



universität  
wien

# DISSERTATION

Titel der Dissertation

## **Combined Ligand- and Structure-based Studies on inhibitors of P-glycoprotein**

Verfasserin

M.Phil. Ishrat Jabeen

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer.nat.)

Wien, 2011

Studienkennzahl lt. Studienblatt:

A 091 419

Dissertationsgebiet lt. Studienblatt:

Dr.-Studium der Naturwissenschaften Chemie

Betreuerin / Betreuer:

Univ.-Prof. Dr. Gerhard F. Ecker



## **Acknowledgements**

The present study was carried out mainly at the Pharmacoinformatics Research group, Department of Medicinal Chemistry, University of Vienna, Austria.

I would like to express my deep gratitude to my supervisor Prof. Gerhard Ecker, for his coherent and illuminating instructions, guidance, fascinating lectures and encouraging support during this project.

I am sincerely thankful to Prof. Manuel Pastor, at the Computer-Assisted Drug Design Laboratory (CADD) of the Research Unit on Biomedical Informatics (GRIB), Universitat Pompeu Fabra, Barcelona, for his guidance and sharing knowledge during my internship in his research group.

I am very grateful to Prof. Thomas Erker and his colleagues, and to Karin Pleban and Penpun Wetwitayaklung, for the synthesis of P-glycoprotein Inhibitors.

I am extremely thankful to Prof. Peter Chiba and his co-workers at the Medical University of Vienna for performing pharmacological studies on the synthesised compounds. I duly acknowledge his critical reading of our manuscripts and helpful suggestions.

I owe my sincere gratitude to my colleagues and friends at the Pharmacoinformatics Research Group, Barbara Zdrazil, Vasanthanathan Poongavanam, Marta Pinto, Michael Demel, Lars Richter, Rita Schwaha, Yogesh Aher, Daniela Digles, Andrea Schiesaro, Freya Klepsch, Rene Weissensteiner, Andreas Jurik, Petra Urach, Daria Tsareva, Amir Seddik, Katharina Prokes, Victoria Slubowski, Sabine Mydza and all members of the administration staff for their cheerful company, help and for sharing their knowledge.

I would like to acknowledge the assistance of all my teachers, colleagues and friends during different stages of my whole academic period.

My hearty thanks go to my family, especially my loving parents, brothers and sisters for their infinite love, moral support and keen interest in my studies.

Finally I am grateful to the Higher Education Commission (HEC) of Pakistan for financial support and for providing me an opportunity to study abroad.

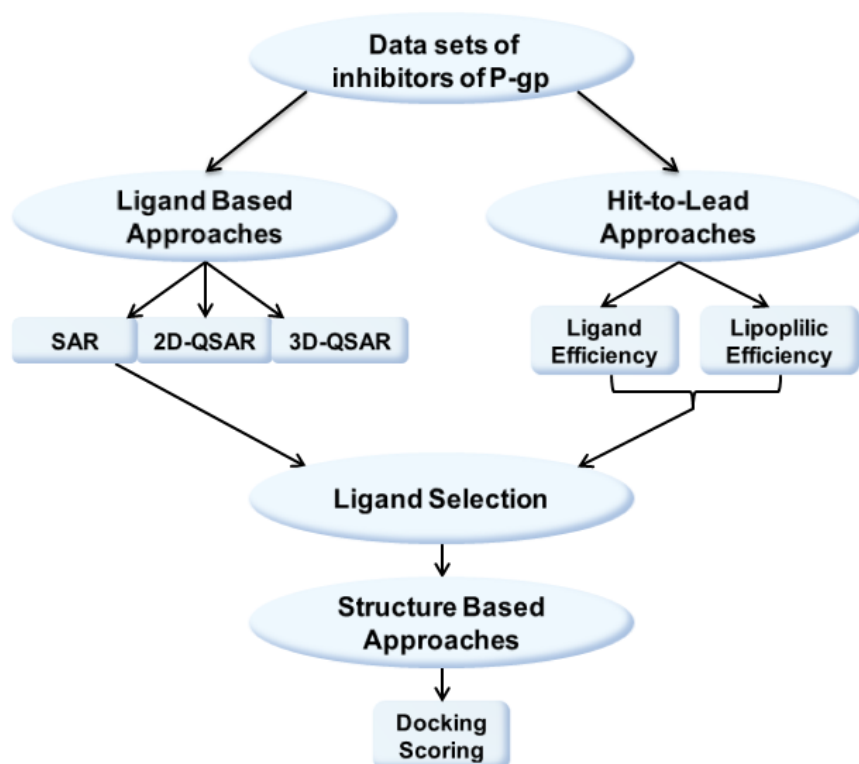
## TABLE OF CONTENTS

|                       |              |
|-----------------------|--------------|
| Acknowledgements      | Page i       |
| Contents              | Page ii-v    |
| Curriculum Vitae      | Page 227-230 |
| Abstract in German    | Page 231-232 |
| Abstract in English   | Page 233-234 |
| List of Abbreviations | Page 235     |

### Contents

#### CHAPTER 1

|   |           |
|---|-----------|
| P-glycoprotein: <i>In Silico</i> Models | Page 1-43 |
|---|-----------|



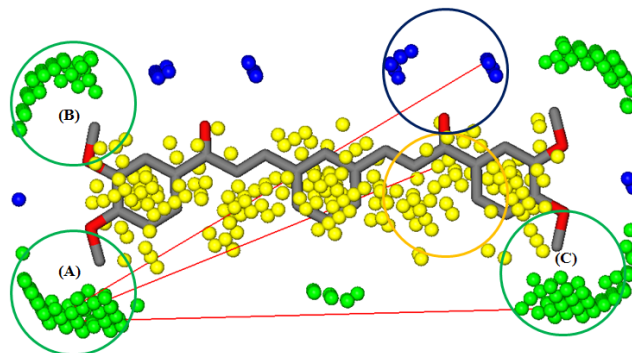
This study comprises both ligand as well as structure based approaches to get an insight about 3D structural requirements of ligands and their interaction pattern in the binding pocket of P-glycoprotein (P-gp). These Models might be applicable both for the design of new inhibitors and for understanding of the structure and function of P-gp.

① **Information:** This chapter summarizes and covers all studies which were done in this thesis

## CHAPTER 2

### Synthesis, ABCB1 Inhibitory Activity and 3D-QSAR Studies of a Series of New Chalcone Derivatives

Page 44-69



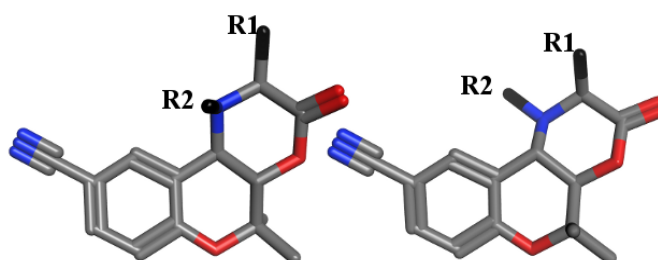
In this part both 2D- and 3D-QSAR models were derived by using general physicochemical and GRID-Independent Molecular Descriptors (GRIND) for prediction of chalcone/ABCB1 interaction. Final models were used to predict the activity of a set of newly synthesized chalcone derivatives.

① **Information:** This chapter was a pre-submission phase of a manuscript by Brunhofer Gerda, **Jabeen Ishrat**, Parveen Zahida, Berner Heinz, Manuel Pastor, Chiba Peter, Erker Thomas and Ecker Gerhard F.

## CHAPTER 3

### Synthesis, Biological Activity and Quantitative Structure-Activity Relationship Studies of a Series of Benzopyranes and Benzopyrano[3,4-b][1,4]oxazines as Inhibitors of the Multidrug Transporter P-glycoprotein

Page 70-103



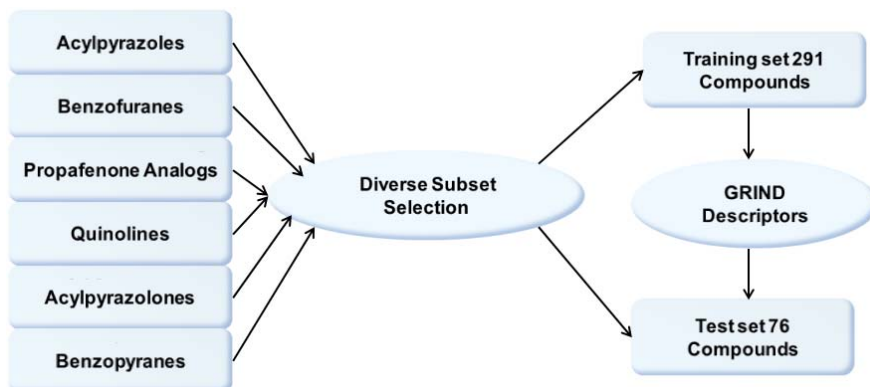
In this part a data set of enantiomerically pure benzopyrano[3,4-b][1,4]oxazines were used to establish predictive models for P-glycoprotein inhibitors. This includes 2D- and 3D-QSAR models, using simple physicochemical as well as GRIND molecular descriptors.

① **Information:** This chapter was submitted to *European Journal of Medicinal Chemistry*, 2011 by **Ishrat Jabeen**, Penpun Wetwitayaklung, Peter Chiba, Manuel Pastor and Gerhard F. Ecker.

## CHAPTER 4

### Development of a Predictive 3D-QSAR Model for a Structurally Diverse Set of Inhibitors of P-glycoprotein (P-gp)

Page 104-116

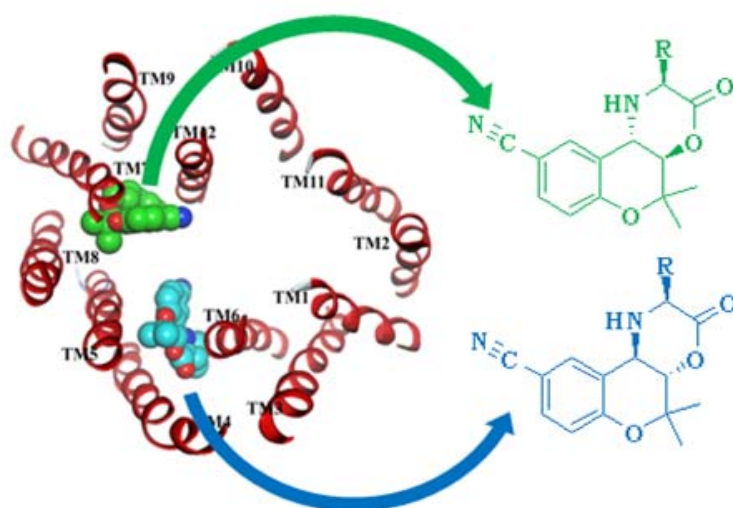


In this chapter the GRIND approach was used to build a predictive 3D-QSAR model for an extended data set of P-glycoprotein inhibitors belonging to different chemical scaffolds.

## CHAPTER 5

### Probing the Stereoselectivity of P-glycoprotein– Synthesis, Biological Activity and Ligand Docking Studies of a Set of Enantiopure Benzopyrano[3,4b][1,4]oxazines

Page 117-121



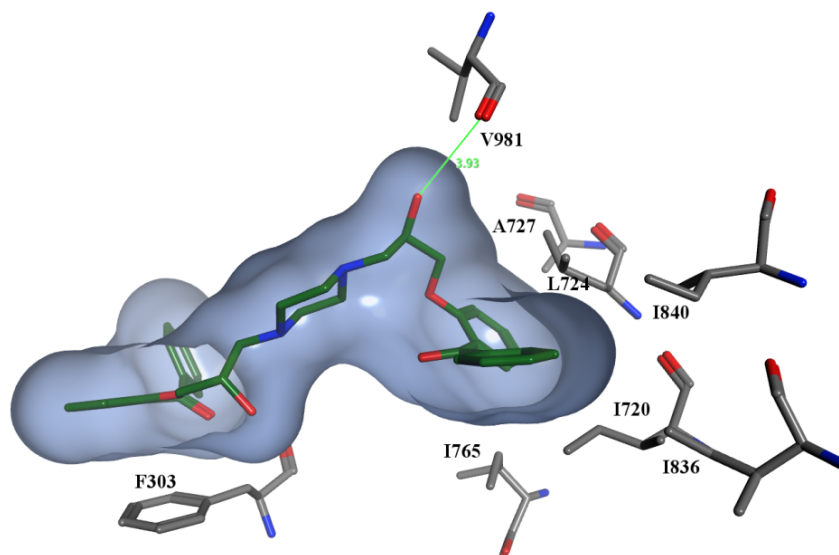
In this chapter a data set of diastereoisomers of benzopyrano[3,4-b][1,4]oxazines were docked into a homology model of P-glycoprotein to probe stereoselective interaction of diastereoisomeric pairs.

① **Information:** This chapter was published in *Chem Comm*, 2011, Volume 47, 2586-2588 by **Jabeen Ishrat**, Wetwitayaklung Penpun, Klepsch Freya, Parveen Zahida, Chiba Peter, Ecker Gerhard F.

## CHAPTER 6

### Structure-Activity Relationships, Ligand Efficiency and Lipophilic Efficiency Profiles of Benzophenone-Type Inhibitors of the Multidrug Transporter P-glycoprotein

Page 122-156



In this chapter a data set of benzophenone analogs along with some compounds in clinical investigations were used for ligand efficiency and lipophilic efficiency profiling, in order to get insights about the importance of these parameters for the design of P-gp inhibitors.

