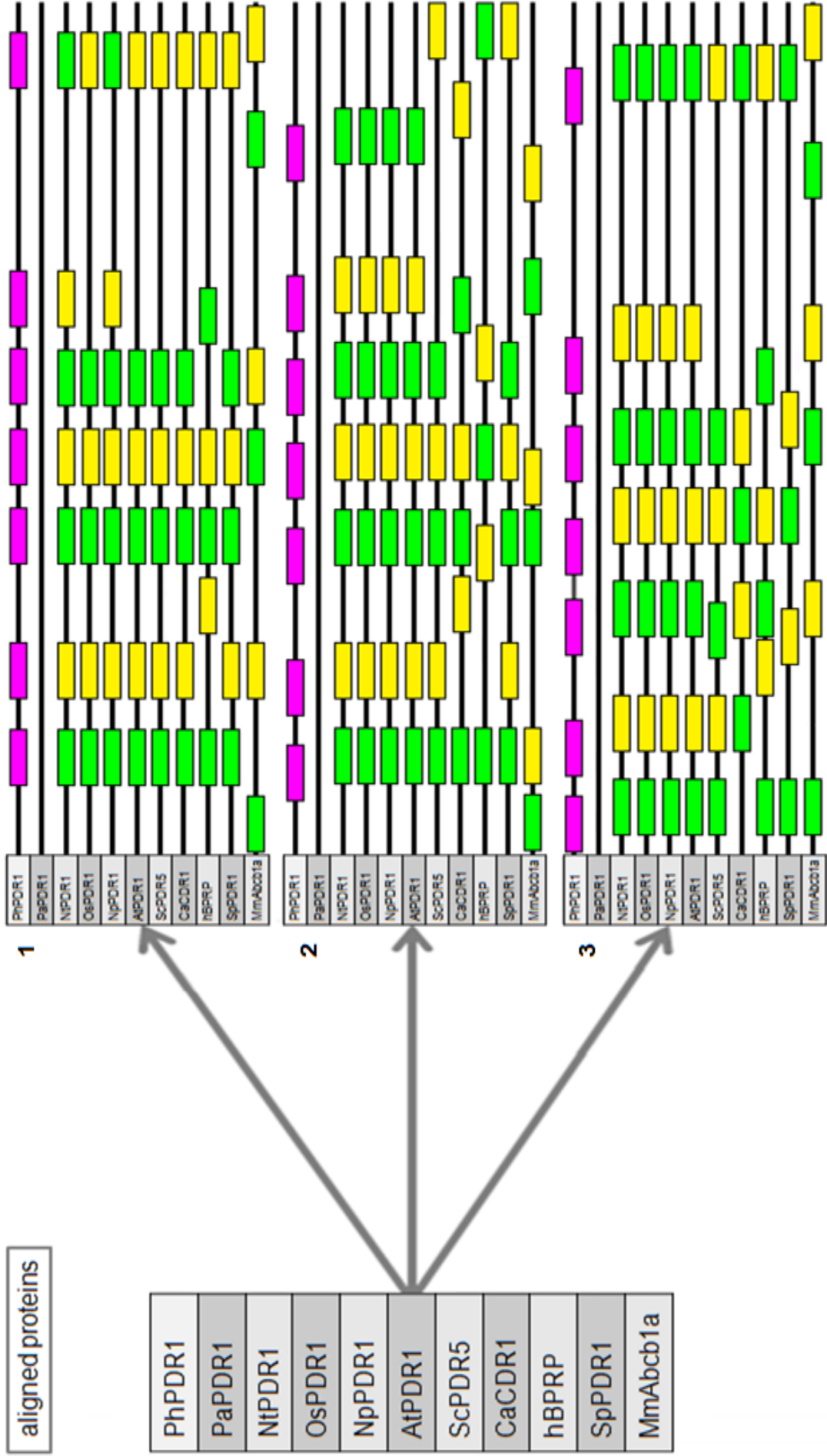
Fig. 7: Cartoon of the 3 alignments to solve the locations of PhPDR1s TMDs. 1. TMD1 in MOE, 2. TMD2 in MOE, 3. TMD2 in MAFFT. PhPDR1 is depicted in pink on the very top of every alignment; the TMDs are the result of the prediction tool TopPred, only to show a relation to the other proteins. The TMDs are colored green and yellow in turn to enable a clearer depiction of their locations. There is no information provided about PaPDR1s TMDs, it was included to improve the alignment. The original alignments with the marked helices in the same color code are included in the appendix (appendix 1, appendix 2, appendix 3).

34

On the one hand, we had collected the results of the prediction tools (Fig. 5) and the information extracted from the multiple sequence alignments on the other (Fig. 7).

According to these investigations, we were facing the problem that the prediction tools provided various (among them mostly odd) numbers of transmembrane areas and that Uniprot [14] also reports odd numbers of helices for some of the proteins we chose (Fig. 8). Uniprot annotates predicted transmembrane regions in these cases; they use TMHMM, Memsat, Phobius and the hydrophobic moment plot method of Eisenberg and co-workers [41]. Apart from our finding that the results of prediction tools are sometimes inconsistent, it is worth to mention that we found a 6 + 6 topology in literature for NtPDR1 [42] and NpPDR1 [43].

At the same time, the TMDs of the ABC transporters extracted from the literature and a crystal structure (MmAbcb1a, 4M1M), provided exclusively even numbers, actually the exact number of 6 helices per region (Fig. 8).

Prediction Tool	TMDs
TopPred	7 + 7
DAS	6 + 7
TMPred	6 + 6
SOSUI	6 + 7
PRED-TMR	6 + 7
SPLIT-SERVER	5 + 7
TMHMM	5 + 7
SACS MEMSAT	4 + 7
MEMSAT-SVM	6 + 7
PredictProtein	7 + 7

Protein	Organism	TMDs
NtPDR1	*Nicotiana tabacum*	7 + 7
NpPDR1	*Nicotiana plumbaginifolia*	7 +7
OsPDR1	*Oryza sativa ssp. japonica*	6 + 7
AtPDR1	*Arabidopsis thaliana*	6 + 7
SpPDR1	*Schizosaccharomyces pombe*	6 + 5

Protein	Organism	TMDs
ScPDR5	*Saccharomyces cerevisiae*	6 + 6
CaCDR1	*Candida albicans*	6 + 6
hBCRP	*Homo sapiens*	6
MmAbcb1a	*Mus musculus*	6 + 6

Fig. 8: The various numbers of TMDs. Left: results of the prediction tools. Upper right: The TMDs of the proteins used for the multiple sequence alignment provided by Uniprot. Lower right : The TMDs of the proteins used for the multiple sequence alignment provided by literature and the PDB [44] (for MmAbcb1a).

As already mentioned in the beginning of this chapter, an odd number of helices is not likely for ABC transporters in plants. According to the current knowledge, their N- and C-termini as well as their nucleotide binding domains (NBDs) are located in the cytosol [36].

Finally, aggregating all the gathered information the following conclusions may be drawn:

For TMD1, the situation was quite obvious. The prediction tools did not provide a consistent result for helix 6 (Fig. 5) and also the alignment showed the highest level of inconsistency in this area (alignment 1 in Fig. 7 and appendix 1). Therefore, we decided to leave out helix 6, which lead to the result of 6 transmembrane segments in a reasonable topology (also according to the longer loop between the last 2 helices).

For TMD2, we extracted the information of 2 alignments (the MAFFT [16] alignment only for the last helix). In the MOE [15] alignment and in the results of the prediction tools, the proposals for the transmembrane segments 8 to 12 were so far consistent. Both alignments showed the highest level of inconsistency in the area of helix 13. Even though all of the prediction tools suggest a transmembrane segment in this area, we decided to leave it out. The weight of the arguments supporting a proper overall topology, the agreement with literature information including experimental results (crystal structure 4M1M) and the result of our multiple sequence alignment counted more to take this decision.

The precise start and end of the transmembrane helices was determined manually based on the results of the prediction tools and the alignments as well.

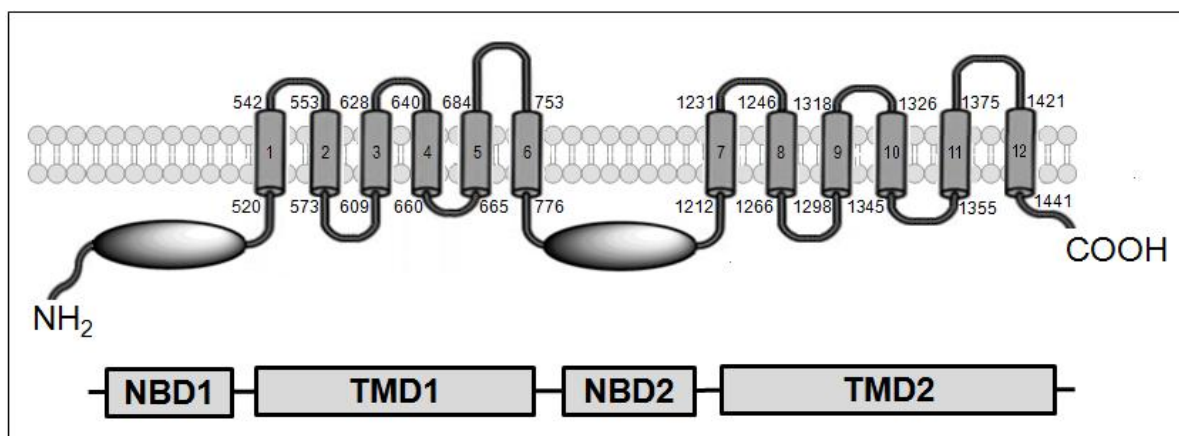The final result is depicted in Fig. 9.

Fig. 9: Proposal of the topology of PhPDR1. The numbers define the start and end of the respective transmembrane helices. For further information, see appendix 4.

4.2 Choice of the template

To create a homology model, a 3D structure of a protein is needed as a template.

It was a challenging task to find a proper one in this case. The PDB [44] provides no structures sharing the reverse topology of PDR proteins (Fig. 1, Fig. 9), but some ABC transporters with the regular topology of these proteins (TMD1-NBD1-TMD2-NBD2).

Rutledge et al. built a homology model of PDR5 in *Saccharomyces cerevisiae* (ScPDR5) [38] (see 1 Introduction – The ABC of phytohormone translocation, section "Elucidation of the structure of a strigolactone transporter"). They solved this problem by cutting the structural elements apart and reassembling them in the correct order. We decided to follow their approach and collected all the provided crystal structures of ABC transporters from the PDB. Furthermore, we included the sequence of ScPDR5 (Fig. 10) because we found another interesting approach by Rawal et al. [39]. They used the homology model of ScPDR5 as a template in their analysis of CDR1 in *Candida albicans*.

The alignment was done with the MAFFT [16] software (gap penalty = 3).

	1	2	3	4	5	6	7	8	9	10
C. elegans, ABCB1		34.0	18.4	17.6	14.4	14.1	13.1	9.2	7.8	7.8
Mus musculus, ABCB1a	30.7		17.5	16.3	12.8	15.0	14.6	8.8	5.6	9.2
Homo sapiens, ABCB10	20.2	21.3		24.9	20.6	20.0	17.2	10.1	9.3	11.1
T. maritima, TM_0288	17.1	17.5	22.0		24.7	22.2	17.8	10.1	5.9	12.9
St. aureus, SAV1866	14.3	14.1	18.6	25.2		19.7	16.9	10.5	9.7	14.3
S. typhimurium, msbA	14.0	16.5	18.1	22.7	19.7		15.0	6.9	9.3	9.7
N. aromaticivorans, atm1	12.7	15.8	15.3	17.9	16.6	14.7		8.2	8.7	10.6
S. cerevisiae, PDR5	8.7	9.3	8.8	9.9	10.0	6.6	8.0		15.3	7.4
P. hybrida PDR1	7.8	6.2	8.5	6.1	9.7	9.4	8.9	16.0		8.8
Escherichia coli, metN	5.3	6.9	6.8	8.9	9.7	6.6	7.3	5.2	5.9	

Fig. 10: Identity percentage matrix derived from the multiple sequence alignment of TMD2 of the possible templates with PhPDR1 [45]. PDR5 has the highest identity percentage.
For the PDB codes see appendix 5.