❶ **Information:** This chapter was submitted in *Journal of Medicinal Chemistry*, 2011 by **Ishrat Jabeen**, Karin Pleban, Peter Chiba and Gerhard F. Ecker.

## CHAPTER 7

### Pharmacoinformatic Approaches to Design Natural Product Type Ligands of ABC-Transporters

Page 157-168

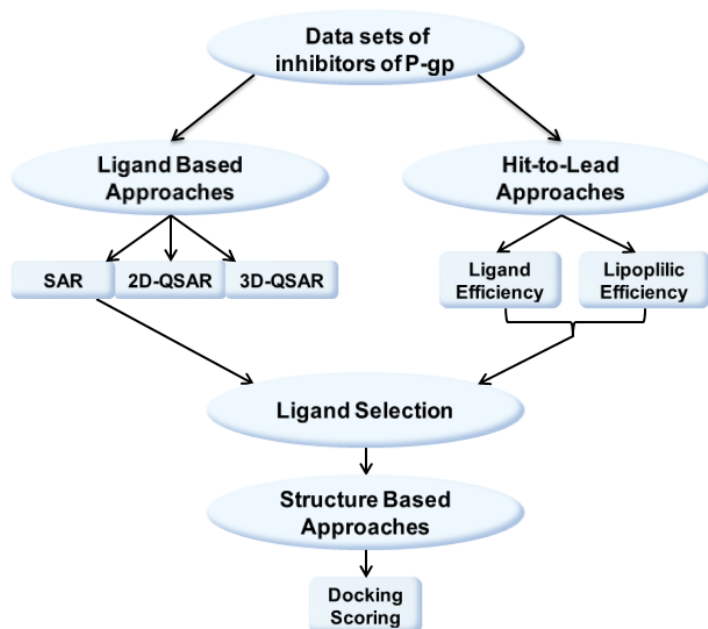
❶ **Information:** This chapter was published in *Current Pharmaceutical Design* (2010), Volume 16, 1742-1752 by Klepsch Freya, **Jabeen Ishrat** and Ecker Gerhard F.

## 📖 APPENDIX

|     |                                  |              |
|-----|----------------------------------|--------------|
| A1. | Supplementary data for chapter 2 | Page 170-172 |
| A3. | Supplementary data for chapter 5 | Page 173-193 |
| A4. | Supplementary data for chapter 6 | Page 194-226 |







This study comprises both ligand as well as structure based approaches to get an insight about 3D structural requirements of the ligands and their interaction pattern in the binding pocket of P-glycoprotein (P-gp). These models might be applicable both for the design of new inhibitors and for understanding of the structure and function of P-gp.

## Contents

1. Introduction
2. P-gp Drug Binding Sites
3. Current State of the Art of P-gp Computational Models
  - 3.1. Ligand Based Approaches
  - 3.2. Structure Based Approaches
4. Aim of the Work
5. 2D-QSAR Models
6. 3D-QSAR Models
7. P-gp and Stereoselectivity
8. Ligand Efficiency/Lipophilic Efficiency Profiles of P-gp Inhibitors
9. Interaction Pattern of P-gp Inhibitors in the Binding Pocket
10. Ligand Efficiency and Lipophilic Efficiency Profiles along Different Entrance Pathways

**11. Summary and Outlook**

**Acknowledgement**

**References**

- ① **Information:** This chapter summarizes and covers all studies which were done in this thesis

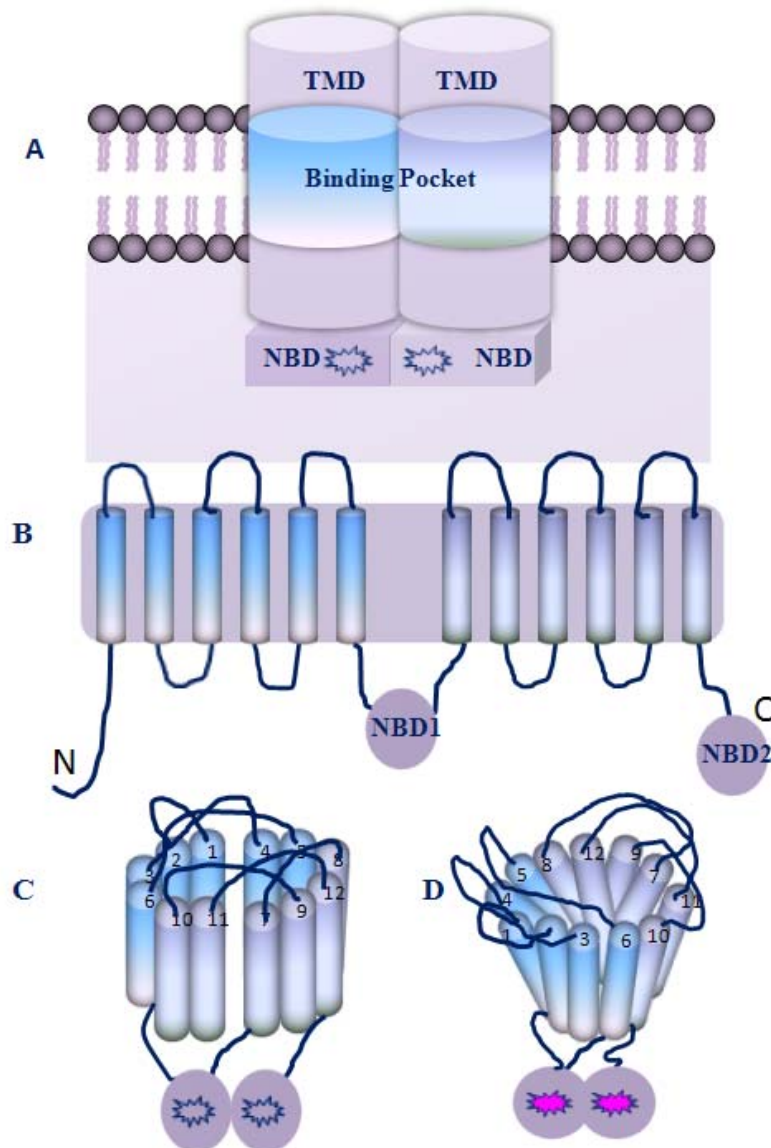
**1. Introduction**

Successful treatment of cancer and infections caused by pathogenic microorganisms is very often compromised by the development of resistance to multiple chemotherapeutic agents, widely known as multidrug resistance (MDR).<sup>1</sup> Molecular biologists, biochemists and oncologists in the last 30 years realized that this phenomenon is due to the expression of plasma membrane “pumps,” which actively extrude various cytotoxic agents from the cells due to an increased active efflux.<sup>2-5</sup> This accelerated efflux is an ATP dependent process resulting from overexpression of membrane bound ATP binding cassette (ABC) transporters.<sup>6-9</sup> In humans, the three major types of multidrug resistance (MDR) transporters include members of the ABCB (ABCB1/MDR1/P-glycoprotein), ABCC (ABCC1/MRP1, ABCC2/MRP2) and the ABCG (ABCG2/MXR/BCRP) subfamily.<sup>10</sup>

P-glycoprotein (P-gp/ABCB1) is a classical ABC (ATP Binding Cassette) transporter which is most intensively studied and has remarkably broad substrate specificity. Being highly promiscuous, it transports numerous structurally and functionally unrelated compounds including substrates/inhibitors of CYP3A4 and of the hERG (potassium ion channel).<sup>8,11,12</sup> It is expressed in epithelial cells of the kidney, liver, intestine, pancreas, colon, as well as at the, blood–tissue barriers (blood–brain barrier, blood–testis barrier,<sup>13</sup> blood cerebrospinal fluid (B-CSF), and blood-placenta barrier), thus underscoring its role in maintaining concentration gradients of toxic compounds at physiological barriers.<sup>14</sup> P-gp and its ligands (substrates and inhibitors) are therefore extensively studied both with respect to reversing multidrug resistance in tumors and for modifying ADME-Tox properties of drug candidates, such as blood-brain barrier penetration.<sup>15</sup>

Models describing the structure and function of P-gp rely on biochemical experiments, mutagenesis data, low resolution X-ray structures, and the atomic level structures of various other ABC transporters. Analysis of primary amino acid sequence of P-gp delineates tandem repeats of transmembrane domains, an ATP binding cassette and a linker region connecting the two homologous parts of the protein. Each repeat consists of a transmembrane domain (TMD), containing six helices, followed by a nucleotide binding domain (NBD).<sup>16,17</sup> The two halves form a single transporter with a pseudo-two fold symmetry, in which the transmembrane helices define a “pore” which

is open to both the cytoplasm and the inner leaflet for substrate translocation, and the nucleotide binding sites harvest the energy of ATP binding and hydrolysis (Figure 1). Recently the first X-ray structure of mammalian P-gp has been published which supports this topology.<sup>18</sup>



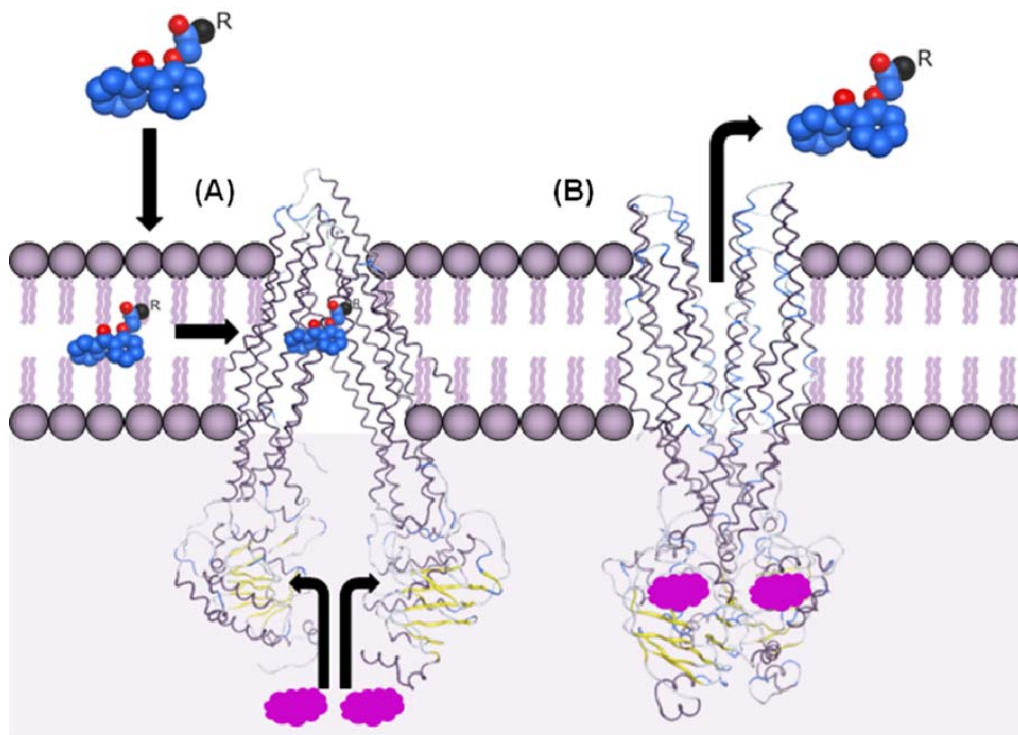
**Figure 1. Schematic representation of the structural topology of P-glycoprotein** (A) Two cylindrical transmembrane domains (TMD) containing a large substrate binding pocket. Two nucleotide binding domains (NBD) holding ATP binding sites, NBDs are responsible for ATP hydrolysis and drug efflux. (B) Representing full architecture of the transmembrane domains, each consisting of six transmembrane helices, followed by cytoplasmic nucleotide binding domains (NBD). A linker region connects NBD1 with TMD2. (C) Binding pocket of P-gp in nucleotide free inward-facing conformation as described by Aller *et al.*,<sup>18</sup> all transmembranes helices are numbered and connected by the linker region. (D) Binding pocket in ATP bound state of protein when it is more exposed to the extracellular fluid and results in translocation of substrates out of the cell.