As can be seen in Fig. 10, ScPDR5 shows the highest sequence identity percentage and was thus chosen as a template for further comparative modelling steps.

The aim of this study is to provide first ideas about the 3D structure of PhPDR1. We are aware of the fact that homology modelling based on such little information on the one hand and the use of another model as a template on the other bears a high degree of hypotheses. Therefore, it is very important to understand that we want to provide an educated guess here, a kind of pioneer project in a vastly uncharted territory.

The final choice of the template was based on a multiple sequence alignment with the potential templates, including crystallized ABC transporters reported in the PDB and the high-quality, reliably validated homology model of ScPdr5. We assume that using ScPDR5 as a template will make more sense than choosing one of the crystallized structures. First, ScPDR5 belongs to the same family and shares therefore the same topology as PhPDR1. Second, the identity percentage of the sequences is distinctly higher than for all of the other proteins. Third, the homology

model of ScPDR5 is based on the crystal structures, so the influence of them will support our model throughout this approach.

As already mentioned, Rawal et al. [39] had also used this model as a template to create a 3D structure of CDR1 in *Candida albicans.* As their analysis leads to promising results and their procedure was validated in a suitable way, this approach was a solid support for our own plans.

4.3   Alignment of PhPDR1 and ScPDR5

Once the template was chosen, the next step was to align the two proteins pairwise. To improve the result, the four structural elements were aligned separately and merged again afterwards (table 3). This approach was chosen because the main focus lies on the very exact alignment of the transmembrane segments. Between them, there are loops that are not well conserved; extrapolated to the whole sequence they would influence the alignment in a way that would remove the emphasis from the transmembrane segments.

Protein	NBD1	TMD1	NBD2	TMD2
PhPDR1	1-499	500-780	781-1198	1199-1452
ScPDR5	1-499	500-795	796-1199	1200-1505

table 3: Amino acid indices of the four structural elements in which the proteins were distinguished for a more precise alignment.

Both NBD1 and NBD2 were aligned in MOE with the default options, the identity of NBD1 was 22,2% and NBD2 35%; therefore, we decided to use them as-is.

The TMDs took a bit more effort. They were aligned in MOE with default options as well (appendix 6 and appendix 7), but they had to be edited manually afterwards to achieve a satisfying result (appendix 8). Our main focuses in this matter were (numbered by priority):

1. to avoid gaps in the transmembrane helices. They need to be a continuous sequence as required for the helical geometry.

2. to align the transmembrane helices of the two proteins as connected in parallel as possible. The better they match; the better is the overall topology of the model.

3. to keep the maximum identity percentage. We tried best to take the identical residues which were aligned at the same position into account.

4. to avoid gaps in the rest of the sequence as well in balance with a reasonable sequence identity; this intention worked quite well as we reached an overall identity percentage of 24,5% for our alignment in the very end (see appendix 8).

4.4 Choice of the final model

As already mentioned, we had calculated 10 models with the modeller software [19], inclusive the validation scores molpdf, DOPE and GA341 (Fig. 11).

```
>> Summary of successfully produced models:
Filename                    molpdf      DOPE score     GA341 score
-----------------------------------------------------------------
Pdr1.B99990001.pdb       10302.38086   -128785.94531      0.15648
Pdr1.B99990002.pdb        9989.53223   -126510.18750      0.10358
Pdr1.B99990003.pdb        9570.34961   -127173.03906      0.15420
Pdr1.B99990004.pdb       10410.49414   -127263.45312      0.07147
Pdr1.B99990005.pdb        9824.38574   -127600.02344      0.11567
Pdr1.B99990006.pdb       10889.45801   -127973.70312      0.17514
Pdr1.B99990007.pdb       10396.91797   -127565.89062      0.11988
Pdr1.B99990008.pdb       10439.18555   -127414.70312      0.14373
Pdr1.B99990009.pdb        9684.64355   -128344.63281      0.12319
Pdr1.B99990010.pdb       10274.01270   -127663.56250      0.12932
```

Fig. 11: List of the calculated models with the related assessment scores molpdf, DOPE and GA341. The top three values and the corresponding models are marked with a grey box.

At first, we chose the four best ranked models according to the scores (model 1, 3, 6 and 9) for further validation (Fig. 11). Then we checked their Ramachandran plots and G-factors via PDBsum [23]. For the complete models, the plots are not very informative, as they are all very similar and moreover, the values are rather low (table 4). As we were focusing on the transmembrane domains, we cut the sequences of the models again into TMD1 and TMD2 and used their PDB-files as input. The results, which are significantly higher, are depicted in table 4.

Model, part of the sequence	Residues in most favoured regions
model 1, complete	83,3%
model 1, TMD1	89,9%
model 1, TMD2	88,4%
model 3, complete	83,5%
model 3, TMD1	91,2%
model 3, TMD2	88,9%
model 6, complete	83,5%
model 6, TMD1	91,6%
model 6, TMD2	86,0%
model 9, complete	83,8%
model 9, TMD1	90,3%
model 9, TMD2	87,4%

table 4: Percentage of the residues in the most favoured region according to the Ramachandran plots for each TMD1 and TMD2 of model 1, 3, 6 and 9.

According to these results (table 4), we decided to choose model 3 for our further investigations. As a good model should have over 90% of the residues in the most favoured region, we excluded model 1 first (both TMD1 and TMD2 under 90%). Then we excluded model 9 because the value for TMD1 is only 90,3% and the two models left have values over 91%. Finally, we excluded model 6 although it has the very best value for TMD1 of all the models but TMD2 has only 86,0% of the residues in the most favoured region, whereas the TMD2 of model 3 has 88,9%; and as this is closer to the intended 90%, we took model 3.
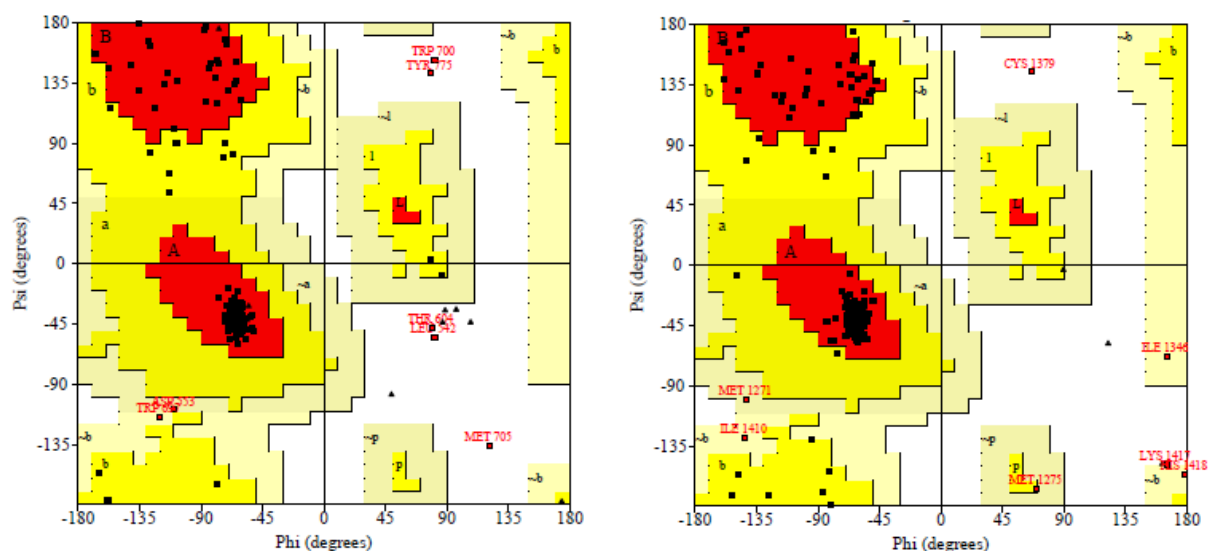
Fig. 12: The Ramachandran plots for model 3. Left: TMD1 (91,2% of the residues in the most favoured regions), right: TMD2 (88,9% of the residues in the most favoured regions).

4.5 Validation of the model

After the final selection of the model, the Protein Preparation Wizard [24] was applied. Before this procedure the model had a number of 91 bad contacts, 9 of these in TMD1 and 7 in TMD2. After this step, the number was reduced to 10 bad contacts in total; 8 of those are located in the NBDs.

However, in the course of this proceeding, the Ramachandran plots decreased from originally 83,5% (TMD1 91,2%. TMD2 88,9%) to 81,5% (TMD1 89,2%, TMD2 85,5%) of residues in the most favoured region, which is probably a matter of weighing the involved parameters. Furthermore, the percentage in the additional and generously allowed region increased, in the disallowed region decreased; this finding also relativises the result.

Due to the significant reduction of the bad contacts within the structure, we decided to deal with the loss and kept the minimized model for further validation.

4.5.1 Proline and glycine residues

In ScPDR5 there are 8, in PhPDR1 5 prolines in the transmembrane domains. In helix 3, there is 1 proline in the template and 1 in the model (ScPDR5: residue 609, PhPDR1: residue 610) which are well aligned and fit into the helical shape. The same

counts for helix 8 and 11. In helix 5, 8 and twice in helix 9, there is a proline only in ScPDR5 but not in PhPDR1; still the helical geometry is not influenced at all. In helix 10, there is only a proline in PhPDR1 but not in ScPDR5 but also here, the helix is well shaped without any irregular kink (Fig. 13).
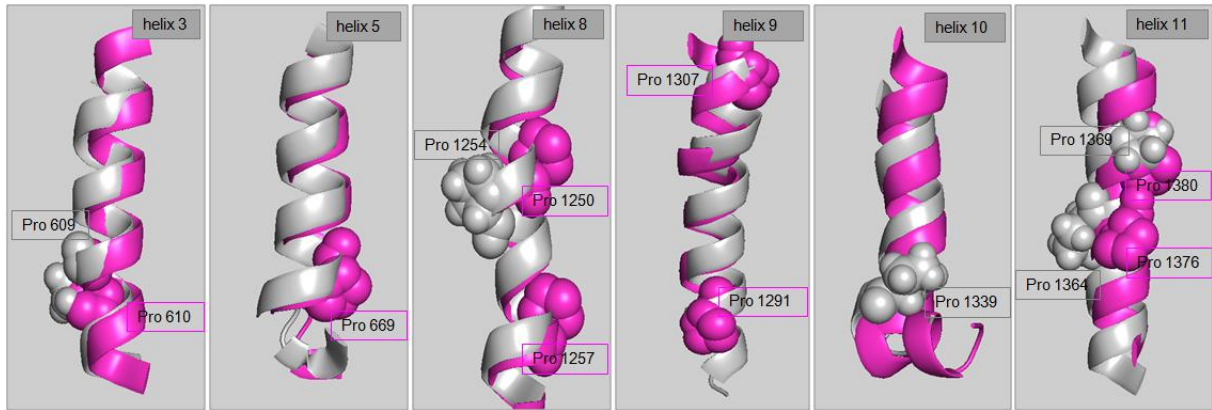

Fig. 13: Depiction of the prolines in the transmembrane domains. PhPDR1 = grey, ScPDR5 = pink

Of the glycines, there are 21 in PhPDR1, in ScPDR5 there are 18. As the approach was completely the same as for the prolines and neither here did occur any problems, we won't report the analysis of every single glycine residue but pick out a two examples (Fig. 14): in helix 1, there are 3 glycines in the template but not in the model; however, helix 1 is straight in both structures. In helix 7, we have a glycine in the model but not in the template and it is causing a kink here. Even though the template doesn't have a glycine in this position, the helix is bowed in the template as well, so the helix is shaped suitably in both cases.
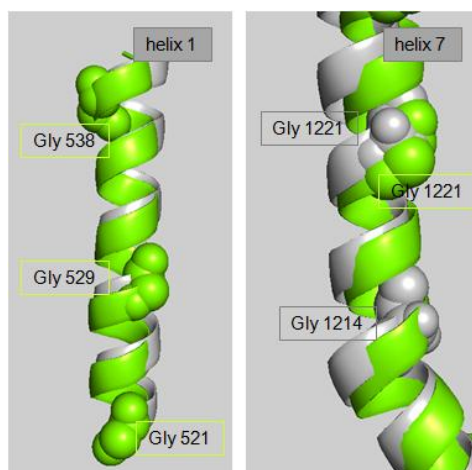
Fig. 14: Examples for the glycines in the transmembrane domains. PhPDR1 = grey, ScPDR5 = green

4.5.2 Charged and polar residues

The investigation of the charged and polar residues in the transmembrane domains led to the conclusion that lysine 638 and aspartate 1219 of the charged, as well as glutamine 1253 and tyrosine 1326 of the polar residues needed further improvement as they pointed directly towards the cell membrane. Therefore, different rotamers were calculated for the 4 mentioned amino acids and the most favourable of them was selected through a manual analysis. The first residue analyzed was lysine 638. From the 25 suggested rotamers, three (9, 23 and 24) were selected for a more detailed investigation. They were mapped in MOE [15] to see the precise position of each rotamer in context with the others (Fig. 15). This led to the conclusion, that rotamer 24 is closest to and most alongside the helix, therefore pointing much less towards the membrane. To make sure that there are no steric hindrances with the other residues in the near surroundings, a search for potential clashes was performed in MOE. As there were no problems found, we decided finally to keep rotamer 24 (Fig. 15).
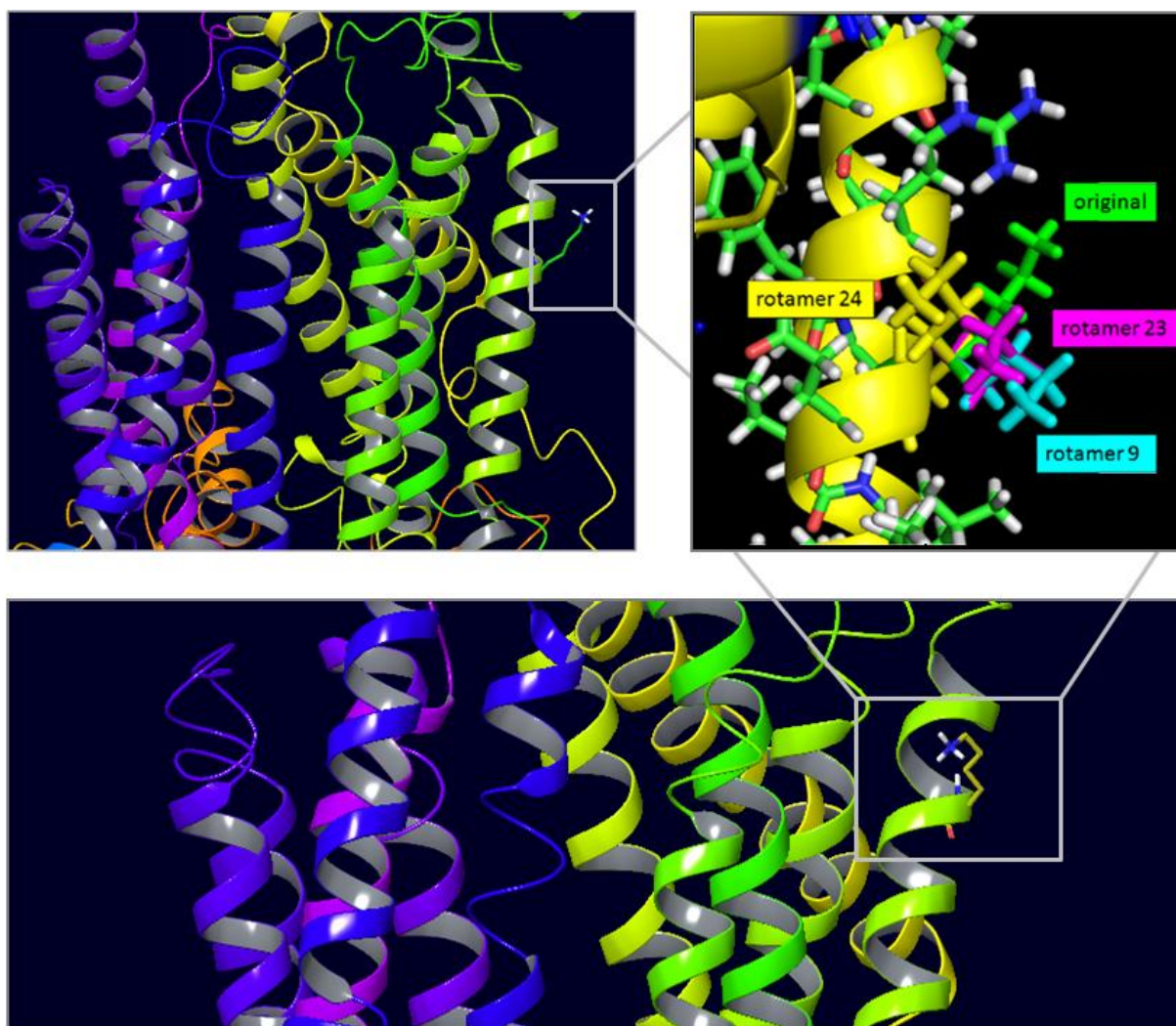
Fig. 15: The analysis of lysine 638. Top left: the original position of the residue pointing directly towards the membrane. Bottom: The final rotamer running alongside the helix. Top right: the original residue with the 3 final rotamer candidates (9 = cyan, 23 = pink and 24 = yellow) for the conclusive selection.

Exactly the same approach was used for aspartate 1219, glutamine 1253 and tyrosine 1326; the residues before and after the correction are depicted in Fig. 16.
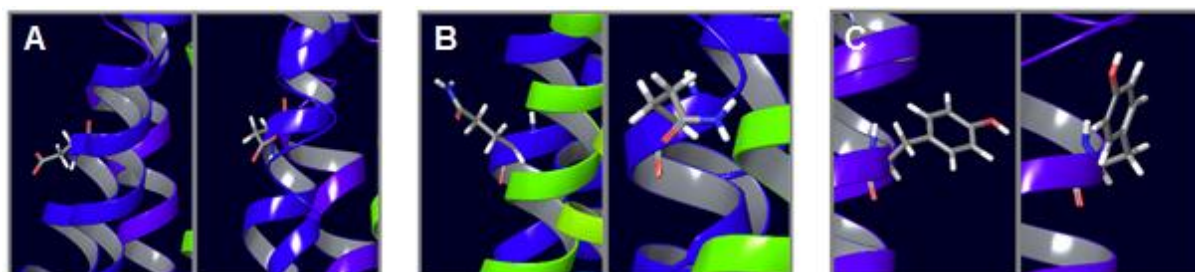

Fig. 16: The selected residues and their adjusted rotamers. In each A, B and C are on the left side the original residues and on the right side the final rotamers. A: aspartate 1219, B: glutamine 1253 and C: tyrosine 1326.

45

Finally, the Preparation Wizard was applied once again to accommodate the overall structure after the changes. The bad contacts were completely removed after this step.

After each adjusted rotamer, the residues in the most favoured regions of Ramachandran plots decreased again; ending up with a model having 78,3% of the residues in the most favoured regions.

We are convinced that the model with the corrected rotamers should be used for further investigation because:

- if the residues point towards the membrane, as they did in the beginning, the model would have obvious weak points, because this just won't be the case in nature. Therefore, the model is more realistic concerning this matter.
- when the bad contacts were decreased within the Preparation Wizard, which improved the protein structure, the residues in the most favoured region of the Ramachandran plots decreased as well. As mentioned before, this is probably due to differences in weighing involved parameters. Finally, the model has no bad contacts at all anymore after these proceedings.
- the residues in the disallowed regions stayed the same.
- most of the residues shifted from most favoured to additional region which sheds a different light on the results in this context.
- The template itself, which is a reliable, very valuably validated model, has "only" 84% of its residues in the most favoured region of the Ramachandran plot.

4.5.3 Electrostatic potentials

The electrostatic surface of PhPDR1 created in PyMOL [25] was compared with those of the template ScPDR5 and ABCB1 in *Arabidopsis thaliana* (AtABCB1). AtABCB1was chosen because it is an auxin transporter on the one hand and is part of an exhaustive analysis of electrostatic potentials by Bailly et al. [46], which was used as basis for these investigations on the other hand. The results are already explained in our paper (see 1 Introduction – The ABC of phytohormone translocation, section "Elucidation of a strigolactone transporter").

The main outcomes of this approach show that the translocation chambers of these 3 proteins display different properties, which could be due to their respective substrate specificities.
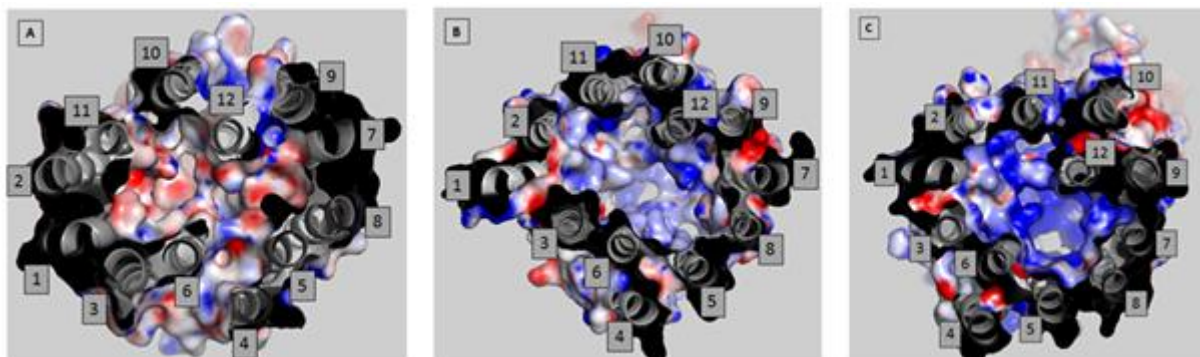
Fig. 17: The translocation chambers of AtABCB1 (A), PhPDR1 (B) and ScPDR5 (C) from the intracellular side [45]. The transmembrane helices are numbered 1-12; colour codes for the electrostatic potentials: red = negative, white = neutral, blue = positive.

4.6 Docking study

For the first time, a 3D model of the transporter PhPDR1 was generated. Now that there is a first idea about the shape and topology of the protein, we wanted to gain a very first glimpse on its transport mechanism. In general, ABC transporters eliminate substances from the cell - so they discharge compounds through the cell membrane.