**2. P-gp Drug Binding Sites**

The role of the TMDs for substrate recognition in P-gp has been subject of many investigations. About a decade ago, two major photo-binding regions were identified using several techniques such as photoaffinity labeling studies, electron microscopic images and epitope mapping.<sup>19-23</sup> The main regions captured comprise TM segments 5/6 in N-terminal and 11/12 in C-terminal part of P-gp. It was further demonstrated that even mutants lacking the NBDs were still able to interact with certain substrates.<sup>24</sup> Cystein scanning mutagenesis in combination with employment of thiol reactive substrates further identified TM segments 4 and 10 to directly interact with certain substrate molecules.<sup>24,25</sup> Later on in our group, photoaffinity labeled benzophenones were used to characterize the drug-binding domain of P-gp. TM 3, 5, 8 and 11 were identified as highly labeled transmembrane regions.<sup>26,27</sup> The question of one or two binding sites remains elusive, but the data suggest that there indeed are more than one drug- interaction sites. The overall assumption in this case is that P-gp possesses a huge binding pocket with at least more than two distinct binding sites, with TM 6 as main interaction helix. Well characterized are the binding sites of Rhodamine and Hoechst 33342, the so called R- and the H-site.<sup>28,29</sup>

The recently published structure of human P-gp using cystine scanning mutagenesis identified two bundles of six transmembrane helices (TMs 1 to 3, 6, 10, 11 and TMs 4, 5, 7 to 9, 12) as shown in figure 1C. This results in a large internal cavity in the lipid bilayer which opens to both the cytoplasm and the inner leaflet. Two portals 4/6 and 10/12 allow access for entry of hydrophobic molecules directly from the membrane and accommodate them at two different positions.<sup>18</sup> This is consistent with a recent observation of two pseudosymmetric drug translocation pathways in the binding cavity.<sup>30</sup> Furthermore, the substrate binding sites appear to exist in two states, a high- and a low- affinity state, which in case of P-gp are in equilibrium. Affinities can be switched from either binding at an alternate site<sup>31</sup> or even during the catalytic cycle.<sup>11,32</sup> Both affinity states have been shown in figure 2, with a proposed mechanism of transport. Although over the past decade many new insights on details of the substrate binding site have been gained,<sup>33,34</sup> ATP-driven dimerization of the NBDs has been recognized as playing a key role in the catalytic cycle of ABC proteins. However, how ATP hydrolysis during the catalytic cycle is coordinated between the two NBDs at the molecular level and how this is coupled to drug transport is still not understood. Also

the exact mechanism of communication between TMs and NBDs remains elusive, but it is strongly suggested that the intracellular loops (ICL) couple drug binding to ATP hydrolysis.<sup>35,36</sup>

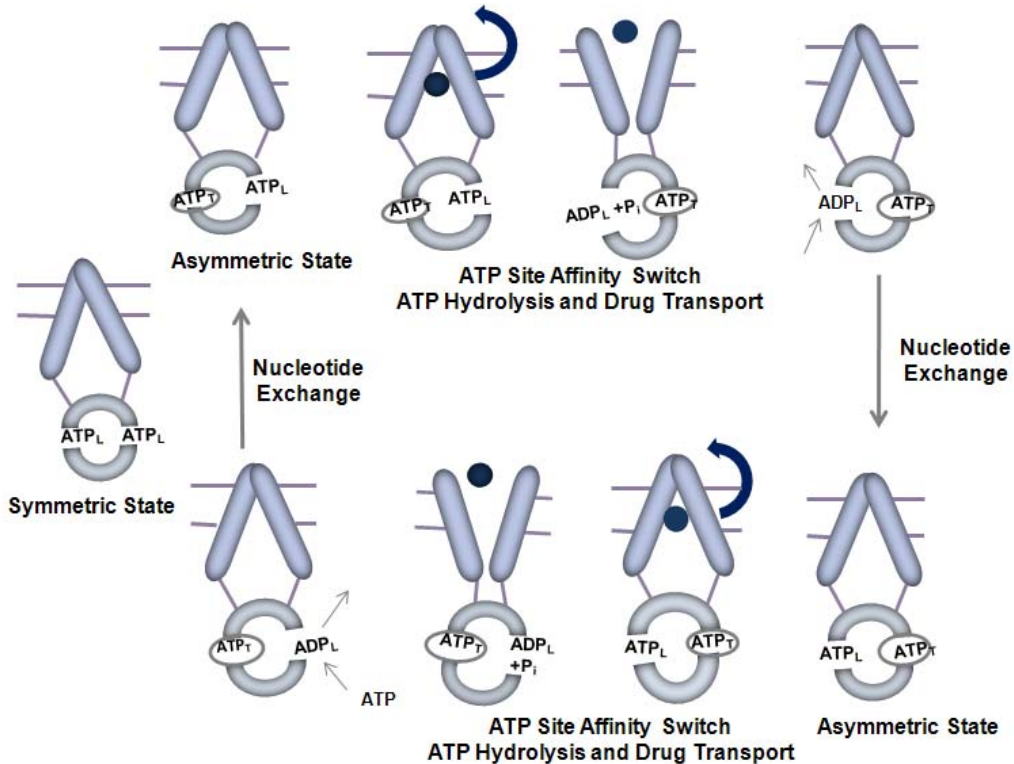


**Figure 2. Proposed model for the mechanism of substrate transport across cell membranes** (A) Representing the high affinity state where the ligand enters the binding cavity from the inner leaflet of the membrane. (B) Low affinity state where ATP (magenta colour) binds to the NBD following a large conformational change and release of the ligand into the extracellular space.

Several models attempt to show how this transport process might work. In the “power stroke model” the substrate enters the binding pocket from the inner leaflet of the membrane and induces nucleotide binding. This promotes the formation of an NBD dimer which results in the power stroke for reorientation of the drug-binding sites from high-affinity inward-facing orientations to outward facing low-affinity sites.<sup>33,37-40</sup> Several studies using ATP analogs have shown that there are alterations in packing of the TM  $\alpha$ -helices in a way that the binding site reorientates towards the extracellular fluid resulting in the release of the substrate.<sup>41-43</sup>

Senior *et al*, proposed that drug transport is coupled to relaxation of a high chemical potential conformation of the catalytic site containing bound  $Mg^{+2}$ , ADP and Pi, which is generated by the process of ATP hydrolysis itself rather than being coupled to nucleotide binding.<sup>44</sup> According to the “alternate site mechanism” one out of two NBD

active sites is able to hydrolyze ATP at any point in time during the catalytic cycle.<sup>45</sup> This mechanism requires that all reaction intermediates are asymmetric, which is in agreement with the recently proposed “site switching model” of substrate transport.<sup>46</sup> This model proposes that substrate translocation across the membrane is driven by ATP hydrolysis. According to this model one of the two NBD dimer interfaces is always in the tightly bound occluded state at all times (Figure 3). The NBD dimer thus never dissociates during catalytic turnover. As only one-half of the interface opens after hydrolysis of an ATP molecule, asymmetry of the structure is maintained continuously throughout the transport cycle.



**Figure 3 Proposed substrate transport mechanism of P-gp**, taken and modified from Siarheyeva *et al.*<sup>46</sup> It has been proposed that catalytically active P-gp maintains its asymmetry and one out of two NBDs active sites is able to hydrolyze ATP at any point in time during the catalytic cycle. One ATP molecule is tightly bound ( $ATP_T$ ) in one of the two NBDs which results in closure of the dimer interface in NBD1. The tightly bound ATP molecule undergoes hydrolysis, which provides the energy for movement of the drug into the extracellular fluid. ATP hydrolysis converts the tightly bound ATP to ADP and  $P_i$ , which are now loosely bound ( $ADP_L$ ), resulting in opening of the dimer interface in NBD1. The other catalytic site simultaneously switches to the high affinity state where a second ATP molecule tightly binds, resulting in closure of the interface at NBD2.  $P_i$  dissociates from the catalytic site of NBD1 first, followed by ADP, which is replaced by another molecule of loosely bound ATP to achieve the asymmetric occluded state once again. A second round of ATP hydrolysis and drug transport then takes place at NBD2.

Recently cross-linking analysis by Loo *et al*, suggests that P-gp cross-linked between residues 175 and 820 in the cytoplasmic portion of transmembrane helices 3 and 9 is able to hydrolyze ATP in the absence of substrates. In addition, basal ATPase activity is stimulated by drug substrates, indicating that under conditions in which the NBDs cannot disassociate completely, the transporter is still able to bind drugs.<sup>47</sup> The crystal structure of a bacterial ABC transporter (Sav1866) in the outward-facing conformation agrees well with the recent cross-linking analysis and most likely reflects a physiologically relevant state.<sup>48</sup> However, the crystal structure of the open conformation of mouse P-gp<sup>18</sup> could only be obtained in the absence of ATP, ATP analogs or magnesium, which is unlikely in the physiological conditions. Nevertheless, this study gives hope to seek crystal structures of other mammalian P-gps that diffract X-rays to higher resolution and that represent more physiologically relevant conformations.<sup>49</sup>

Within the past two decades numerous modulators of P-gp mediated drug efflux have been identified and several entered clinical studies up to phase III.<sup>50,51</sup> However, up to now no compound achieved approval, which is mainly due to severe side effects and lack of efficacy. This further emphasizes the physiological role of efflux transporters in general and P-gp in particular<sup>52</sup> and stresses the need for a more detailed knowledge on the structure and function of these proteins and the molecular basis of their interaction with small molecules. Discovery of new drug entities is costly and time demanding, for this reason reliable *in silico* tools for recognition of P-gp substrates and inhibitors can be valuable during the early phases of drug discovery.

### 3. Current State of the Art of P-gp Computational Models

Both ligand- and structure-based approaches have been undertaken to explore the molecular basis of ligand-protein interactions. In order to probe structural features important for P-gp inhibitor activity, extensive SAR and QSAR studies have been performed. These include Hansch analyses, GRIND, CoMFA and CoMSIA, HQSAR studies, pharmacophore modeling as well as neural network based classification approaches. Key amino acid residues involved in ligand interaction have been identified by homology modeling, site directed mutagenesis and docking protocols.

In the present work propafenone analogs and related compounds were used to study 3D pharmacophoric features of inhibitors of P-gp and their ligand-protein interaction profiles. We also analyzed a set of dihydrobenzopyranes, which, in contrast to our main



lead compound propafenone, offer the advantage of remarkably reduced conformational flexibility and thus might be versatile molecular tools for probing stereoselective differences of drug/P-gp interaction. Finally, compounds were prioritized by ligand efficiency and lipophilic efficiency profile studies. These parameters normalise biological activity towards size and logP, thus helping to identify the derivatives with the best activity/logP (or size) ratio. In this part of the discussion, the current status of *in silico* models for prediction of P-gp inhibitors will be addressed.

### 3.1. Ligand Based Approaches

P-glycoprotein and its congeners are membrane-spanning proteins and thus until very recently only little structural information is available. Therefore, in lead optimization programs, mainly ligand-based approaches have been pursued. These include both 2D- and 3D-QSAR studies on structurally homologous series of compounds, such as verapamil analogs, triazines, acridonecarboxamides, phenothiazines, thioxanthenes, flavones, dihydropyridines, propafenones and cyclosporine derivatives.<sup>53-56</sup> Extensive QSAR studies have been mainly performed on phenothiazines and propafenones. They include Hansch- and Free-Wilson analyses,<sup>57</sup> hologram QSAR, CoMFA, and CoMSIA studies,<sup>58</sup> as well as nonlinear methods of classification,<sup>59</sup> similarity-based approaches<sup>60</sup> and most recently ligand efficiency and lipophilic efficiency profiling of ABC transporters. Most of these studies point towards an importance of H-bond acceptors and their strength, a certain distance between aromatic moieties and H-bond acceptors, as well as the influence of global physicochemical parameters, such as lipophilicity and molar refractivity.<sup>57,61,62</sup>

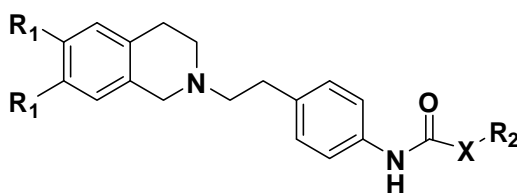
However, several studies showed that lipophilicity might influence pharmacological activity in a space-directed manner rather than as a general physicochemical determinant.<sup>63,64</sup> This space-directedness might be indicative of different orientations of molecules within the binding pocket of P-gp. Pleban *et al.*,<sup>65</sup> further reported that distribution of hydrophobicity within the molecules influences their mode of interaction with P-gp. Moreover, she hypothesized that the substructure with the higher partial lipophilicity acts as an anchor for the compounds in the lipid membrane and thus influences the orientation of the molecule in the binding pocket, which is further confirmed by König *et al.*, by using hydrophobic moments as QSAR descriptors.<sup>66</sup> Hydrophobic distribution within the molecules along with molecular weight, number of

rotatable bonds and energy of highest occupied orbital  $E_{\text{homo}}$ , was further identified as determinant for P-gp inhibitory potency by Wang and colleagues.<sup>67</sup>

Seelig<sup>68,69</sup> more explicitly defined the number of H-bond acceptors and their fixed spatial distance while working on a database of 100 P-gp ligands. She defined two patterns of H-bond acceptor groups. Type I units have two H-bond acceptors separated by a distance of  $2.5 \pm 0.3 \text{ \AA}$  and type II units have two or three H-bond acceptors where two groups are spaced  $4.6 \pm 0.6 \text{ \AA}$  apart. She proposed that P-gp ligands contain at least one type I or type II unit, or may contain both. Ecker *et al.*,<sup>70</sup> followed up on Seelig's findings and found an excellent correlation between the summed electron donating strength of the substituent at the nitrogen atom and its potency as an inhibitor. Pajeva and Wiese further considered the strengths of the H-bond groups, rather than just their numbers, by calculating H-bond acceptor capacities.<sup>71</sup>

Tariquidar, which is among the third generation representatives and one of the most active MDR modulators, has been identified to bind at the same site as the P-gp substrate Hoechst 33342.<sup>31,72</sup> Globisch *et al.*,<sup>73</sup> identified the structural features of tariquidar analogues<sup>74</sup> contributing to MDR activity by Free-Wilson analysis and 3D-QSAR, such as CoMFA and CoMSIA. H-bond acceptor, steric, and hydrophobic fields were identified as most important 3D properties, which is in line with previous studies about the role of the steric, hydrophobic and H-bond interactions of MDR modulators.<sup>63,64,75</sup> Later on, similar features were identified by Müller *et al.*, for a set of compounds comprising a tetrahydroisoquinoline-ethyl-phenyl-amide substructure.<sup>76</sup> Overall no correlation has been identified between logP and pIC<sub>50</sub> values of tariquidar analogs. In contrast to logP, the hydrophobic indices, generated by CoMSIA models, showed a relatively high correlation with pIC<sub>50</sub> values. The hydrophobic field alone produced models with acceptable statistics.<sup>77</sup>

Recently, Gadhe *et al.*, reported topomer CoMFA<sup>78</sup> and HQSAR models for newly synthesized third generation MDR modulators,<sup>79</sup> containing a tetrahydroisoquinoline-ethyl-phenylamine nucleus as shown in figure 4.