Based on this knowledge the receptor grids for this docking study were placed in a way that covers the whole transmembrane region of the protein. According to the size of this area, 2 different boxes were needed (Fig. 18).
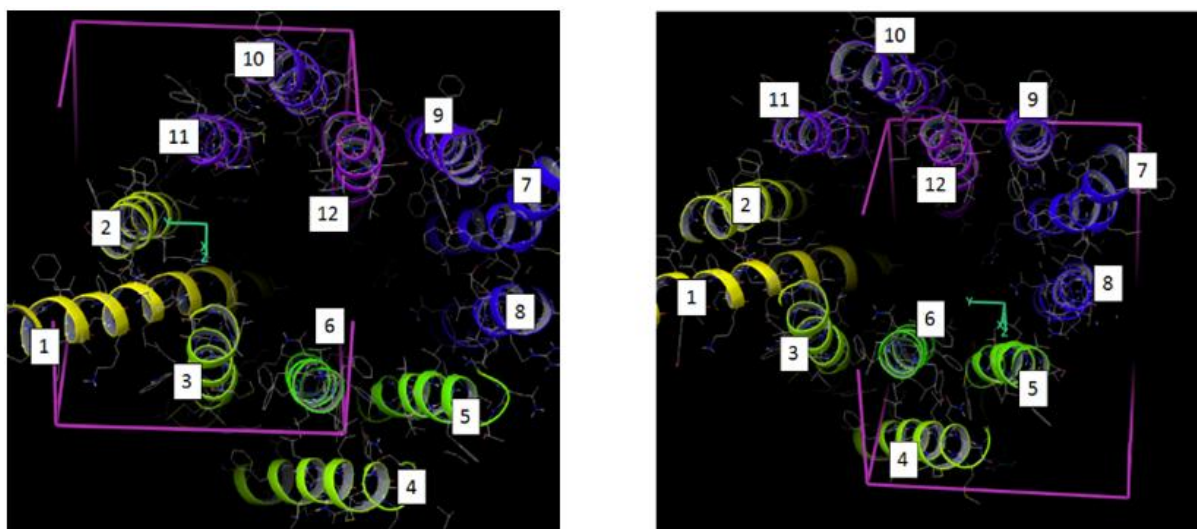
Fig. 18: The two receptor grids generated in Maestro [26] to cover the whole transmembrane region for the docking study with orobanchol.

To define these grids precisely, the relevant residues were extracted from the 3D structures in MOE [15] (see table 5). For the first one the helices 1, 2, 3, 10, 11 and 12 were used as an input; for the second one 4, 5, 6, 7, 8 and 9.

TMH	Residues
1	513-542
2	553-581
3	604-625
4	632-654
5	668-687
6	750-773

TMH	Residues
7	1203-1225
8	1241-1263
9	1295-1312
10	1323-1338
11	1346-1371
12	1419-1439

table 5: The residues extracted from the 3D structure to define the receptor grids.

Finally, orobanchol (Fig. 19) was docked into the 3D structure of PhPDR1 prepared with the respective grids. For each run, the ligand poses were set to 100; 98 and 94 poses were reported respectively for the first and second box.
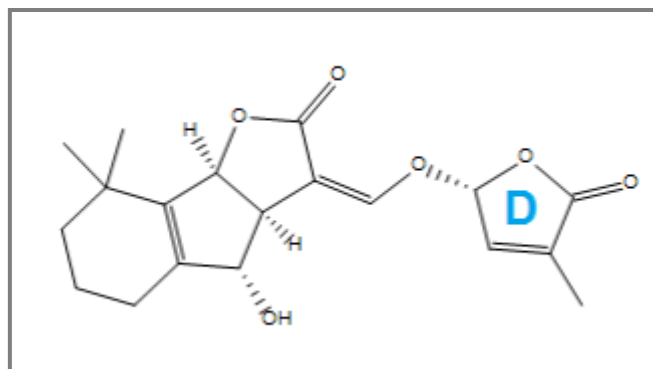
Fig. 19: The 2D structure of orobanchol. The D-ring seems to play an important role for the interaction between ligand and protein.

To analyze the results, the two separate docking outputs were merged and the poses were clustered to draw further conclusions. The clustering approach lead to 60 clusters containing various poses. We decided to consider each cluster containing at least seven poses because the more poses are clustered in a certain area, the higher are the chances that this could be a binding site. However, the exact number of seven poses was arbitrarily chosen.

Therefore, we extracted six different clusters (2, 4, 39, 41, 45, 55) from the original result (Fig. 20).
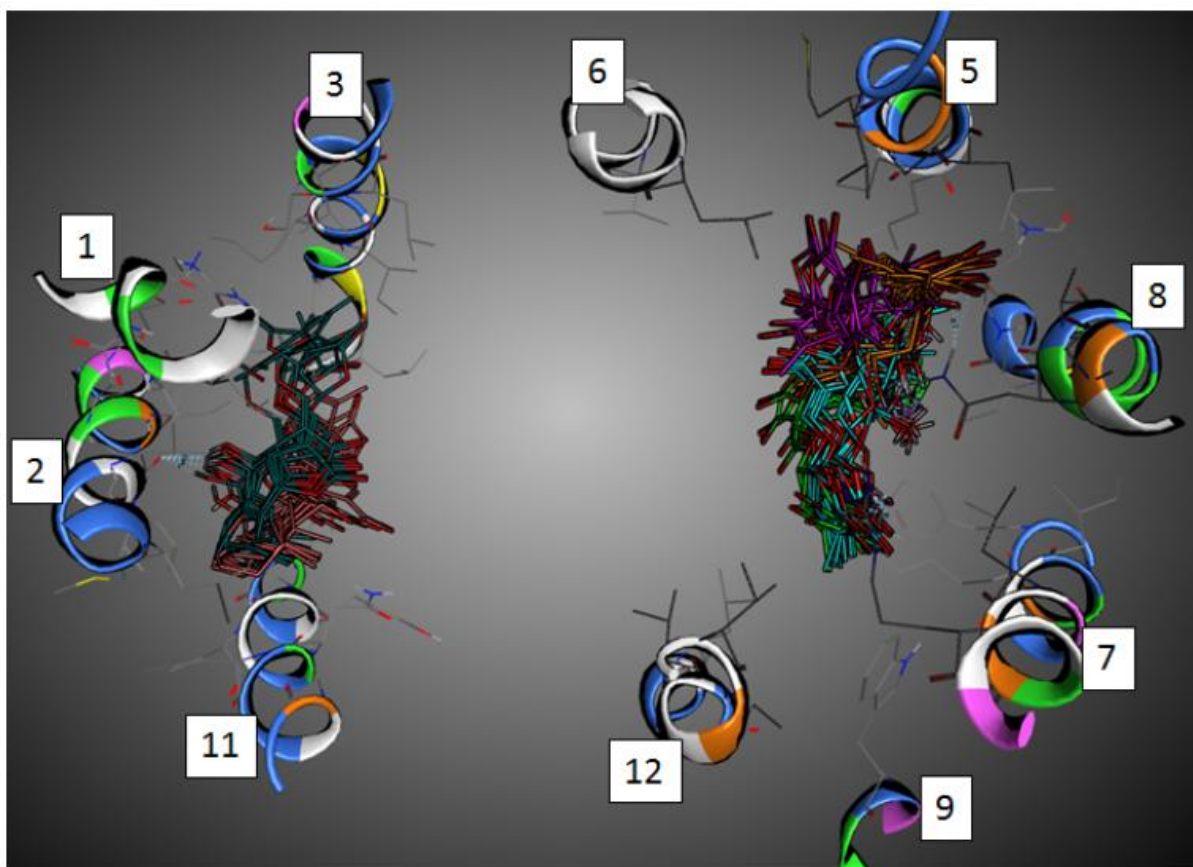
Fig. 20: The six considered clusters of the obtained orobanchol docking poses. The helices are numbered from 1-12. Colour coding of the depicted clusters: petrol = cluster 2, salmon = cluster 4 (both on the left side); orange = cluster 39, cyan = cluster 41, pink = cluster 45, green = cluster 55 (all on the right side).

Short description of the selected clusters and the reported interactions of orobanchol (Fig. 19) with PhPDR1:

At first, we analyzed the interactions between the poses and the protein (table 6). There are 4 amino acids involved in H-bonds to the ligand. These interactions take place with the hydroxyl group in pos. 4 and the carboxyl-O of the D-ring of the orobanchol molecule (Fig. 19).

Cluster	No. of poses	Ile568	Trp605	Asn1248	Arg1224	Avg. docking score
2	7	<u>6</u>	2	/	/	-5,31161
4	7	<u>1</u>	1	/	/	-4,62302
39	14	/	/	$\frac{1}{2}$	/	-4,0895
41	11	/	/	1	/	-3,74562
45	12	/	/	1	/	-3,40481
55	8	/	/	/	4	-2,98437

table 6: Analysis of the 6 chosen docking clusters. Orobanchol is connected to Ile 568, Trp 605, Arg 1224 and Asn 1248 by H-bond interactions. In the columns, the number of interacting poses with the respective amino acid is reported. Underlined numbers represent an H-bond with the hydroxyl group in pos. 4 of the orobanchol molecule; the others with the carboxyl-O of the D-ring.

Furthermore, the average docking scores of the clusters were analyzed (table 6). Cluster 2 and cluster 4 have remarkable average scores and apart from that they are both on the same side of the protein (between helix 1, 2, 3, 10, 11 and 12; see Fig. 20); the four other clusters have lower scores and are all located on the other side. This result could be an interesting finding for further investigations - maybe this is the first indication to a preferential binding site.

Finally, the 6 clusters were analyzed according to their environment within the protein (Fig. 21). This was done by selecting the top scored pose of each cluster and calculating the molecular surface colored by lipophilicity for them in MOE [15]. On the one hand, this approach enables to draw conclusions on the steric fitting of the ligand in its certain surroundings; and allows an impression how protein and ligand match with respect to hydrophilic, neutral and lipophilic properties on the other.
As the figure shows (Fig. 21), the poses fit perfectly into their environment without any steric clashes. The colored surfaces emphasise the facility of the hypothetical poses due to the hydrophilic, neutral and lipophilic properties.
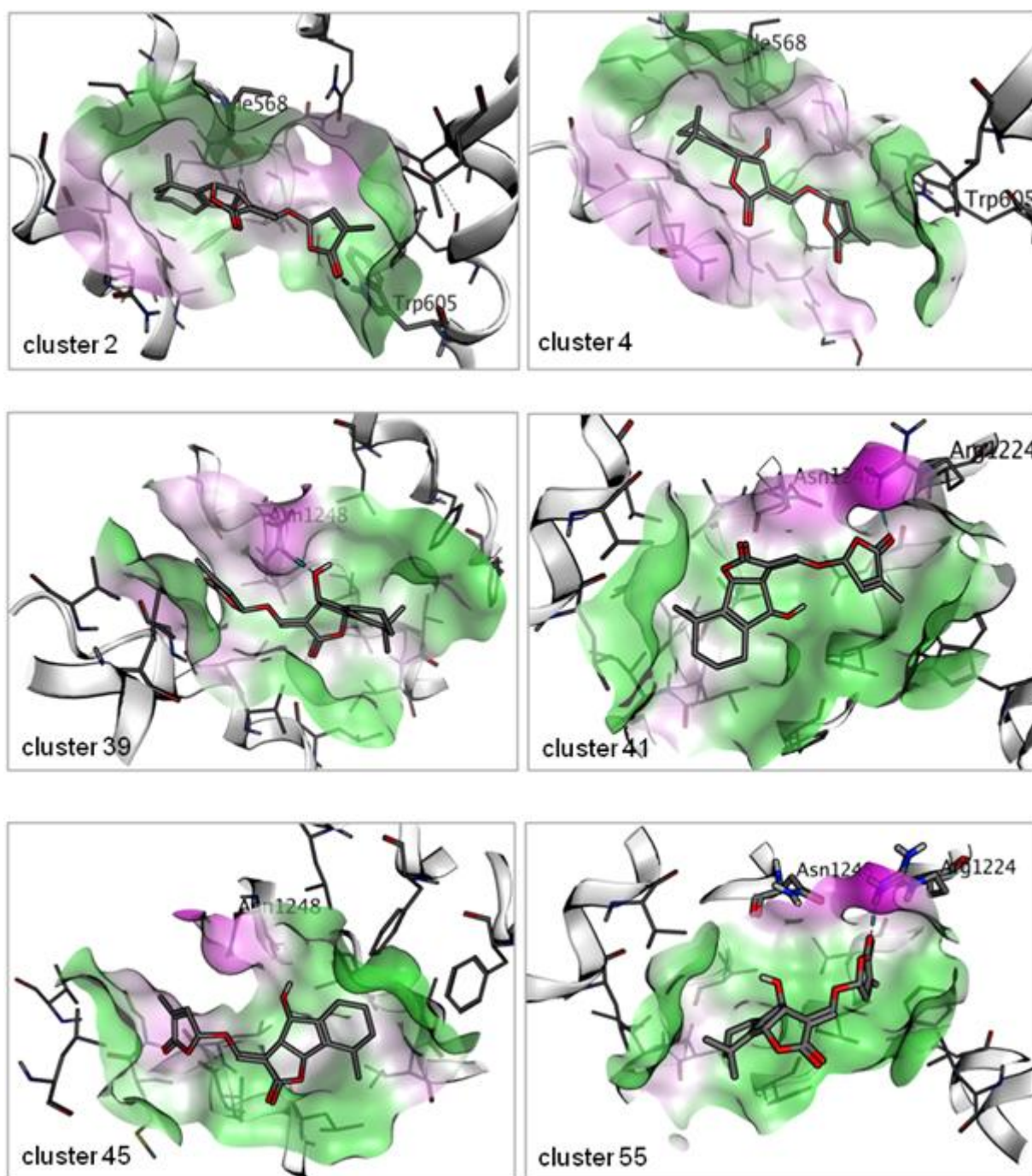
Fig. 21: the top orobanchol pose of each cluster depicted with the molecular surface of the closest surroundings within PhPDR1. Color codes for the molecular surfaces: pink = hydrophilic, white = neutral, green = lipophilic. The H-bonds (cluster 2, 39, 41 and 55) are depicted in cyan.

It is important to mention, that this docking study is a very first attempt to collect primary ideas of PhPDR1s binding mechanisms. Our approach lead to two main areas (Fig. 20) where the considered clusters accumulate. Furthermore, the amino

acids interacting with the docked orobanchol poses found in this areas could easily make sense when imagining a compound being discharged from inside to outside the cell. Our first investigations show a promising strategy as there were no inconsistencies detected through the first validation steps. However, further investigations will be needed to refine the binding hypothesis.

5 Conclusions and Outlook

As mentioned in the very beginning, the information about plant ABC transporters is scant. According to the breaking news from Kretzschmar et al. [30] about PhPDR1's strigolactone transport we picked up the topic with great interest and started a thorough investigation.

We did comprehensive literature search and put a lot of effort on defining the transmembrane domains to create the most optimal homology model as possible with the available resources. The sequence alignments were carefully considered through extensive experiments including various approaches and software packages. Finally, the performance of an intense validation process with many different tools and methods pointing out a lot of aspects and background information behind our operations as well, round off this diligently conducted study.

All in all, a lot of aspects and capabilities were explored and investigated to gain a first idea of the 3D structure of a PDR plant hormone transporter including an extensive validation process of our approach.

Of course, the whole story is rather an educated guess - but where would the humanity be without creating and expand exciting ideas? We built our hypotheses on reliable literature and carefully conducted and validated experiments, so the chances are quite high that these theoretical assumptions will be proven real facts one day.

The model and the first ideas about the docking poses could serve as a solid basis for further investigations in this direction; the next interesting task could be to conduct a docking study with the other strigolactone mimics.

6   References

[1] *Claros MG, von Heijne G*. TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci CABIOS 1994; 10: 685–686.

[2] *Nakashima H, Nishikawa K*. The amino acid composition is different between the cytoplasmic and extracellular sides in membrane proteins. FEBS Lett 1992; 303: 141–146.

[3] *Cserzö M, Wallin E, Simon I, von Heijne G, Elofsson A*. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 1997; 10: 673–676.

[4] *Hofman, K, Stoffel, W*. TMbase - a database of membrane spanning proteins segments. Biol Chem Hoppe-Seyler 1993; 347: 166.

[5] *Hirokawa T, Boon-Chieng S, Mitaku S*. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinforma Oxf Engl 1998; 14: 378–379.

[6] *Kyte J, Doolittle RF*. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982; 157: 105–132.

[7] *Pasquier C, Promponas VJ, Palaios GA, Hamodrakas JS, Hamodrakas SJ*. A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. Protein Eng 1999; 12: 381–385.

[8] *Juretić D, Zoranić L, Zucić D*. Basic charge clusters and predictions of membrane protein topology. J Chem Inf Comput Sci 2002; 42: 620–632.

[9] *Krogh A, Larsson B, von Heijne G, Sonnhammer EL*. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001; 305: 567–580.

[10] *Jones DT, Taylor WR, Thornton JM*. A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry (Mosc) 1994; 33: 3038–3049.

[11] *Nugent T, Jones DT*. Transmembrane protein topology prediction using support vector machines. BMC Bioinformatics 2009; 10: 159.

[12] *Rost B, Yachdav G, Liu J*. The PredictProtein server. Nucleic Acids Res 2004; 32: W321–W326.

[13] *Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ*. Basic local alignment search tool. J Mol Biol 1990; 215: 403–410.

[14] The UniProt Consortium, Activities at the Universal Protein Resource (UniProt), Nucleic Acids Res. 42: D191-D198 (2014). Available at http://uniprot.org

[15] Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.

[16] *Katoh K, Misawa K, Kuma K, Miyata T*. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002; 30: 3059–3066.

[17] *Katoh K, Standley DM*. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 2013; 30: 772–780.

[18] *Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ*. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinforma Oxf Engl 2009; 25: 1189–1191.

[19] *Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A*. Comparative protein structure modeling using Modeller. Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al 2006; Chapter 5: Unit 5.6.

[20] *Melo F, Sánchez R, Sali A.* Statistical potentials for fold assessment. Protein Sci Publ Protein Soc 2002; 11: 430–448.

[21] *Shen M-Y, Sali A.* Statistical potential for assessment and prediction of protein structures. Protein Sci Publ Protein Soc 2006; 15: 2507–2524.

[22] *Ramachandran GN, Ramakrishnan C, Sasisekharan V.* Stereochemistry of polypeptide chain configurations. J Mol Biol 1963; 7: 95–99.

[23] *Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM.* PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends Biochem Sci 1997; 22: 488–490.

[24]Schrödinger Release 2014-3: Schrödinger Suite 2014-3 Protein Preparation Wizard; Epik version 2.9, Schrödinger, LLC, New York, NY, 2014; Impact version 6.4, Schrödinger, LLC, New York, NY, 2014; Prime version 3.7, Schrödinger, LLC, New York, NY, 2014.

[25]The PyMOL Molecular Graphics System, Version 1.6. Schrödinger, LLC. Available at http://www.pymol.org

[26]Schrödinger Release 2014-3: Maestro, version 9.9, Schrödinger, LLC, New York, NY, 2014.

[27] *Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA.* Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 2001; 98: 10037–10041.

[28] *Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA.* PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res 2004; 32: W665–W667.

[29] *Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA.* PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res 2007; 35: W522–W525.

[30] *Kretzschmar T, Kohlen W, Sasse J, Borghi L, Schlegel M, Bachelier JB, Reinhardt D, Bours R, Bouwmeester HJ, Martinoia E.* A petunia ABC protein controls strigolactone-dependent symbiotic signalling and branching. Nature 2012; 483: 341–344.

[31] *Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004; 47: 1739–1749.

[32] Small-Molecule Drug Discovery Suite 2014-3: Glide, version 6.4, Schrödinger, LLC, New York, NY, 2014.

[33] Schrödinger Release 2014-3: LigPrep, version 3.1, Schrödinger, LLC, New York, NY, 2014.

[34] (3E,3aS,4S,8bS)-4-Hydroxy-8,8-dimethyl-3-({[(2R)-4-methyl-5-oxo-2,5-dihydro-2-furanyl]oxy}methylene)-3,3a,4,5,6,7,8,8b-octahydro-2H-indeno[1,2-b]furan-2-one| $C19H22O6$ | ChemSpider 2014.

[35] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[36] *Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, Klein M, Kolukisaoglu Ü, Lee Y, Martinoia E, Murphy A, Rea PA, Samuels L, Schulz B, Spalding EP, Yazaki K, Theodoulou FL.* Plant ABC proteins – a unified nomenclature and updated inventory. Trends Plant Sci 2008; 13: 151–159.

[37] *O'Donovan C, Martin MJ, Gattiker A, Gasteiger E, Bairoch A, Apweiler R*. High-quality protein knowledge resource: SWISS-PROT and TrEMBL. Brief Bioinform 2002; 3: 275–284.

[38] *Rutledge RM, Esser L, Ma J, Xia D*. Toward understanding the mechanism of action of the yeast multidrug resistance transporter Pdr5p: A molecular modeling study. J Struct Biol 2011; 173: 333–344.

[39] *Rawal MK, Khan MF, Kapoor K, Goyal N, Sen S, Saxena AK, Lynn AM, Tyndall JDA, Monk BC, Cannon RD, Komath SS, Prasad R*. Insight into Pleiotropic Drug Resistance ATP-binding Cassette Pump Drug Transport through Mutagenesis of Cdr1p Transmembrane Domains. J Biol Chem 2013; 288: 24480–24493.

[40] *Wang H, Lee E-W, Cai X, Ni Z, Zhou L, Mao Q*. Membrane topology of the human breast cancer resistance protein (BCRP/ABCG2) determined by epitope insertion and immunofluorescence. Biochemistry (Mosc) 2008; 47: 13778–13787.

[41] http://www.uniprot.org/help/transmem; date of access: 2015-04-06.

[42] *Sasabe M, Toyoda K, Shiraishi T, Inagaki Y, Ichinose Y*. cDNA cloning and characterization of tobacco ABC transporter: NtPDR1 is a novel elicitor-responsive gene1. FEBS Lett 2002; 518: 164–168.

[43] *Jasiński M, Stukkens Y, Degand H, Purnelle B, Marchand-Brynaert J, Boutry M*. A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion. Plant Cell 2001; 13: 1095–1107.

[44] *Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE*. The Protein Data Bank. Nucleic Acids Res 2000; 28: 235–242.

[45] *Hellsberg E, Montanari F, Ecker GF*. The ABC of Phytohormone Translocation. Planta Med 2015.

[46] *Bailly A, Yang H, Martinoia E, Geisler M, Murphy AS*. Plant lessons: exploring ABCB functionality through structural modeling. Plant Traffic Transp 2012; 2: 108.