**Figure 4.** Common scaffold of newly synthesized third generation MDR modulators<sup>79</sup>

Contour map analysis of the topomer CoMFA model showed that bulky and more electropositive substituents on the amide linker (X) are responsible for higher potency, which was further supported by the HQSAR atomic contribution map. However, HQSAR additionally revealed that a dimethoxy group at R1 position is important for an inhibitory effect. Both HQSAR and topomer CoMFA underscore the importance of the aromatic dimethoxy and nitro groups for the inhibitory effect. This information was utilized to design some compounds which were predicted more potent than existing ones.<sup>80</sup>

Ekins *et al.*,<sup>81,82</sup> used chiral compounds for model development, but stereochemical aspects of ligand-protein interactions were not addressed and/or racemic mixtures were used instead of the single enantiomers. Both studies described similarity of pharmacophoric features of MDR modulators of digoxin and vinblastine transport and verapamil binding. Important features of these substrates include multiple hydrophobic and hydrogen bond acceptor features, which, after alignment suggest commonality in their binding sites. However, despite some similarities between the two models presented, a direct comparison of the distance matrices of the pharmacophoric features revealed some differences. This reflects that the large binding site in P-gp might have several points able to participate in hydrophobic and H-bond interactions and might undergo conformational changes so that both ligand and receptor adopt a best fit for each other. Langer and co-workers,<sup>83</sup> additionally highlighted the importance of a positive ionizable feature, corresponding to the tertiary nitrogen atom in the molecules. However, Ecker *et al.*,<sup>70</sup> demonstrated that the basic nitrogen atom in propafenone-type modulators of P-gp represents an important pharmacophoric group and that it interacts on basis of H-bond acceptor features rather than charged ones. Later on, pharmacophore models in combination with *in vitro* approaches were used to filter out P-gp substrates or inhibitors and identify potential therapeutic P-gp inhibitors.<sup>84</sup>

Although the majority of pharmacophoric models are reasonably accurate in predicting inhibitors and substrates of P-gp<sup>71,81-83,85-87</sup> and in identifying the minimal structural requirements of MDR modulators, they rarely reached high accuracy when applied to large data sets with nonlocal series. 3D-QSAR methods on the other hand need a proper alignment of the molecules. Moreover, pharmacophoric as well as CoMFA and CoMSIA models do not consider ADME (e.g membrane permeability) properties of the compounds. This can be overcome by using descriptors derived from

molecular interaction fields (MIF), such as Volsurf or GRID,<sup>88</sup> which are alignment-independent and thus allow the analysis of structurally diverse data series. In addition, GRID MIFs have been applied to many areas of drug discovery, including 3D-QSAR,<sup>89</sup> ADME profiling, pharmacokinetic modeling<sup>90</sup> and metabolism prediction.<sup>91</sup>

Boccard *et al.*,<sup>92</sup> used VolSurf descriptors based on molecular interaction fields (MIFs) related to hydrophobic interaction forces, polarizability and hydrogen-bonding capacity, to predict affinity variation of flavonoid derivatives<sup>93,94</sup> towards P-gp. Optimal shape of the ligands and hydrophobicity was identified as major physicochemical parameters responsible for the affinity of flavonoid derivatives for P-gp. Interestingly, hydrogen-bonding capacities showed minor contribution towards activity. Moreover, authors also suggested the use of a 3D linear solvation energy approach<sup>95</sup> to predict pharmacokinetic properties such as permeability of the molecules in the membrane or blood- brain barrier penetration.

Cianchetta *et al.*,<sup>88</sup> used GRIND and simple physicochemical-based descriptors to identify pharmacophoric as well as physicochemical properties of substrates/inhibitors of P-gp. The GRIND model consists of two hydrophobic groups 16.5 Å apart from each other, two hydrogen-bond-acceptors at a distance of 11.5 Å, as well as the dimension of the molecule. The GRIND model has been shown to be more robust than the one obtained from physicochemical-based descriptors. Later on a similar approach was used to discriminate the structural features of P-gp substrates from non-substrates.<sup>96</sup> The model interpretation was in good agreement with molecular features proposed by Cianchetta and co-workers. However, the two models can't be compared directly as Almond descriptors are highly conformational dependent and both studies used different data sets of compounds. Nevertheless, a key recognition element, two hydrogen-bond acceptors groups around 11.5-15 Å apart from each other, is present in P-gp substrates in both studies. The model was able to discriminate substrates from non substrates with 82% accuracy and can be used as a virtual screening tool in early discovery programs.<sup>96</sup>

Recently, Broccatelli *et al.*,<sup>15</sup> studied P-gp with respect to its target as well as antitarget properties. VolSurf molecular descriptors were optimized to model pharmacokinetic properties along with a pharmacophoric method (FLAP), by using a structurally diverse data set of 1275 compounds. With VolSurf descriptors, a number of important pharmacokinetic parameters such as membrane permeability (size,

hydrophobic surface area, flexibility and logP) of substrates and inhibitors of P-gp was elucidated, while the pharmacophore-based method of fingerprints for ligands and proteins (FLAP) identified the most important pharmacophoric features around the optimal molecular shape (one H-bond acceptor and two large hydrophobic regions). Finally, the molecules in a conformation that sufficiently fits the optimal shape were classified as P-gp inhibitors. The model was used to predict P-gp inhibition with 88% accuracy and can be used to evaluate ADME-Tox properties and as a guide to design new drug candidates.

Apart from Volsurf and GRIND methods, Similarity-Based descriptors (SIBAR) as well as electrotopological state descriptors can be used to predict (ADME) properties. The SIBAR approach is based on selection of a highly diverse reference compound set and calculation of similarity values to these reference compounds. This approach was used to predict P-gp inhibitory activity of a series of propafenone analogs.<sup>60</sup> Zdrzil *et al.*,<sup>97</sup> focused on improvements regarding the reference set and concluded that a combination of high diversity and an interaction of the reference compounds with the biological target are beneficial for good models. Electrotopological state (ES) descriptors revealed that the ability to penetrate into membranes, molecular size, atom counts and electrotopological values of certain isolated and bonded hydrides are important structural attributes of substrates of P-gp. The “Gombar-Polli Molecular E-state (MoLES) Rule” states that molecules with MoLES value greater than 110 were identified as substrates and those with MoLES less than 48 as nonsubstrates.<sup>98</sup> Didziapetris and co-workers postulated the “rule of four”, which states that compounds with the number of hydrogen bond acceptors less than 8, a molecular weight less than 400 and having an acidic  $pK_a > 4$  are likely to be P-gp substrates, whereas compounds having less than 4 H-bond acceptors, a molecular weight less than 400 and a basic  $pK_a < 8$  are more likely to be non substrates.<sup>99</sup> This is in line with the findings of Gleeson *et al.*,<sup>100</sup> that neutral or basic molecules showing a molecular weight greater than 400 and logP value greater than 4 are more likely to be transported by P-gp rather than acidic or zwitterionic compounds. The topological substructural molecular design approach (TOPS-MODE), which is based on the calculation of spectral moments of molecular bond matrices, has also been shown to give accurate prediction of P-gp activity and helps in identification and quantification of fragment contributions that are responsible for P-gp activity for any molecular structure.<sup>101</sup>

In recent years, nonlinear methods such as artificial neural networks and support vector machines have been successfully applied for prediction of polyspecific drug-protein interactions. Kaiser *et al.*,<sup>102</sup> used self-organizing maps to separate high- and low-active propafenone-type inhibitors of P-gp. Compounds in the near neighborhood of highly active compounds in the training set were considered as potential hits. They showed structural scaffolds differing from those used in the training set, which indicates that this method is suitable for identification of structurally unrelated, diverse hits and thus represents a versatile tool for scaffold hopping. Huang *et al.*,<sup>103</sup> used Support Vector Machine (SVM) classification with a particle swarm-optimization algorithm as feature selection pre-processing step, which gave a model with an accuracy of greater than 90 %. This model was based on seven non-correlated and simple descriptors including three constitutional descriptors (molecular weight, number of H atoms and number of O atoms), two functional group counts (number of ring/tertiary C, number of substituted benzene) and two molecular properties (TPSA and logP). This further strengthens the important role of molecular weight, H-bonds, and the polar surface area for substrate recognition of P-gp.<sup>67,99</sup> Some other features such as number of ring tertiary C-atoms and the numbers of substituted benzene C atoms were also identified as characteristics of P-gp substrates.

Finally, Demel *et al.*,<sup>104</sup> classified P-gp substrates and nonsubstrates by using Friedman's RuleFit algorithm.<sup>105</sup> Models derived were based on simple, physicochemical descriptors. It has been suggested that substrates are associated with large hydrophobic surfaces, expressed as PEOE\_VSA\_HYD, are quite flexible and show less than four H-bond donors. Nonsubstrates are mainly characterized by small vsa\_acc and logP(o/w) values and are enriched in H-bond donors. The method allowed accurate and fast classification of P-gp substrates for a highly diverse set of compounds.

To summarize this part, ligand based studies on P-gp ligands, including QSAR, pharmacophore modeling and classification methods, shared some descriptors such as hydrophobicity, steric and electrostatic interactions, size or shape and H-bond properties. Although this information is still not giving a global picture, it may help medicinal chemists in the design and synthesis of novel compounds with more inhibitory potencies and better ADME properties.

## 3.2 Structure Based Approaches

Structure-based design mainly relies on the availability of crystal structures of the target protein at atomic level. P-glycoprotein and its homologues being membrane embedded proteins are mostly resistant to forming diffracting crystals which complicates the crystallization process of such proteins. Therefore, protein homology modeling based on templates of bacterial homologues or mouse P-gp, representing different catalytic states, is the method of choice for structure-based studies. Many structure based studies have been reviewed in detail by several authors in the past.<sup>55,106-109</sup> Table1 gives an overview of currently available templates and homology models of P-glycoprotein.

**Table 1.** Currently available homology models of P-glycoprotein, taken and modified from Klepsch *et al.*<sup>107</sup>

| Templates (PDB Codes) | Organism              | Ligand                  | Resolution [Å] | Similarity (%) <sup>a</sup> | Ref.    | Homology Models |
|-----------------------|-----------------------|-------------------------|----------------|-----------------------------|---------|-----------------|
| MsbA (1JSQ)           | <i>E. coli</i>        | Apo-open <sup>b</sup>   | 4.50           | 36 / 57                     | 110 (R) | 111-113         |
| MsbA (1PF4)           | <i>V. cholerae</i>    | Apo-closed <sup>c</sup> | 3.80           | 33 / 55                     | 114 (R) | 27              |
| MsbA (1Z2R)           | <i>S. typhimurium</i> | ADP·V <sub>i</sub>      | 4.20           | 37 / 57                     | 115 (R) |                 |
| Sav1866 (2HYD)        | <i>S. aureus</i>      | ADP                     | 3.00           | 34 / 52                     | 48      | 116-122         |
| MsbA (3B5W)           | <i>E. coli</i>        | Apo-open                | 5.30           | 36 / 57                     | 123     | 118,124         |
| MsbA (3B5X)           | <i>V. cholerae</i>    | Apo-closed              | 5.50           | 33 / 55                     | 123     | 118,124         |
| MsbA (3B5Y)           | <i>S. typhimurium</i> | AMP-PNP                 | 4.50           | 37 / 57                     | 123     |                 |
| MsbA (3B5Z)           | <i>S. typhimurium</i> | ADP·V <sub>i</sub>      | 4.20           | 37 / 57                     | 123     | 117             |
| MsbA (3B60)           | <i>S. typhimurium</i> | AMP-PNP                 | 3.70           | 37 / 57                     | 123     | 118             |
| MalK (1Q1B)           | <i>E. coli</i>        | Apo-semi open           | 2.80           | 31 / 50                     | 117     | 117             |
| MalK (1Q1E)           | <i>E. coli</i>        | Apo-open                | 2.90           | 31 / 50                     | 117     | 117             |
| ABCB1 (3G5U)          | <i>M. musculus</i>    | Apo-closed              | 3.80           | 87 / 93                     | 18      | 30,122          |
| ABCB1 (3G60)          | <i>M. musculus</i>    | QZ59-RRR                | 4.40           | 87 / 93                     | 18      |                 |
| ABCB1(3G61)           | <i>M. musculus</i>    | QZ59-SSS                | 4.35           | 87 / 93                     | 18      | 125             |

<sup>a</sup> Sequence identity/ homology with human P-gp, <sup>b</sup>Apo-open represents the nucleotide-free protein with the NBDs far apart, <sup>c</sup>Apo-closed describes the nucleotide-free protein with NBDs that lie close together, (R) Retracted models.