7 Appendix

7.1 Supplemental material

appendix 1: alignment of TMD1 in MOE with default options. PhPDR1 is depicted in pink on the very top of the alignment; the TMDs are the result of the prediction tool TopPred, only to show a relation to the other proteins. The TMDs are colored green and yellow in turn to enable a clearer depiction of their locations. There is no information provided about PaPDR1s TMDs, it was included to improve the alignment.

```
Block 1
 1: Ph_Pdr1|1200-1450        QHWSYWRNPPYTAVRIMFTFFIALMFGTIFWDLGSRRERQQDLLNAIGSMYIAVLF
 2: Petaxi_Pdr1|1200-1450                  TAVRIMFTFFIALMFGTIFWDLGSRRERQQDLLNAIGSMYIAVLF
 3: Nictab_Pdr1_TMD2|1180…             AYTAVRFIFTTFIALIFGTMFWDLGTKVSKSQDLLNAMGSMYAAVLF
 4: Orysat_TMD2|1210-1460        WRNPPYNAIRLFFTTVIALIFGTIFWDLGGKMGQSQDLFNAMGSMYAAVLF
 5: Nicplu_Pdr1_TMD2|1180…       PAYTAVRLIFTTFIALIFGTMFWDIGTKVSRNQDLVNAMGSMYAAVLF
 6: Aratha_Pdr1_TMD2|1160…            LARFFFTLAAAVMLGSIFWKVGTKRENANDLTKVIGAMYAAVLF
 7: sp|PDR5_TMD2|1200-15…        WRSPDYLWSKFILTIFNQLFIGFTFFKAGTSL···QGLQNQMLAVFMFTVI
 8: Canalb_Cdr1|1196-1487          YIYSKIFLVVSAALFNGFSFFKAKNNM···QGLQNQMFSVFMFFIP
 9: Abcg2|385-665            FKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS···TGIQNRAGVLFFLTTN
10: Schizpom_Pdr1_TMD2|1…          VAFNIVAGLIIGFSFYKQGVGV···EETQNKMFSAYMLTVA
11: Pgp_mouse_TMD2|705-…     WPYFVVGIFCAIINGGLQPAFSVIFSKVVGVFTNGGPPETQRQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTKRLRYMVFKSML

Block 2
 1: Ph_Pdr1|1200-1450        LGVQNATTVQPVIAIERTVFYRERAAGMYSAMPYAFGQVMIELPYLFLQTIIYGVIVYAMIGFEWTVAKF·······FWYLFFMYFTLLY
 2: Petaxi_Pdr1|1200-1450    LGVQNATTVQPVIAIERTVFYRERAAGMYSAMPYAFGQVMIELPYLFLQTIIYGVIVYAMIGFEWTVAKF·······FWYLFFMYFTLLY
 3: Nictab_Pdr1_TMD2|1180…   LGVQNASSVQPVVAIERTVFYRERAAGMYSAIPYAFGQVSIEIPYIFVQSVFYGIIVYAMIGFEWDVGKF·······FWYLFIMFFTLLY
 4: Orysat_TMD2|1210-1460    IGVLNGQSVQPVVSVERTVFYRERAAGMYSALPYAFGQVAIEFPYTLVQSVIYSIIVYSMIGFQWTVAKF·······FWYLFFMFFTLLY
 5: Nicplu_Pdr1_TMD2|1180…   LGVQNSSSVQPVVSVERTVFYREKAAGMYSAIPYAFAQVLIEIPYIFVQATVYGLIVYSMIGFEWTVAKF·······FWDFFFMFFTFLY
 6: Aratha_Pdr1_TMD2|1160…   VGVNNSSSVQPLIAVERSVFYRERAAEMYSALPYALAQVVCEIPYVLIQTTYYTLIIYAMMCFEWTLAKF·······FWFYFVSFMSFLY
 7: sp|PDR5_TMD2|1200-15…    FNPILQQYLPSFVQQRDLYEARERPSRTFSWISFIFAQIFVEVPWNILAGTIAYFIYYYPIGFYSNASAAGQLHERGALFWLFSCAFYVY
 8: Canalb_Cdr1|1196-1487    FNTLVQQMLPYFVKQRDVYEVREAPSRTFSWFAFIAGGQITSEIPYQVAVGTIAFFCWYYPLGLYNNATPTDSVNPRGVLMWMLVTAFYVY
 9: Abcg2|385-665            QCFSSVSAVELFVVEKKLFIHEYISGYYRVSSYFLGKLLSDLLPMRMLPSIIFTCIVYFMLGLKPKADAF·······FVMMFTLMMVAYS
10: Schizpom_Pdr1_TMD2|1…    STSTMNGLQPKFIYFRSIYEQYEQNTAIYSRTAFIIAFFLVEAVINCCFATLFFFGWYYPSGFYEF···NHNIPFYGGFAWLMLMIFTLY
11: Pgp_mouse_TMD2|705-…     RQDVSWFDDPKNTTGALTTRLANDAAQVKGATGSRLAVIFQNIANLGTGIIISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDK

Block 3
 1: Ph_Pdr1|1200-1450        FTLYGMMTVAVTPNQSIAAIISSAFYAVWNLFCGFIVPKTRMPVWWRWYYYICPISWTLYGLIASQFGDIQDRL······DTNETVEQFI
 2: Petaxi_Pdr1|1200-1450    FTLYGMMTVAVTPNHSIAAIISSAFYAVWNLFCGFIVPKTRMPVWWRWYYYICPISWTLYGLIASQFGDIQDRL······DTNETVEQFI
 3: Nictab_Pdr1_TMD2|1180…   FTFYGMMGVAVTPNQNVASIVAAFFYGVWNLFSGFIIPRPRMPVWWRWYYWANPVAWTLYGLVASQFGDIQTKLS·····D·NETVEQFL
 4: Orysat_TMD2|1210-1460    FTFYGMMAVGLTPSYHVASIVSSAFYAIWNLFTGFVISRPATPVWWRWYCWICPVAWTLYGLIVSQYGDIVTPM······DDGIPVNVFV
 5: Nicplu_Pdr1_TMD2|1180…   FTFFGMMTVPNQNVASIVAGFFYTVWNLFSGFIVPRPRIPIWWRWYYWGCPIAWTLYGLVASQFGDLQDPLT·····DQNQTVEQFL
 6: Aratha_Pdr1_TMD2|1160…   FTYYGMMTVALTPNQQVAAVFAGAFYGLFNLFSGFVIPRPRIPKWIWYYWICPVAWTVYGLIVSQYGDVEDTIKVPGMA·NDPTIKWYI
 7: sp|PDR5_TMD2|1200-15…    VGSMGLLVISFNQVAESAANLASLLFTMSLSFCGVMTTPSAMPRFWIFMYRVSPLTYFIQALLAVGVANVDVKCADYELLEFTPPSGMTC
 8: Canalb_Cdr1|1196-1487    TATMGQLCMSFSELADNAANLATLLFTMCLNFCGVLAGPDVLPGFWIFMYRCNPFTYLVQAMLSTGLANTFVKCAEREYVSVKPPNGESC
 9: Abcg2|385-665            ASSMALAIAAGQSVVSVATLLMTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNF·CPGLNATGNNPCNYATC
10: Schizpom_Pdr1_TMD2|1…    YTTLGIGIATISPSIGTASIISGTAFVFIQYFNGMIQLPGVIVGFWKWMDALSPYKYFLEGMIGGVLHDAPITCEKFEIHYVDPPPNYSC
11: Pgp_mouse_TMD2|705-…     KELEGSGKIATEAIENFRTVVSLTREQKFETMYAQSLQIPYRNAMKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVTQQLMTFENVLLV

Block 4
 1: Ph_Pdr1|1200-1450        ENFFDFKHDFVGYVALILVGISVLFLFIFAFSIKTFNFQ
 2: Petaxi_Pdr1|1200-1450    ENFFDFKHDFVGYVALILVGISVLFLFIFAFSIKTFNFQ
 3: Nictab_Pdr1_TMD2|1180…   RRYFGFKHDFLGVVAAVLTAYVFMFAFTFAFAIKAFN
 4: Orysat_TMD2|1210-1460    ENYFDFKHSWLGFVAVVIVAFTMLFAFLFGFAI
 5: Nicplu_Pdr1_TMD2|1180…   RSNFGFKHDFLGVVAAVIVAFAVVFAFTFALGIKA
 6: Aratha_Pdr1_TMD2|1160…   ENHYGYDADFMIPIATVLVGFTLFFAFMFAFGIRTLNFQQ································R
 7: sp|PDR5_TMD2|1200-15…    G··QYMEPYLQLAKTGYLTDENATDTCSFCQISTTNDYLANVNSFYSERWRNYGIFICYIAFNYIAGVFFYWLARVPKKN
 8: Canalb_Cdr1|1196-1487    S··TYLDPYIKFA··GGYFETRNDGSCAFCQMSSTNTFLKSVNSLYSERWRNFGIFIAFIAINIILTVIFYWL
 9: Abcg2|385-665            TGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS
10: Schizpom_Pdr1_TMD2|1…    G··EYFSSFLNSSGHGIVYNPEAYSSCQYCPYKNADELMVGFGYHYNHKWRNFCIMIGYTAFNLGAAIALYYI
11: Pgp_mouse_TMD2|705-…     FSAIVFGAMAVGQVSSFAPDYA
```