The first complete ABC crystal structure was published by Chang and co-workers in 2001, namely the MsbA lipid A half-transporter from *Escherichia coli* in the nucleotide free state. This was followed by two further MsbA structures reported by the same group in 2003 and 2005, in nucleotide free (NBDs lie close to each other) and ADP

bound state, respectively. Later on these structures were retracted because of some discrepancies between the MsbA structures and other structural and biochemical data on complete ABC transporters and isolated dimeric NBDs,<sup>37,126</sup> which, according to Chang and co-workers, was due to errors in crystallographic data-processing. In this context it has to be noted that several authors<sup>111-113, 27</sup> used wrong structures for generation of protein homology models that also fulfill a significant amount of biochemical data.<sup>111,113</sup> This is because in case of highly promiscuous membrane transporters cysteine cross linking studies and ligand photoaffinity labeling could be interpreted in several ways and thus might lead to quite convincing hypotheses even when based on partially wrong assumptions on the structure of a protein. Therefore, the results of these studies have to be carefully reconsidered, as some of them simply might be wrong.

In 2006, a high resolution (3 Å) X-ray structure of the *Staphylococcus aureus* transporter (PDB ID: SAV1866) in the ADP bound “outward-facing” state was published by Dewson *et al.*,<sup>48</sup> which served as a template for most of the homology models.<sup>116-122</sup> These models were found to be more consistent with the structural restraints obtained by cross-linking<sup>127,128</sup> and electron microscopic studies.<sup>129</sup> Later on in 2007, Chang and co-workers<sup>123</sup> revised their previously published crystal structures of the bacterial transporter MsbA and reported X-ray structures of MsbA in nucleotide free (PDB code: 3B5W, *E. coli*, resolution: 5.30 Å ; PDB code: 3B5X, *V. cholerae*, resolution: 5.50 Å) as well as “outward facing” ADP or AMP- PNP bound (PDB code: 3B5Z, *S. typhimurium*, resolution: 4.20 Å; PDB code: 3B60, *S. typhimurium*, resolution: 3.70 Å) structures. All four X-ray structures of MsbA represent different catalytic states of the transport cycle and are in agreement with the SAV1866 architecture. Becker *et al.*,<sup>115</sup> reported four homology models of different catalytic states of P-gp by using 2HYD and 3B60 as templates for the nucleotide bound state, and 3B5W and 3B5X as templates for the nucleotide free state. The measured interresidue distances in all four models correlate well with distances derived from cross-linking data.<sup>130</sup> Although the resolution of these MsbA X-ray structures are rather low and are thus insufficient for a detailed investigation of drug-transporter interactions, they provided important insights in our understanding of the complete structural picture of P-gp at different stages of the catalytic cycle. These conformational changes in the MsbA structures are further supported by structural studies of the ABC transporter MalK<sup>131</sup> and by the domain swapping topology suggested by the Sav1866 structures.<sup>48</sup> Further, O'Mara *et al.*,<sup>117</sup>



created homology models of different catalytic states of P-gp representing, semi-open, open and ADP bound states by using MalK (PDB ID: 1Q1B), (PDB ID: 1Q1E) and Sav1866 (PDB ID: 2HYD), respectively, as templates. As P-gp must go through the ADP-bound state to reset the NBDs for the next catalytic cycle, the flexibility of the ADP-bound states in MalK<sup>131</sup> suggests that the models may represent three stages of the catalytic cycle, an ATP bound closed state, an ADP-bound (nucleotide free) semi-open state and an open state (nucleotide free) conformation.

In the absence of a high-resolution crystal structure of human P-gp, since 2009 most of the homology models are based on Sav1866 and MsbA<sup>116-121</sup> and are consistent with cross linking and electron microscopic data, showing close association of TM segments 5 with 8<sup>127</sup>, 2 with 11<sup>128</sup> and 1 with 11.<sup>132</sup> Amino acid residues predicted to line the drug-translocation pathway were also consistent with cystein scanning and mutagenesis data.<sup>25,28,133,134</sup> However, most of these homology models represent P-gp in the closed conformation as the NBDs are close to each other and the predicted drug binding cavity is open to the outside of the cell.

In 2009 the first X-ray structure of a eukaryotic ABC efflux pump, P-glycoprotein (PDB code: 3G5U, *M. musculus*, resolution: 3.8 Å), was published by Aller *et al.*<sup>18</sup> Additionally the structure was published together with two co-crystallised enantiomeric cyclic peptide inhibitors (CPPIs; QZ59-RRR/QZ59-SSS, resolution: 4.40 Å /4.35 Å). This new information sheds light on possible ligand binding areas as well as on stereoselectivity of P-gp. Stereoselectivity has been also observed recently for a series of benzopyrano[3,4-b][1,4]oxazines, as well as for flupenthixol.<sup>135</sup> The first crystal structure of mouse P-gp represents a huge step forward for structure-based studies on this transporter, but it is still difficult to determine the exact orientation of many side chain residues at this (3.8 Å) resolution. However, having 87 % sequence identity to human P-gp it may serve as a good template for homology modeling.

Pajeva and co-workers built a homology model of human P-gp by using the X-ray structure of mouse P-gp as template (PDB ID: 3G61, resolution 4.35 Å) and docked quinazolinones into the binding region, which was defined by extending 14 Å around the position of the co-crystallized ligands. The ligand-protein interaction profile of quinazolinones suggested interaction with TM5 (Tyr307), TM6 (Phe336) and TM11 (Tyr953, Phe957), which was further validated and confirmed by models based on pharmacophoric features.<sup>136</sup> Pajeva *et al.*<sup>125</sup> in another study compared the residues

exposed to the binding cavity of the “inward-facing” homology model of human P-gp (3G61) with that of the “outward-facing” homology model based on the Sav1866 structure.<sup>121</sup> It has been elucidated that the ligands remain bound to the same residues during the transition from the inward- to the outward-facing conformation of the protein. Further analysis of docking poses of cyclic peptides QZ59-RRR and QZ59-SSS<sup>125</sup> confirms the X-ray data about the functional role of TM4, TM6, TM10 and TM12 for the entrance gates (portals) to the cavity.<sup>18</sup> This is in agreement with recent findings of Klepsch and co-workers about ligand-protein interaction profiles of propafenones in two different catalytic states of P-gp.<sup>122</sup> She extensively docked some selected propafenones into a homology model, based on 3G5U (mouse P-gp without QZ59 isomer) as well as in the nucleotide-bound conformation 2HYD based on the Sav1866 structure. Transmembrane helices 5, 6, 7 and 8 showed interaction with propafenones in most of the clustered poses in both models. Moreover, amino acid residue Tyr307 has been identified to play a crucial role in H-bond interaction: Most of the homology models and docking studies published so far are in agreement with experimental studies therefore, information from both structure based as well as ligand based approaches could pave the way for a deeper understanding of the molecular basis of ligand/transporter interaction.

#### 4. Aim of the Work

The general objective of this thesis is to get detailed knowledge about the molecular basis of ligand/P-gp interaction by using both ligand and structure based *in silico* model: The specific aims are:

- To establish predictive models for P-gp inhibitors, such as 2D- and 3D- QSAR models, using simple physicochemical as well as GRIND molecular descriptors.
- To explore the stereoselectivity of P-gp by docking of enantiomerically pure benzopyrano[3,4-b][1,4]oxazines into the apo state homology model of human P-gp.
- To identify the most promising P-gp inhibitors out of our compound library by using hit to lead tools such as ligand efficiency (LE)<sup>137,138</sup> and lipophilic efficiency (LipE).<sup>138</sup>
- To explore the ligand-protein interaction pattern of P-gp inhibitors by docking of those ligands showing the best activity/logP and size ratio into an open state (nucleotide free) homology model of P-gp.

- To explore the lipophilic efficiency (LipE) distribution profiles for three targets showing fundamental differences in the way how the ligand enters the binding site: P-gp (via the membrane bilayer), serotonin transporter (SERT; from outside the cell) and the hERG (human *Ether-à-go-go* Related Gene) potassium channel (from inside the cell).

## 5. 2D-QSAR Models

Within this thesis we established 2D-QSAR models for two different data series of inhibitors of P-gp, including a series of chalcones and conformationally rigid diastereoisomers of benzopyrano-[3,4-b][1,4]-oxazines. Complete details on the 2D-QSAR models of these two series are provided in chapter 2 and 3 of this thesis (submitted manuscripts). The following section contains a brief summary of the results obtained.

In order to determine the influence of physicochemical properties of the compounds on their biological activity, a pool of molecular descriptors consisting of those supplied by the program MOE<sup>139</sup> version 2009-10 (atom and bond counts, connectivity indices, partial charge descriptors, pharmacophore feature descriptors, calculated physical property descriptors) were computed for Hansch analyses.<sup>140</sup> QSAR-Contingency,<sup>141</sup> a statistical application in MOE, was used for selection of those descriptors which best describe the molecules in the training set. PLS analysis was performed to determine the relationship between 2D molecular descriptors and biological activity of the compounds. The predictive ability of the models were determined by leave one out cross validation (LOO), as well as by an external test set. Our QSAR models for two data series end up with the following equations,

$$\text{Log1/IC}_{50} = 0.02 (\text{vsa\_hyd}) - 0.36 (\text{b\_rotN}) + 0.93 \quad \text{Chalcone derivatives}$$

$$\text{R}^2 = 0.79; \text{q}^2 (\text{LOO}) = 0.71; \text{RMSE} = 0.51; \text{n} = 22 \quad (\text{Equ.1})$$

$$\text{Log1/IC}_{50} = 0.01 (\text{vsa\_hyd}) - 4.74 \quad \text{Benzopyrans}$$

$$\text{R}^2 = 0.67; \text{q}^2 = 0.63; \text{RMSE} = 0.48; \text{n} = 35 \quad (\text{Equ.2})$$

Vsa\_hyd, which describes the sum of VDW surface areas of hydrophobic atoms ( $\text{\AA}^2$ ), is identified as the most contributing descriptor towards biological activity within both series of P-gp inhibitors. This is perfectly in line with previous studies which showed that distribution of hydrophobicity within the molecules influences their mode of interaction with P-gp,<sup>65</sup> and that lipophilicity needs to be considered as a space

directed property.<sup>63,64</sup> This is also in line with a recent analysis of the binding area, which shows a large internal cavity in the lipid bilayer that allows the access for entry of hydrophobic compounds via the protein/membrane interface.<sup>18</sup> In addition, overall lipophilicity (logP) of the compounds (i.e. to enrich in biological membranes) plays an important role, as stressed out in numerous publications.<sup>57,61,62</sup> A separate logP (o/w) analysis of the two diastereoisomeric series reveals a positive correlation towards biological activity. However, diastereoisomers of benzopyrano[3,4-b][1,4]oxazines having 4aS,10bR-configuration showed a better correlation ( $R^2 = 0.60$ ) as compared to the ones having 4aR,10bS-configuration ( $R^2 = 0.40$ ). This might be due to steric constraints caused by a benzyl moiety in a compound having 2S,4aR,10bS-configuration and thus strengthens different binding modes for these two types of diastereoisomers.<sup>142</sup> Interestingly, in contrast to several other compound classes, a poor correlation has been observed between overall lipophilicity of the chalcone derivatives and their P-gp inhibitory activity ( $r^2 = 0.18$ ). This indicates that the variance in the biological activity of chalcone derivatives is mainly driven by the concrete pattern of hydrophobicity distribution within the molecules

The QSAR studies for a set of chalcone derivatives demonstrate that hydrophobic distribution along with number of rotatable bonds in the molecule influence the potency of the compounds. This confirms the finding of Wang *et al*, on the contribution of the hydrophobic distribution within the molecules along with molecular weight and number of rotatable bonds for P-gp inhibitory potency.<sup>67</sup> 2D-QSAR models showed good predictive power, however in order to get an insight about 3D structural requirements of P-gp inhibitors, we computed several 3D-QSAR models using MIF based descriptors.

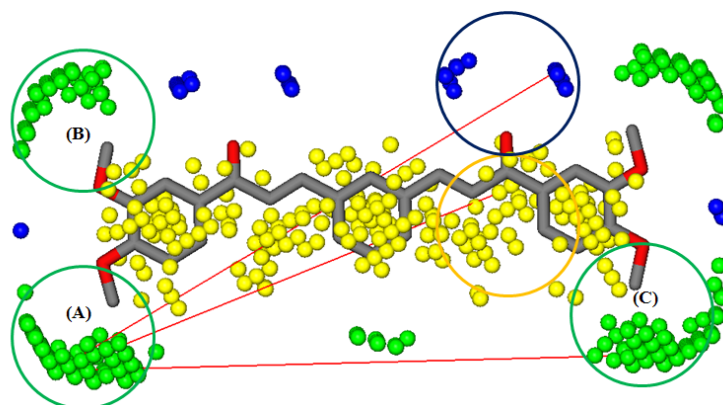
## 6. 3D-QSAR Models

The computational tool Pentacle version 1.06<sup>143</sup> was used for computing alignment-free molecular descriptors or GRID-independent molecular descriptors (GRIND)<sup>144</sup> using different compound series active as inhibitors of P-gp. Within this thesis we explore the capability of the GRIND approach to derive predictive 3D-QSAR models for different data sets of inhibitors separately and then also to combine the sets and to create one general model (Chapter 2-4). Our 3D-QSAR models using GRIND descriptors identified two hydrogen bond acceptors, one hydrogen bond donor,

hydrophobicity and shape of the ligand as most important common features for high biological activity of P-gp inhibitors.

Favorable interacting regions of two H-bond acceptor groups 8.80-9.20 Å apart from each other have been identified as being highly beneficial for high P-gp inhibitory activity in local models (Chapter 2,3). Interestingly, the same distance between two hydrogen bond acceptors has been identified in a GRIND model containing an extended training set of 292 compounds of different chemical scaffolds ( $q^2 = 0.61$ ) (Chapter 4). However, this distance range is not fully consistent in the combined model and could not separate completely the highly active ( $IC_{50} < 1\mu M$ ) compounds from low active ( $IC_{50} > 1\mu M$ ) ones. This might reflect the highly promiscuous binding site of P-gp, which possesses multiple spots able to participate in hydrophobic and H-bond interactions. Thus, different chemical series most probably utilize different H-bond interaction patterns.

We identified three important boundaries (A, B and C, Figure 5) of inhibitors of P-gp. Distances of favorable interacting regions including one hydrogen bond acceptor, one hydrogen bond donor and one large hydrophobic group from different edges of the molecules have been measured and compared in all series of P-gp inhibitors separately as well as in a combine model containing structurally diverse compounds (Chapter 4). Highly active benzopyrano[3,4-b][1,4]oxazines and chalcone derivatives showed a hydrophobic moiety at a distance of 15.20-15.60 Å or 17.60-18.00 Å, respectively, apart from one edge of the molecule. Interestingly, the same pharmacophores have been identified separated by a distance of 16.00-16.80 Å in our GRIND model containing diverse data series. Thus, most important pharmacophoric features, their mutual distances and distances from different edges of the molecules are comparable in individual models as well as in one combined model. This indicates the demand of a specific shape as well as a particular pharmacophoric pattern for this chemotype for P-gp inhibitors. Additionally it points out the usefulness of the GRIND approach for deriving predictive models across diverse chemical scaffolds.



**Figure 5.** Important pharmacophoric features and their mutual distances for high biological activity of P-gp inhibitors (as proposed by our 3D-QSAR models). A, B and C represent three important shape probes and their mutual distances, yellow and blue probes indicate favorable hydrophobic and hydrogen bond acceptor areas and their distance from different edges of the molecules.

## 7. P-gp and Stereoselectivity

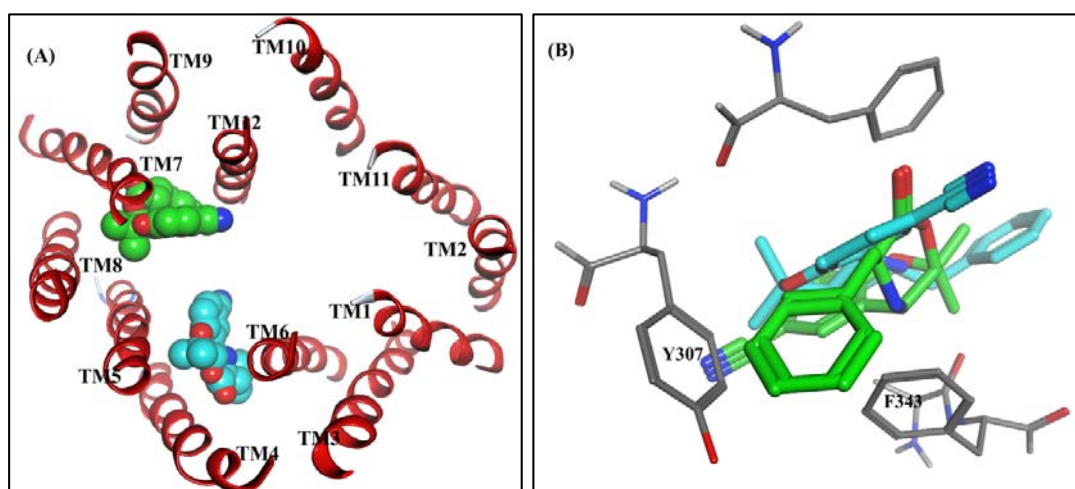
Lack of significant stereoselectivity in drug/P-gp interaction was observed for P-gp substrates/inhibitors, such as verapamil, nifedipine, nitrendipine, felodipine, carvedilol, propranolol, zosuquidar and propafenone.<sup>145,146</sup> However, there are a few reports of remarkable stereospecificity.<sup>135,147</sup> Furthermore, the recently published crystal structure of mouse P-gp co-crystallised with the two enantiomeric cyclopeptides QZ59-RRR and QZ59-SSS revealed distinct binding sites for the two enantiomers.<sup>18</sup> In contrast to our main lead compound propafenone, the dihydrobenzopyranes offer the advantage of remarkably reduced conformational flexibility and thus might be versatile molecular tools for probing stereoselective differences of drug/P-gp interaction (Table. 2). Especially annelation of a third ring leading to benzopyrano[3,4-b][1,4]oxazines and introduction of large substituents at position 2 of the tricyclic system should lead to compounds with pronounced configurational differences. Complete details about stereoselective interactions of benzopyrano[3,4-b][1,4]oxazines are described in chapter 5.<sup>142</sup>

**Table 2.** Chemical structure and biological activity of enantiomerically pure benzopyrano[3,4-b][1,4]oxazines.

| #   | R  | Log P | IC <sub>50</sub> (μM) |
|-----|--|-------|-----------------------|
| 5a  | CH <sub>3</sub>                                  | 2.84  | 29.85                 |
| 5b  | CH <sub>3</sub>                                  | 2.84  | 14.55                 |
| 6a  | CH(CH <sub>3</sub> ) <sub>2</sub>                | 3.82  | 2.40                  |
| 6b  | CH(CH <sub>3</sub> ) <sub>2</sub>                | 3.82  | 2.70                  |
| 7a  | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 4.38  | 0.55                  |
| 7b  | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 4.38  | 0.77                  |
| 11a | CH <sub>3</sub>                                  | 1.98  | 1241.65               |
| 11b | CH <sub>3</sub>                                  | 1.98  | 76.89                 |
| 12a | CH(CH <sub>3</sub> ) <sub>2</sub>                | 2.95  | 15.32                 |
| 12b | CH(CH <sub>3</sub> ) <sub>2</sub>                | 2.95  | 59.33                 |
| 13a | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 3.51  | 2.69                  |
| 13b | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 3.51  | 259.78                |

Complete details about structure activity relationship (SAR) as well as correlation to logP (o/w) are provided in chapter 3 of this thesis. A remarkable difference in biological activity of two types of diastereoisomers is might be due to their different binding modes at P-gp. This is further supported by results of docking studies performed on a homology model of human P-gp based on the X-ray structure of mouse P-gp (PDB ID: 3G5U). Agglomerative Hierarchical Cluster analysis of the docking poses based on consensus RMSD of their common scaffold identifies mainly interactions of **5a,b-7a,b** with amino acid residues of TM5 and TM6, including Y307, Y310, F343, F336 and Q347. For tricyclic diastereoisomers **11a,b-13a,b** two types of clusters have been identified (Figure 6 A, B). Clusters of type one containing only compounds with (4aS,10bR)-configuration (**11a-13a**) are located close to the potential entry pathway consisting of TM 4, 5, and 6, interacting in particular with amino acid residues Y307, F343, A342, and F303. The second type of clusters contain all compounds with (4aR,10bS)-configuration (**11b-13b**). These (second type) clusters are located at two different positions. One position is identical with those of **11a-13a**, the second position is located close to TM 7, 8, 9 and 12 (Figure 6 B), surrounded by amino acid residues A985, I765 and L724. Comparing the main positioning of the benzopyrano[3,4-b][1,4]oxazines with those of QZ59 some overlap could be observed. Especially interaction with Y307, F343, F336, A985, A342, M69 and F728 was observed for all

ligands. A closer look of ligand–protein interaction profiles of compounds **13a,b** and **7a,b** identified some docking poses of **13b** showing a steric constraint of the benzyl moiety of **13b**, which is about 2 Å apart from Y307 and about 2.5 Å apart from F343. All these poses are located at the entry gate (Figure 6 C). No such steric constraint has been observed for **13a** or for **7a,b**. In the case of **7b** this is most probably due to its conformational flexibility, which allows adopting a conformation to minimize the steric interactions. This indicates that the differences observed for the biological activities of phenylalanine derivatives **13a** and **13b** might be due to steric constraints at the entry path rather than differences in drug/transporter binding. Of course, at the current stage this has to be taken very cautiously, as P-gp undergoes major conformational changes during the transport cycle and docking experiments represent only a single snapshot of this complex movement. However, our ligand docking studies into a homology model of P-gp could provide first evidence for different binding areas of the two diastereomeric compound series.



**Figure 6.** (A) Docking poses of **13a** (blue) and **13b** (green) viewed from outside into the TM region. (B) Steric constraints of **13b** (green) with amino acids residues Y307 and F343 near the entry gate which is supposed to be the preferable interaction position of **13a** (blue).

## 8. Ligand Efficiency/Lipophilic Efficiency Models

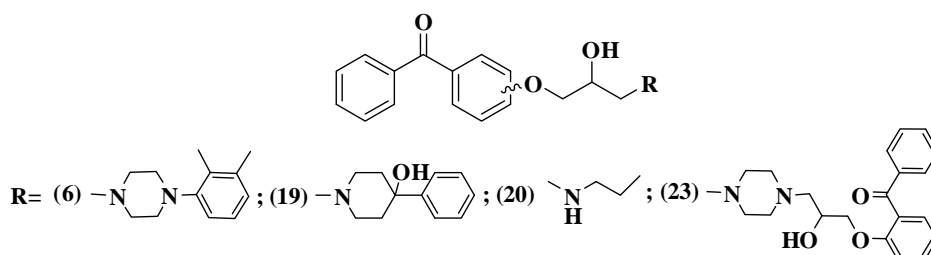
Concept of “*Binding energy of the ligand per atom*” or ligand efficiency (LE)<sup>137</sup> and lipophilic efficiency (LipE),<sup>148</sup> which combines both “*potency and lipophilicity*”, represent useful hit to lead tools to identify the derivatives with the best activity/logP (or size) ratio and provide insights for the design of new ligands.<sup>149,150</sup> Complete details about LE and LipE calculations are described in chapter 6 of this thesis.



**Ligand Efficiency (LE)** is a simple metric for assessing whether a ligand derives its potency from optimal fit with the target protein or simply by virtue of making many contacts.<sup>151</sup> Hopkins *et al.*,<sup>152</sup> in 2004 proposed a ligand efficiency of 0.29 kcal mol<sup>-1</sup> per non-hydrogen atom for a promising drug candidate possessing a potency of about 10 nM, which was widely accepted by several other authors later on.<sup>149,153,154</sup> We calculated ligand efficiency values for a dataset of inhibitors of P-gp, including benzophenones, some selected propafenones and 8 compounds in different stages of clinical investigation. A basic trend has been observed where ligand efficiencies drop dramatically as the size increases above 50, which has also frequently been observed in literature.<sup>155</sup> However, this dependency has a disadvantage when using the LE measure to guide the design of new compounds, as the size of the ligands is likely to increase during this process. Various schemes have been designed in the literature to solve this problem.<sup>156, 157</sup> LE values of the P-gp inhibitors and substrates have been used for subsequent scaling to get a size-independent ligand efficiency scale as described by Reynolds *et al.*<sup>155</sup> Finally, the ratio of ligand efficiency over normalized ligand efficiency scale gives a scoring function called “Fit Quality” (FQ), where more efficient binders in the data set were scaled to have a score of 1.0 across a wide range of molecular size. Implementing this to our data set, we could observe that most of the compounds in clinical investigation showed FQ score above 1 including zosuquidar, ONT093, elacridar, and tariquidar, along with some benzophenones and propafenone analogues. Although they showed low values for the size dependent LE, in the normalized LE fit quality scale they are considered to be more efficient ligands covering a wide range of molecular size. Subsequently, the same data set was normalized for their lipophilicity, which may provide some guidance towards promising drug candidates in the future.

**Lipophilic Efficiency (LipE)**, is a parameter that combines both potency and lipophilicity, and has been introduced for the first time in 2007 by Leeson *et al.*<sup>138</sup> LipE is defined as a measure of how efficiently a ligand exploits its lipophilicity to bind to the target. clogP values of benzophenones, selected propafenones and inhibitors/substrates of P-gp in clinical investigation varies from 15.09 to 2.66 with lipophilic efficiency covering a range between -8.79 to + 3.08. Ligand lipophilic efficiency values greater than 5, lipophilicity of about 2.5 and activity of less than 10 nM have been reported as standard thresholds for an oral drug.<sup>138,148</sup> Interestingly, only

4-hydroxy-4-phenyl-piperidine analogous propafenone GPV0062 as well as the dimer **23** exhibit values slightly higher than 3, while the rest of the compounds exhibit LipE values below 3. Although P-gp inhibitors are highly lipophilic, they showed LipE values below the standard threshold.<sup>138,148</sup> This might be due to the fact that the access path of substrates/inhibitors of P-gp is most likely via the membrane bilayer. This is additionally supported by the recent X-ray structure of mouse P-gp, which shows a large inner cavity accessible from the membrane via putative entry ports composed of transmembrane helices 4/6 and 10/11.<sup>18</sup> Thus, for P-gp and other ABC transporter the thresholds should be reconsidered and adjusted to this target class. Nevertheless, from the benzophenone data set of P-gp inhibitors, compounds **15**, **16**, **19**, **20**, and **23**, might be the most promising ones (Chapter 6) as their LipE values are between 2 and 3, a range where most of the compounds which entered clinical trials, are located. To get insights into the potential binding mode of the most promising compounds, we selected compounds **19**, **20**, and **23**, which are ranked high both in LipE and FQ scores and **6** (Figure 7), as it is top ranked with respect to FQ for further structure based studies.