appendix 2: alignment of TMD2 in MOE, gap start penalty =23. PhPDR1 is depicted in pink on the very top of the alignment; the TMDs are the result of the prediction tool TopPred, only to show a relation to the other proteins. The TMDs are colored green and yellow in turn to enable a clearer depiction of their locations. There is no information provided about PaPDR1s TMDs, it was included to improve the alignment.

```
1: Ph_Pdr1|1200-1450      QHWSYWRNPPYTAVRIMFTFFIALMFGTIFWDLGSRRERQQDLLNAIGSMYIAVLFLGVQNA····TTVQPVIAIERTVF·YRERAAG
2: Petaxi_Pdr1|1200-1450          TAVRIMFTFFIALMFGTIFWDLGSRRERQQDLLNAIGSMYIAVLFLGVQNA····TTVQPVIAIERTVF·YRERAAG
3: Nictab_Pdr1_TMD2|1180...    AYTAVRIFTTFIALIFGTMFWDLGTKVSKSQDLLNAMGSMYAAVLFLGVQNA····SSVQPVVAIERTVF·YRERAAG
4: Orysat_TMD2|1210-1460   WRNPPYNAIRLFFTTVIALIFGTIFWDLGGKMGQSQDLFNAMGSMYAAVLFIGVLNG····QSVQPVVSVERTVF·YRERAAG
5: Nicplu_Pdr1_TMD2|1180...    PAYTAVRLIFTTFIALIFGTMFWDIGTKVSRNQDLVNAMGSMYAAVLFLGVQNS····SSVQPVVSVERTVF·YREKAAG
6: Aratha_Pdr1_TMD2|1160...         LARFFFTLAAAVMLGSIFWKVGTKRENANDLTKVIGAMYAAVLFVGVNNS····SSVQPLIAVERSVF·YRERAAE
7: sp|PDR5_TMD2|1200-15...    WRSPDYLWSKFILTIFNQLFIGFTFFKAGTSLQGLQN······QMLAVFMFTVIFNPI·LQQYLPSFVQQRDLYEARERPSR
8: Canalb_Cdr1|1196-1487     YIYSKIFLVVSAALFNGFSFFKAKNNMQGLQN······QMFSVFMFFIPFNTL·VQQMLPYFVKQRDVYEVREAPSR
9: Abcg2|385-665          FKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDSTGIQN·······RAGVLFFLTTNQCFSSVSAVELFVVEKKLF·IHEYISG
10: Schizpom_Pdr1_TMD2|1...          VAFNIVAGLIIGFSFYKQGVGVEETQN······KMFSAYMLTVASTST·MNGLQPKFIYFRSIYEQYEQNTA
11: Pgp_mouse_TMD2|705-...        WPYFVVGIFCAIINGGL················QPAFSVIFSKVVGVF··TNGGPPETQRQNSNL········


1: Ph_Pdr1|1200-1450      SAMPYAFGQVMIE·LPYLFLQTIIYGVIVYAMIGFEWTVA·········KFFWYLFFMYFT············LLYFTLYGMMTVAVTP
2: Petaxi_Pdr1|1200-1450   SAMPYAFGQVMIE·LPYLFLQTIIYGVIVYAMIGFEWTVA·········KFFWYLFFMYFT············LLYFTLYGMMTVAVTP
3: Nictab_Pdr1_TMD2|1180...   SAIPYAFGQVSIE·IPYIFVQSVFYGIIVYAMIGFEWDVG·········KFFWYLFIMFFT············LLYFTFYGMMGVAVTP
4: Orysat_TMD2|1210-1460   SALPYAFGQVAIE·FPYTLVQSVIYSIIVYSMIGFEWTVA·········KFFWYLFFMFFT············LLYFTFYGMMAVGLTP
5: Nicplu_Pdr1_TMD2|1180...   SAIPYAFAQVLIE·IPYIFVQATVYGLIVYSMIGFEWTVA·········KFFWDFFFMFFT············FLYFTFFGMMTVAVTP
6: Aratha_Pdr1_TMD2|1160...   SALPYALAQVVCE·IPYVLIQTTYYTLIIYAMMCFEWTLA·········KFFWFYFVSFMS············FLYFTYYGMMTVALTP
7: sp|PDR5_TMD2|1200-15...   SWISFIFAQIFVE·VPWNILAGTIAYFIYYYPIGFYSNASAAGQLHERGALFWLFSCAFY············VYVGSMGLLVISFNQ
8: Canalb_Cdr1|1196-1487   SWFAFIAGGITSE·IPYQVAVGTIAFFCWYYPLGLYNNATPTDSVNPRGVLMWMLVTAFY············VYTATMGQLCMSFSE
9: Abcg2|385-665          RVSSYFLGKLLSDLLPMRMLPSIIFTCIVYFMLGLKPKAD·········AFFVMMFTLMMV············AYSASSMALAIAAGQS
10: Schizpom_Pdr1_TMD2|1...   SRTAFIIAFFLVE·AVINCCFATLFFFGWYYPSGFYEFNH···NIPFYGGFAWLMLMIFT·············LYYTTLGIGIATISP
11: Pgp_mouse_TMD2|705-...   ·············FSLLFLILGIISFITFFLQGFTFGKA···GEILTKRLRYMVFKSMLRQDVSWFDDPKNTTGALTTRLANDAAQV


1: Ph_Pdr1|1200-1450      SIAAIISSAFYAVWNLFCGFIVPKTRMPVWWRWYYYICPISWTLYGLIASQFGDIQDRLDT··········NETVEQFIENFFDFK··
2: Petaxi_Pdr1|1200-1450   SIAAIISSAFYAVWNLFCGFIVPKTRMPVWWRWYYYICPISWTLYGLIASQFGDIQDRLDT··········NETVEQFIENFFDFK··
3: Nictab_Pdr1_TMD2|1180...   NVASIVAAFFYGVWNLFSGFIIPRPRMPVWWRWYYWANPVAWTLYGLVASQFGDIQTKLSD··········NETVEQFLRRYFGFK··
4: Orysat_TMD2|1210-1460   HVASIVSSAFYAIWNLFTGFVISRPATPVWWRWYCWICPVAWTLYGLIVSQYGDIVTPMDD··········GIPVNVFVENYFDFK··
5: Nicplu_Pdr1_TMD2|1180...   NVASIVAGFFYTVWNLFSGFIVPRPRIPIWWRWYYWGCPIAWTLYGLVASQFGDLQDPLTD··········QNQTVEQFLRSNFGFK··
6: Aratha_Pdr1_TMD2|1160...   QVAAVFAGAFYGLFNLFSGFVIPRPRIPKWWIWYYWICPVAWTVYGLIVSQYGDVEDTIKVPGM·····ANDPTIKWYIENHYGYD··
7: sp|PDR5_TMD2|1200-15...   ESAANLASLLFTMSLSFCGVMTTPSAMPRFWIFMYRVSPLTYFIQALLAVGVANVDVKCADYELLEFTPPSGMTCGQYMEPYLQLAK·
8: Canalb_Cdr1|1196-1487   DNAANLATLLFTMCLNFCGVLAGPDVLPGFWIFMYRCNPFTYLVQAMLSTGLANTFVKCAEREYVSVKPPNGESCSTYLDPYIKFAG·
9: Abcg2|385-665          SVATLLMTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTAL···················QHNEFLGQNFCPGLNATG·
10: Schizpom_Pdr1_TMD2|1...   GTASIISGTAFVFIQYFNGMIQLPGVIVGFWKWMDALSPYKYFLEGMIGGVLHDAPITCEKFEIHYVDPPPNYSCGEYFSSFLNSSG·
11: Pgp_mouse_TMD2|705-...   ATGSRLAVIFQNIANLGTGIIISL·····IYGWQLTLLLLAIVPIIAIAGVVEMKMLSGGQALKDKKELEGSGKIATEAIENFRTVVSL


1: Ph_Pdr1|1200-1450      ···························HDFVGYVALILVGISVLFLFIFAFSIKTFNFQ
2: Petaxi_Pdr1|1200-1450   ···························HDFVGYVALILVGISVLFLFIFAFSIKTFNFQ
3: Nictab_Pdr1_TMD2|1180...   ···························HDFLGVVAAVLTAYVFMFAFTFAFAIKAFN
4: Orysat_TMD2|1210-1460   ···························HSWLGFVAVVIVAFTMLFAFLFGFAI
5: Nicplu_Pdr1_TMD2|1180...   ···························HDFLGVVAAVIVAFAVVFAFTFALGIKA
6: Aratha_Pdr1_TMD2|1160...   ···························ADFMIPIATVLVGFTLFFAFMFAFGIRTLNFQQR
7: sp|PDR5_TMD2|1200-15...   YLTDENATDTCSFCQISTTNDYLANVNSFYSE················RWRNYGIFICYIAFNYIAGVFFYWLARVPKKN
8: Canalb_Cdr1|1196-1487   YFETRN·DGSCAFCQMSSTNTFLKSVNSLYSE················RWRNFGIFIAFIAINIILTVIFYWL
9: Abcg2|385-665          ··NNPCNYATC······TGEEYLVKQGIDLSP················WGLWKNHVALACMIVIFLTIAYLKLLFLKKYS
10: Schizpom_Pdr1_TMD2|1...   IVYNPEAYSSCQYCPYKNADELMVGFGYHYNH················KWRNFCIMIGYTAFNLGAAIALYYI
11: Pgp_mouse_TMD2|705-...   EQKFETMYAQSLQIPYRNAMKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVTQQLMTFENVLLVFSAIVFGAMAVGQVSSFAPDYA
```

appendix 3: alignment of TMD2 in MAFFT, gap start penalty = 3. PhPDR1 is depicted in pink on the very top of the alignment; the TMDs are the result of the prediction tool TopPred, only to show a relation to the other proteins. The TMDs are colored green and yellow in turn to enable a clearer depiction of their locations. There is no information provided about PaPDR1s TMDs, it was included to improve the alignment.

Feature	Sequence	Residues
N-terminus, NBD1	1-520	520
TMH1	521-542	22
ECD1	553-552	10
TMH2	553-573	21
ICD1	574-608	35
TMH3	609-628	20
ECD2	629-639	11
TMH4	640-660	21
ICD2	661-664	4
TMH5	665-684	20
ECD3	685-752	68
TMH6	753-776	24
NBD2	777-1211	435
TMH7	1212-1231	20
ECD4	1232-1245	14
TMH8	1246-1266	21
ICD3	1267-1297	31
TMH9	1298-1318	21
ECD5	1319-1325	7
TMH10	1326-1345	20
ICD4	1346-1354	9
TMH11	1355-1375	21
ECD6	1376-1420	44
TMH12	1421-1441	21
C-terminus	1442-1452	11

appendix 4: Arrangement of PhPDR1s topology. NBD = nucleotide binding domain, TMH = transmembrane helix, ECD = extracellular domain, ICD = intracellular domain.

Protein	PDB code
Caenorhabditis elegans, ABCB1	4F4C
Mus musculus, ABCB1a	4M1M
Homo sapiens, ABCB10	4AYX
Thermotoga maritima, TM_0288	3QF4
Staphylococcus aureus, SAV1866	2HYD
Salmonella typhimurium, msbA	3B60
Novosphingobium aromaticivorans, atm1	4MRN
Escherichia coli, metN	3TUI

appendix 5: PDB codes of the proteins that were used in the multiple sequence alignment to find a template for the homology model.

appendix 6: Alignment of TMD1 of ScPDR5 and PhPDR1. The transmembrane helices are colored in green and yellow in turn to enable a clearer depiction of their locations. The alignment was created in MOE with the default options.