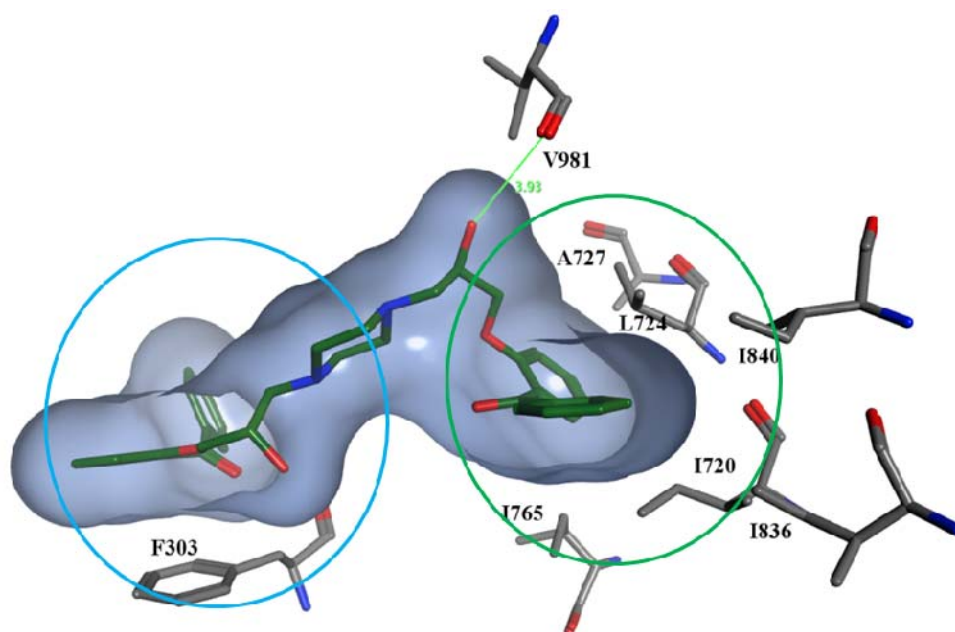


**Figure 7.** Benzophenone analogs which showed best activity/logP (or size) ratio in their LE, LipE profiles.

### 9. Interaction Pattern of P-gp Inhibitors in the Binding Pocket

Compound **6**, **19**, **20** and **23** were docked in their neutral form into an open state homology model of human P-gp<sup>122</sup> based on the X-ray structure of mouse P-gp (PDB ID: 3G5U)<sup>18</sup> by using software package GOLD. Complete details about docking protocol are provided in methods section of chapter 6. Ligand protein interaction pattern of selected poses of compounds **6**, **19**, **20** and **23** further strengthen our structure-activity relationship studies as well as previous docking studies.<sup>122,142</sup> The benzophenone scaffold interacts with F343 and F303 near the entry gate, whereas the lipophilic substituents in the vicinity of the basic nitrogen atom are surrounded by hydrophobic amino acid residues L724, I720, V981, I840, I836 and I765 located at TM 7, 9, and 12 (Figure 8). This further supports the importance of high lipophilicity and

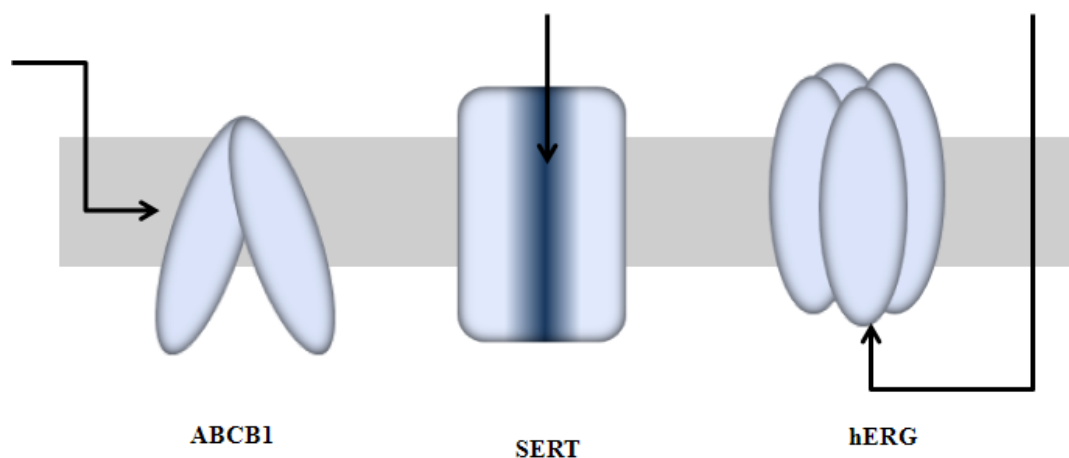
also is in line with previous studies performed by Pajeva and Wiese.<sup>63</sup> The top ranked cluster also support our previously proposed binding positions for benzopyrano[3,4-*b*][1,4]oxazines, where compounds having 4aS,10bR configuration interact mainly with amino acid residues of TM4, 5 and 6 near the entry gate, while compounds having 4aR,10bS configuration go deeper inside the binding cavity and are mainly surrounded by hydrophobic amino acid residues of TM7, 8, 9 and 12.<sup>142</sup> Interestingly, the top scored cluster for dimer **23** is positioned in a way to bridge these two positions (Figure 8). Selected benzophenone analogs have been previously used as photo-affinity ligands to characterize the drug-binding domain of propafenone-type analogs. In these studies, TM 3, 5, 6, 8, 10, 11, and 12 were identified as potential interacting helices.<sup>26,27,158,159</sup> This is well in line with our docking studies, which show main interactions with TM 5, 6 near the entry gate and TM 7, 8 and 12 deeper inside the cavity thus, making 5/8 interface. No significant cluster of poses has been identified on the second wing (2/11 interface), which might be due to the asymmetry in the homology model of P-gp, thus narrowing the available space at this side.



**Figure 8.** Ligand-protein interaction profile of best scored pose of benzophenone dimer **23** making a bridge between interaction positions of benzopyrano[3,4-*b*][1,4]oxazine having 4aS,10bR-configuration, represented by a blue circle, while green circle indicates the preferable interaction position of diastereoisomers with 4aR,10bS-configuration.

### 10. Ligand Efficiency (LE) and Lipophilic Efficiency (LipE) Distribution Profiles Along Different Entrance Pathways.

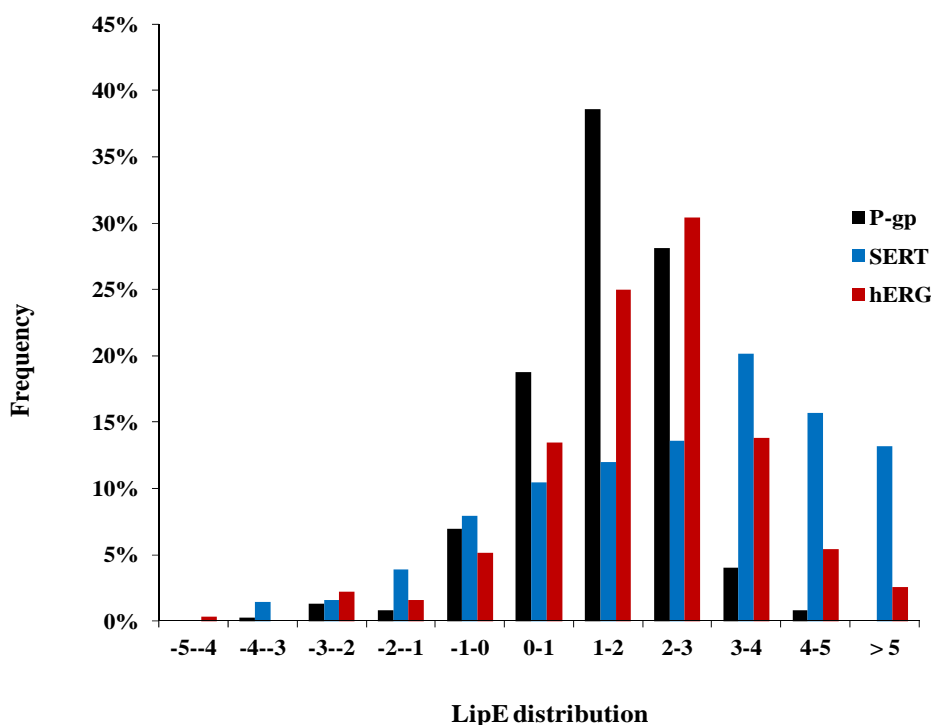
LipE values by definition quantify the extent to which ligands ‘prefer’ to bind to the protein or to be solvated in octanol. Inhibitors/substrates of P-gp are highly lipophilic, and are supposed to easily get access to the binding cavity which is directly exposed to the membrane bilayer. Therefore, substrates/inhibitors of P-gp are more likely solvated in octanol. This may provide one of the bases why LipE values of ligands of P-gp are below the standard threshold of 5 as described in detail in chapter 6 of this thesis. Therefore, we also study the distribution of LipE and LE profiles for compounds taking three different access pathways. (1) The ligand gets access to the binding pocket via the membrane bilayer (P-gp), (2) the ligand directly accesses to the binding chamber from the extracellular environment (SERT), (3) the ligand reaches the binding cavity via the cytoplasm (hERG) (Figure 9). LipE values of inhibitors of serotonin transporter (SERT), (ChEMBL data base),<sup>160</sup> hERG blockers<sup>161</sup> and propafenone derivatives of inhibitors of P-gp (in-house data) were calculated as described in chapter 6 of this thesis.



**Figure 9.** Schematic representation of access of ligands into the binding chambers of P-gp, SERT and hERG along three different translocation pathways, P-gp ligands approach the binding cavity via the membrane bilayer, however in SERT the ligands get direct access in to the binding chamber from the extracellular environment while in hERG the access route is via the cytoplasm.

The LipE distribution profile of SERT inhibitors identified about 13% compounds that cross the LipE threshold of 5 (Figure 10). These compounds cover a wide range of activity (0.01 nM- 10 mM) and clogP (-3.42 to 4.66) distribution. Moreover, 15 lead compounds for SERT inhibition have been identified with clogP ~2.5, LipE > 5 and

$IC_{50} < 10$  nM. However, none of them was listed as a marketed drug. These include ChEMBL compounds (IDs: 196468,<sup>162</sup> 190937, 1140708, 253184,<sup>163</sup> 1138354, 241646, 1138354, 138354,<sup>164</sup> 1147706, 1144290,<sup>165</sup> 1153074,<sup>166</sup> 394158,<sup>167</sup> 511500, 1152751<sup>168</sup> and 1152805.<sup>169</sup> In case of hERG only 2.5 % of the compounds could cross the LipE threshold  $> 5$ . They showed a potency distribution between 5-18000 nM and clogP values from -0.77 to 2.21. Moreover, only two compounds (Almokalant and Dofetilide) showed  $clogP \sim 2.5$ ,  $LipE > 5$  and potency values  $< 10$ nM. Dofetilide is a registered class III antiarrhythmic agent, while almokalant is in phase II clinical investigations.<sup>170,171</sup> LipE profiles of P-gp inhibitors could not identify any compound that reaches the standard threshold value of 5. Most of these ligands fall in the LipE range of 1-2 (39%) or 2-3 (28%) with a wide range in distribution of their clogP values (1.32 to 15.09) as well as  $IC_{50}$  (5.6 nM to 1.8 mM) values.



**Figure 10.** LipE distribution profiles of ligands of the hERG potassium channel, SERT and P-gp, representing targets with three different access pathways.

These observations are in line with our hypothesis about LipE distribution along different ligand access pathways. As the binding cavity of P-gp is directly exposed to the membrane bilayer, the inhibitors/substrates of P-gp are highly lipophilic and their LipE values although below threshold are considerable to get an access to the binding site. In case of hERG the ligands have to pass first through the membrane bilayer and

then get an access to the target from the cytoplasmatic side, therefore they need some specific range of lipophilicity ( $\text{clogP} \sim 2\text{-}3$ ) to cross the membrane bilayer as well as potency to preferably interact with the target. In this case the most efficient ligand will be the one which exploits its lipophilicity to get an access into the binding cavity. For this reason only a few hERG blockers could maintain this balance between inhibitory potency and lipophilicity and reach the standard LipE value. Finally, for monoamine transporters such as SERT, the binding cavity is directly exposed to the extracellular fluid, which allows ligands to enter in the binding chamber without being highly lipophilic. Therefore, the percentage of compounds that reached the standard threshold of LipE is greater in SERT (13%) as compared to hERG (2.5) and P-gp (0%).

The same data sets were subjected to ligand efficiency profiling. Maximum ligand efficiency values for most promising candidates have been observed for inhibitors of SERT ( $0.3\text{-}0.8 \text{ kcal mol}^{-1}$ ) followed by hERG blockers, which possess LE in the range of ( $0.3\text{-}0.65 \text{ kcal mol}^{-1}$ ). The most efficient inhibitors of P-gp exhibit an LE range of ( $0.3\text{-}0.4 \text{ kcal mol}^{-1}$ ) as shown in figure 11. This difference in the range of LE values might be due to a difference in number of non hydrogen atoms or due to a difference in potency of inhibitors in all three cases. The first one can be ruled out, as it has been observed that most of the ligands of all three targets exhibit similar average range of number of non hydrogen atoms (10-50).