appendix 7: Alignment of TMD1 of ScPDR5 and PhPDR1. The transmembrane helices are colored in green and yellow in turn to enable a clearer depiction of their locations. The alignment was created in MOE with the default options.

```
1: Pdr5  MPEAKLNNNVNDVTSYSSASSSTENAADLHNYNGFDEHTEARIQKLARTLTAQSMQNSTQSAPNKSDAQSIFSSGVEGVN
2: Pdr1         MEGGEELFRVSSARLSSSNVWRNSAMDVFSRSSREADDEEALKWAALEKLPTYLRIRRGILTEEEGQSREVDIT

1: Pdr5  PIFSDPEAPGYDPKLDPNSENFSSAAWVKNMAHLSAADPDFYKPYSLGCAWKNLSASG·ASADVAYQSTVVNIPYKILKS
2: Pdr1  KL·DLVERRNLLERLIKITDEDNEKFLLKLKERIDRVGLDL···PTIEVRFEHLSVDAEARVGSRALPTVFNFTVNILED

1: Pdr5  GLRKFQ·RSKETNTFQILKPMDGCLNPGELLVVLGRPGSGCTTLLKSISSNTHGFDLGADTKISYSGYSGDDIKKHFRGE
2: Pdr1  FLNYLHILPNRKQPLPILHDVSGIIKPGRMTLLLGPPSSGKTTLLLALAGKLDK·DLKVSGRVTYNGHDMNEFVA··QRS

1: Pdr5  VVYNAEADVHLPHLTVFETLVTVARLKTPQNRIKGVDRESY························ANHLAEVAMATYG
2: Pdr1  SAYISQYDLHIGEMTVRETLAFSARCQGVGAKYEILAELSRREKEANIKPDPDVDIFMKAAWNEGQEANVVTDYTLKILG

1: Pdr5  LSHTRNTKVGNDIVRGVSGGERKRVSIAEVSICGSKFQCWDNATRGLDSATALEFIRALKTQADISNTSATVAIYQCSQD
2: Pdr1  LEICADTIVGDEMVRGISGGQRKRLTTGEMMVGPARALFMDEISTGLDSSTTYQIVNSIRQSIHILQGTAVISLLQPAPE

1: Pdr5  AYDLFNKVCVLDDGYQIYYGPADKAKKYFEDMGYVCPSRQTTADFLTSVTSPSERTLNKDMLKKGIHIPQTPKEMNDYWV
2: Pdr1  TYDLFDDIILLSDGQIVYQGPRENVLEFFEYMGFICPERKGVADFLQEVTS··········RKDQEQYWARREESYKFI

1: Pdr5  KSPNYKELMKEVDQRLLNDDEASREAIKEAHIAKQSKRARPSSPYTVSYMMQVKYLLIRNMWRLRNNIGFTLFMILGNCS
2: Pdr1  TVREFSEAFQAFHIGRKLGDELAVPFDK····SKSHPAALTTKRYGVSKKELLKACTAREYLLMKRNSFVYIFKMIQLTL

1: Pdr5  MALILGSMFFKI·MKKGDT·STFYFRGSAMFFAILFNAFSSLLEIFSLYEARPITEKHRTYSLYHPSADAFASVLSEIPS
2: Pdr1  MASITMTLFLPTEMHRNTTI·DGAVFLGALFYALIMIMFNGFSELALSIMKLPSFYKHRDLLFFPPWAYALPTWILKIPI

1: Pdr5  KLIIAVCFNIIFYFLVDFRRNGGVFFFYLLINIVAVFSMSHLFRCVGSLTKTLSEAMVPASMLLLALSMYTGFAIPKKK
2: Pdr1  TLVEVAIWVCMTYYVIGFEADVGRFF·KQLLLLICVNQMASGLFRLMGALGRNIIVANTFGSFVLLTVLVMGGFVLSRDD

1: Pdr5  ILRWSKWIW··YINPLAYLFESLLINEFHGIKFPCAEYVPRGPAYANISSTESVCTVVGAVPGQDYVLGDDFIR·GTYQY
2: Pdr1  VKKW··WIWGYWISPMMYAQNAIAVNEFLGKSWA···HVP··P···NSTSTET··············LGVSFLKSRGIFP

1: Pdr5  YHKDKWRGFGIGMAYVVFFFFVYLFLCEY····NEGAKQKGEILVFPRSIVKR·MKKRGVLTEKNANDPENVGERSDLSS
2: Pdr1  DARWYWIGAGALIGYVFLFNFLFAVALAYLNPF······KPQAVLSEETVAERNASKRGEVIELSSLGKSSSEKGNDVRR

1: Pdr5  DRKMLQESSEEESDTYGEIGLSKSEAI······FHWRNLCYEVQIKAETR·········RILNNVDGWVKPGTLTALMGA
2: Pdr1  SASSRSMSSRVGSITAADLSKRRGMILPFEPLSITFDDIRYAVDMPQEMKAQGFTEDRLELLRGVSGAFRPGVLTALMGV

1: Pdr5  SGAGKTTLLDCLAERVTMGVITGDILVNGIPRDK·SFPRSIGYCQQQDLHLKTATVRESLRFSAYLRQPAEVSIEEKNRY
2: Pdr1  SGAGKTTLMDVLAGRKTGGYIDGTISISGYPKQQETFARIAGYCEQTDIHSPHVTVYESLQFSAWLRLPREVDTATRKMF

1: Pdr5  VEEVIKILEMEKYADAVVGVAG·EGLNVEQRKRLTIGVELTAKPKLLVFLDEPTSGLDSQTAWSICQLMKKLANHGQAIL
2: Pdr1  IEEVMELIELIPLRDALVGLPGVNGLSTEQRKRLTVAVELVANPSI·IFMDEPTSGLDARAAAIVMRTVRNTVDTGRTVV

1: Pdr5  CTIHQPSAILMQEFDRLLFMQRGGKTVYFGDLGEGCKTMIDYFES·HGAHKCPADANPAEWMLEVVGAAPGSHANQDYYE
2: Pdr1  CTIHQPSIDIFDAFDELLLLKRGGEEIYVGPLGRQSSHLIKYFEGIDGVPKIKDGYNPATWMLEITSVAQEGALGNDFTE

1: Pdr5  VWRNSEEYRAVQSELDWME··RELPKKGSITAAEDKHEFSQSII·YQTKLVSIRLFQQYWRSPDYLWSKFILTIFNQLF
2: Pdr1  LYKNSELYRRNKALIKELSVPASCSKDLYFPTKYSQSFFTQCMACFWKQHWSYWRNPPYTAVRIMFTFFIALMFGTI·FW

1: Pdr5  IGFTFFKAGTSLQGLQN······QMLAVFMFTVIFNPILQQYLPSFVQQRDL····YEARERPSRT······FSWISFIF
2: Pdr1  DLGSRRERQQDLL···NAIGSMYIAVLFLGVQNATTVQPVIAIERTVFYRERAAGMYSAMPYAFGQVMIELPYLFLQTI·

1: Pdr5  AQIFVEVPWNILAGTIAYFIYYYPIGFYSNASAAGQLHERGALFWLFSCAFYVYVGSMGLLVISFNQVAESAANLASLLF
2: Pdr1  ··IYGVIVYAMIGFEWTVAKFFWYL················FFMYFTLLYFTLYGMMTVAVTPNQSIAA···IISSAFY

1: Pdr5  TMSLSFCGVMTTPSAMPRFWIFMYRVSPLTYFIQALLAVGVANVDVKCADYELLEFTPPSGMTCGQYMEPYLQLAKTGYL
2: Pdr1  AVWNLFCGFIVPKTRMPVWWRWYYYICPISWTLYGLIASQFGDIQDRLDTNETVE·········QFIENFFDFK·····

1: Pdr5  TDENATDTCSFCQISTTNDYLANVNSFYSERWRNYGIFICYIAFNYIAGVFFYWLARVPKKNGKLSKK
2: Pdr1  ························HDFVGYVALILVGISVLFLFIFAFSIKTFNFQKR
```

appendix 8: The final alignment of ScPDR5 and PhPDR1. The transmembrane helices are colored in green and yellow in turn to enable a clearer depiction of their locations. The framed areas distinguish between the TMDs and the NBDs, the sequences were precisely at these positions cut for the respective alignment (the alignments of TMD1 and TMD2 are shown in appendix 6 and appendix 7). This alignment was used as an input for the modeller software to create the homology model.

7.2 Abstract

The knowledge about phytohormones and their transport is growing steadily and there is a need for an understanding of the molecular basis of substrate and inhibitor interaction. The class of strigolactones is one of the current centers of attention. They have several roles as stimulation and recognition signals in plants, but the knowledge about their regulating function in shoot branching is quite new. In 2012, the ABC transporter PDR1 in *Petunia hybrida* (PhPDR1) was identified as a strigolactone transporter. The PDR proteins belong to the ABCG family, which exhibits a reverse topology to the other ABC members.

The scientific aim of this project is to elucidate the molecular 3D structure of PhPDR1. At first, we performed a comprehensive investigation to define the transmembrane domains, their containing helices and their locations, which was a challenging task according to the little information yet known. We combined the results extracted from prediction tools, multiple sequence alignments and information from literature to draw the determining conclusions.

The final choice of the template was based on a multiple sequence alignment with the potential templates, including crystallized ABC transporters reported in the PDB and the high-quality, reliably validated homology model of PDR5 in *Saccharomyces cerevisiae*. The latter was already used as a template to model the 3D structure of Cdr1 in *Candida albicans*. PDR5 shows the highest sequence identity percentage and thus was chosen as template for further comparative modelling steps.

We created the homology models with the modeller software and took the decision for a final model based on the assessment methods (molpdf, DOPE, GA341) and by careful analysis of Ramachandran plots and G-factors provided by PDBsum.

In the validation process, we examined the polar residues of the transmembrane helices and the electrostatic potentials of the model to characterize the translocation chamber.

Finally, we conducted a small docking study, where we docked the strigolactone orobanchol (which is evidently transported by PhPDR1) into our model and clustered the results into several groups. This led to first ideas of binding poses and amino acids which could be involved in the binding mode.

7.3 Zusammenfassung

Das Wissen über Phytohormone und deren Transport wächst permanent an, woraus sich die Notwendigkeit eines Verständnisses der molekularen Basis von Substrat und Inhibitor Interaktionen ergibt. Die Klasse der Strigolactone ist derzeit eines der Zentren der Aufmerksamkeit in diesem Zusammenhang. Die Strigolactone spielen eine Rolle als Stimulations- und Erkennungssignale in Pflanzen; die Bekanntheit über ihre regulierende Funktion in der Sprossverzweigung ist hingegen relativ neu. Im Jahr 2012 wurde der ABC-Transporter PDR1 in *Petunia hybrida* (PhPDR1) als Strigolactontransporter identifiziert. Die PDR Proteine gehören zur ABCG Familie, welche eine umgekehrte Topologie im Vergleich zu den anderen ABC Mitgliedern aufweist.

Das wissenschaftliche Ziel dieses Projektes ist die Beleuchtung der 3D Struktur von PhPDR1. Zuerst wurde eine umfangreiche, aufgrund der mangelnden Informationen sehr herausfordernde, Untersuchung durchgeführt, um die transmembranären Domänen mit den enthaltenen Helices bzw. deren Lokalisationen zu definieren. Dazu wurden die Ergebnisse von Vorhersagetools, von multiplen Sequenzalignments und die Kenntnisse aus der Literatur kombiniert um die notwendigen Schlüsse zu ziehen.

Die endgültige Wahl des Templates basierte auf einem multiplen Sequenzalignment mit den vorhandenen Kristallstrukturen der ABC-Transporter in der PDB und dem hochqualitativen, sorgfältig validierten Homologiemodell von PDR5 in *Saccharomyces cerevisiae*, welches auch schon als Template für ein Homologiemodell von Cdr1 in *Candida albicans* gedient hatte.

PDR5 weist die höchste Sequenzidentität auf und wurde deshalb für die weiteren Modellingschritte ausgewählt.

Die 3D Modelle wurden in der Modeller Software erstellt, die endgültige Wahl erfolgte mit Hilfe der implementierten Assessmentmethoden (molpdf, DOPE, GA341) und durch eingehende Analyse der Ramachandranplots.

Im Validierungsprozess wurden die polaren Aminosäurereste der transmembranären Helices und die elektrostatischen Potenziale untersucht.

Zum Schluss wurde eine kleine Dockingstudie durchgeführt, in welcher das von PhPDR1 transportierte Strigolacton Orobanchol in das finale 3D Modell gedockt

wurde. Die daraus resultierenden Posen wurden in Gruppen geclustert, was zu ersten Ideen über Bindungsposen und in den Bindungsmechanismus involvierte Aminosäuren führte.

7.4 Curriculum vitae

Education and Qualifications

Jan-Jun 2015	participation in "Naturtalente", a high potential program for extraordinary science students, organized by the University of Vienna and UNIPORT career service
Jan 2012	Diploma of "Tutoring and Mentoring" for first year students
since 2007	Diploma study Pharmacy at the University of Vienna
2004-2006	extra occupational University Course for Elementary Music Education at the University of Music and Performing Arts Vienna Graduation with Distinction 2006-03-09
1998-2003	school for kindergarten teachers (BAKIP) "Maria Regina" "Reife- und Diplomprüfung für Kindergärten und Horte" Graduation with Distinction 2003-06-16

Practical experience in the pharmaceutical field

since Sep 2012	research internship in the Pharmacoinformatics Research Group, Univ.-Prof. Dr. Gerhard Ecker, Department of Pharmaceutical Chemistry, University of Vienna
Jul/Aug 2012	internship in the division "Quality Control/Chemistry Photo-spectrometry and wet chemistry, Baxter AG, 1220 Vienna
Jul/Aug 2011	internship in the division QC/Chemistry PSN, Baxter AG, 1220 Vienna
Oct 2007-Oct 2012	employed pharmacy assistance for 10h per week, "Apotheke zur Universität", 1090 Vienna, Dr. Tivadar Mikes
2006-2007	fulltime employee in the "Apotheke zur Universität" , 1090 Vienna, Dr. Tivadar Mikes; second year of the school for "Pharmazeutisch-kaufmännische Assistenz", Graduation with Distinction

Posters

May 2015	10th European Workshop in Drug Design, Certosa di Pontignano, Siena, Italy Poster: Elucidation of the structure of a strigolactone ABC-transporter in *Petunia hybrida* (PhPDR1)
Sep 2014	7th SFB 35 - Transmembrane Transporters in Health and Disease - Symposium in Vienna Poster: Towards a model for the structure of a strigolactone ABC-transporter: PDR1 in *Petunia hybrida* (PhPDR1)
Jan 2014	47th week of further education organized by the Austrian Chamber of Pharmacy in Schladming Poster: Towards higher crop productivity: Predicting inhibitors for AtD14 to block the Strigolactone pathway in plants

Engagement in the Austrian student association

Since Sep 2012	vice chair of the working group to develop a new pharmacy curriculum at the University of Vienna
2012-2013	Tutoring and Mentoring for peers in the first year
since Oct 2011	member of the student's representation Pharmacy (various projects, many activities in different committees like study and faculty conferences or appointment committees)

Others

Sep 2011-Dec 2014	member of the "Supernannies.at", an association of independently operating educators which afford quality childcare and competent family support. http://supernannies.jimdo.com/
2003-2006	kindergarden teaching and after-school care at "Kindertagesheim Waldeckgasse", 1180 Vienna (institute of the "Wiener Kindergärten", MA 10)