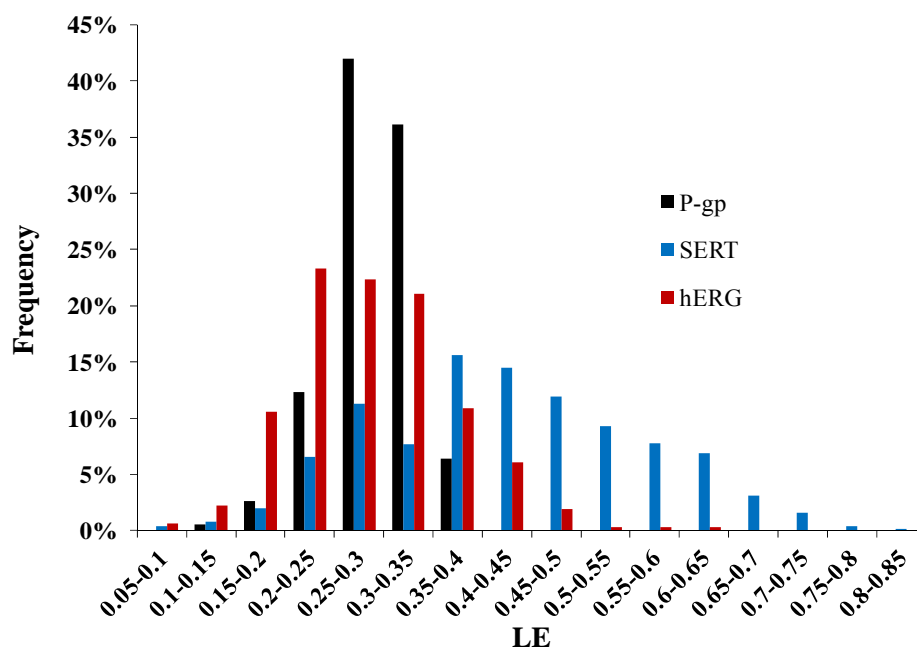
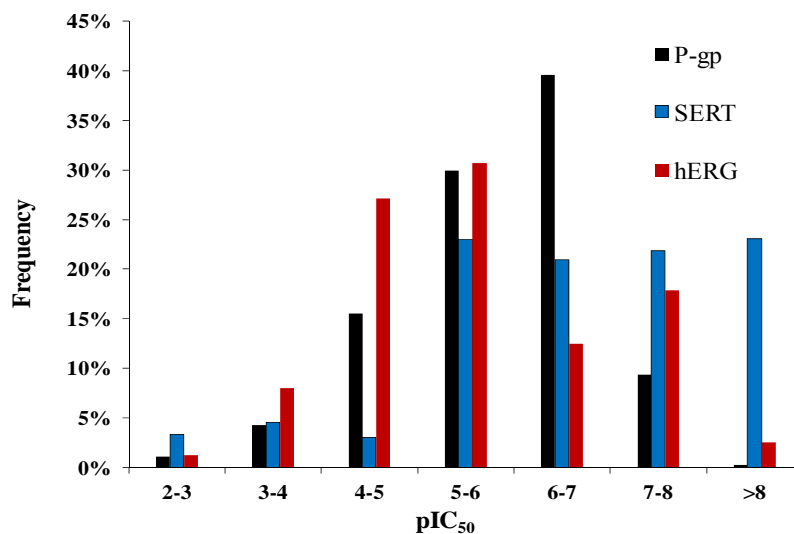


Figure 11. Ligand efficiency (LE) distribution profile of ligands of P-gp, SERT and hERG



**Figure 12.** pIC<sub>50</sub> distribution of ligands of P-gp, SERT and hERG.

Therefore the difference in LE values might be due to a difference in their pIC<sub>50</sub> values. The activity distribution profile of inhibitors of all three targets (Figure 12) showed a clear difference of compound frequencies in the high activity range (pIC<sub>50</sub>: 7-8) (SERT: 23%; hERG: 18%; P-gp: 10%). About 23% of the ligands of SERT were identified as highly potent (pIC<sub>50</sub> > 8), however 3% hERG blockers and only GPV576, which is a propafenone derivative of P-gp could reached the highest category of pIC<sub>50</sub> values (> 8). Remarkably, 41% of P-gp inhibitors were identified to belong to the low activity range (pIC<sub>50</sub>: 6-7). This further facilitates our understanding of the drug-protein interaction for these three targets and elucidates the promiscuity of P-gp as compared to SERT and also even to hERG. Due to the highly promiscuous nature of the binding pocket of P-gp, which possesses multiple spots able to participate in hydrophobic and H-bond interactions, its substrates and inhibitors do not get one particular optimal fit within the binding cavity as compared e.g. to SERT. Therefore, LE values of P-gp inhibitors, although in the range of the widely accepted threshold (LE > 0.3), are lower than LE values of SERT inhibitors and hERG blockers. There is definitely a need for more detailed studies on the ligand-protein interaction profile of inhibitors of P-gp which exhibit high LE and LipE values, as this might facilitate future efforts to design more potent inhibitors of P-gp.

## 11. Summary and Outlook

The primary aim of this thesis was to explore the molecular basis of the interaction of P-gp with small molecules by using both ligand and structure based *in silico* modelling techniques. Starting from ligand based approaches several 2D- and 3D-QSAR models were established to elucidate the interaction forces responsible for high affinity of small molecules. GRIND analysis revealed the importance of a particular shape of inhibitors of P-gp and provided preferred distances of important pharmacophoric features (such as hydrophobic and H-bond acceptors) from different edges of the molecules. Furthermore, in order to gain high activity, two H-bond acceptors at a distance of at least 8.80-9.20 Å should be present in the scaffold. This global GRIND model for P-gp inhibitors can be used for the generation of a web-based application for prediction of inhibition of the P-gp efflux pump.

Benzopyrano-[3,4-*b*][1,4]oxazines are versatile molecular tools to probe the stereoselectivity of P-glycoprotein. Ligand docking studies into a homology model of P-gp provide first evidence for different binding areas of the two types of diastereomeric pairs and thus help to explain a large difference in their potency to inhibit P-gp mediated drug efflux. Docking studies of a set of selected benzophenones provide evidence that the benzophenones seem to bridge the two distinct binding sites proposed for diastereoisomers of benzopyrano[3,4-*b*][1,4]oxazines. This further supports the general hypothesis of a huge binding zone with distinct, but overlapping binding sites for individual scaffolds as basis for the promiscuity of P-gp.

Although ligand efficiency (LE) and lipophilic efficiency (LipE) are routinely used in lead optimisation programs, up to now no reports on LE and LipE profiles of inhibitors and substrates of ABC transporters have been published. We thus analyzed the LE and LipE profiles of a series of benzophenone-type inhibitors of P-gp and compare them with P-gp inhibitors in clinical trials. Some of the benzophenones showed ligand efficiency and lipophilic efficiency behavior comparable with the compounds in different stages of clinical investigations. Interestingly, although P-gp inhibitors are highly lipophilic, they showed LipE values below the threshold considered to be necessary for promising drug candidates. This might be due to the unique entrance pathway directly from the membrane bilayer rather than from the intra- or extracellular compartment. All information from both structure based as well as ligand based approaches used in this study aid in the understanding of the molecular



basis of ligand/P-gp interaction and thus could pave the way for design of new lead compounds.

**Acknowledgements:** This work was supported by financial support of Austrian Science Fund (grant SFB F3502) and Higher Education Commission of Pakistan.

### References

1. Gottesman, M. M.; Fojo, T.; Bates, S. E. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* **2002**, *2*, 48-58.
2. Dano, K. Active outward transport of daunomycin in resistant Ehrlich ascites tumor cells. *Biochim Biophys Acta* **1973**, *323*, 466-83.
3. Juliano, R. L.; Ling, V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta* **1976**, *455*, 152-62.
4. Goda, K.; Bacso, Z.; Szabo, G. Multidrug resistance through the spectacle of P-glycoprotein. *Curr Cancer Drug Targets* **2009**, *9*, 281-97.
5. Keppler, D. Multidrug resistance proteins (MRPs, ABCCs): importance for pathophysiology and drug therapy. *Handb Exp Pharmacol* **2011**, 299-323.
6. Dean, M.; Fojo, T.; Bates, S. Tumour stem cells and drug resistance. *Nat Rev Cancer* **2005**, *5*, 275-84.
7. Higgins, C. F. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **1992**, *8*, 67-113.
8. Szakacs, G.; Paterson, J. K.; Ludwig, J. A.; Booth-Genthe, C.; Gottesman, M. M. Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* **2006**, *5*, 219-34.
9. Labeed, F. H.; Coley, H. M.; Thomas, H.; Hughes, M. P. Assessment of multidrug resistance reversal using dielectrophoresis and flow cytometry. *Biophys J* **2003**, *85*, 2028-34.
10. Sarkadi, B.; Homolya, L.; Szakacs, G.; Varadi, A. Human multidrug resistance ABCB and ABCG transporters: participation in a chemoimmunity defense system. *Physiol Rev* **2006**, *86*, 1179-236.
11. Ambudkar, S. V.; Dey, S.; Hrycyna, C. A.; Ramachandra, M.; Pastan, I.; Gottesman, M. M. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu Rev Pharmacol Toxicol* **1999**, *39*, 361-98.
12. Benet, L. Z. The drug transporter-metabolism alliance: uncovering and defining the interplay. *Mol Pharm* **2009**, *6*, 1631-43.
13. Kemper, E. M.; van Zandbergen, A. E.; Cleypool, C.; Mos, H. A.; Boogerd, W.; Beijnen, J. H.; van Tellingen, O. Increased penetration of paclitaxel into the brain by inhibition of P-Glycoprotein. *Clin Cancer Res* **2003**, *9*, 2849-55.
14. Colabufo, N. A.; Berardi, F.; Contino, M.; Niso, M.; Perrone, R. ABC pumps and their role in active drug transport. *Curr Top Med Chem* **2009**, *9*, 119-29.
15. Broccatelli, F.; Carosati, E.; Neri, A.; Frosini, M.; Goracci, L.; Oprea, T. I.; Cruciani, G. A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields. *J Med Chem* **2011**.
16. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Nucleotide binding, ATP hydrolysis, and mutation of the catalytic carboxylates of human P-glycoprotein cause distinct conformational changes in the transmembrane segments. *Biochemistry* **2007**, *46*, 9328-36.

17. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Human P-glycoprotein is active when the two halves are clamped together in the closed conformation. *Biochem Biophys Res Commun* **395**, 436-40.
18. Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L.; Chang, G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **2009**, *323*, 1718-22.
19. Bruggemann, E. P.; Currier, S. J.; Gottesman, M. M.; Pastan, I. Characterization of the azidopine and vinblastine binding site of P-glycoprotein. *J Biol Chem* **1992**, *267*, 21020-6.
20. Greenberger, L. M. Major photoaffinity drug labeling sites for iodoaryl azidoprazosin in P-glycoprotein are within, or immediately C-terminal to, transmembrane domains 6 and 12. *J Biol Chem* **1993**, *268*, 11417-25.
21. Demmer, A.; Thole, H.; Kubesch, P.; Brandt, T.; Raida, M.; Fislage, R.; Tummeler, B. Localization of the iodomyacin binding site in hamster P-glycoprotein. *J Biol Chem* **1997**, *272*, 20913-9.
22. Wu, Q.; Bounaud, P. Y.; Kuduk, S. D.; Yang, C. P.; Ojima, I.; Horwitz, S. B.; Orr, G. A. Identification of the domains of photoincorporation of the 3'- and 7-benzophenone analogues of taxol in the carboxyl-terminal half of murine mdr1b P-glycoprotein. *Biochemistry* **1998**, *37*, 11272-9.
23. Isenberg, B.; Thole, H.; Tummeler, B.; Demmer, A. Identification and localization of three photobinding sites of iodoarylazidoprazosin in hamster P-glycoprotein. *Eur J Biochem* **2001**, *268*, 2629-34.
24. Loo, T. W.; Clarke, D. M. The Packing of the Transmembrane Segments of Human Multidrug Resistance P-glycoprotein Is Revealed by Disulfide Cross-linking Analysis. *Journal of Biological Chemistry* **2000**, *275*, 5253-5256.
25. Loo, T. W.; Clarke, D. M. Defining the Drug-binding Site in the Human Multidrug Resistance P-glycoprotein Using a Methanethiosulfonate Analog of Verapamil, MTS-verapamil. *Journal of Biological Chemistry* **2001**, *276*, 14972-14979.
26. Ecker, G. F.; Csaszar, E.; Kopp, S.; Plagens, B.; Holzer, W.; Ernst, W.; Chiba, P. Identification of ligand-binding regions of P-glycoprotein by activated-pharmacophore photoaffinity labeling and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Mol Pharmacol* **2002**, *61*, 637-48.
27. Pleban, K.; Kopp, S.; Csaszar, E.; Peer, M.; Hrebicek, T.; Rizzi, A.; Ecker, G. F.; Chiba, P. P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. *Mol Pharmacol* **2005**, *67*, 365-74.
28. Loo, T. W.; Clarke, D. M. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem* **2002**, *277*, 44332-8.
29. Qu, Q.; Sharom, F. J. Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet. *Biochemistry* **2002**, *41*, 4744-52.
30. Parveen, Z.; Stockner, T.; Bentele, C.; Pferschy, S.; Kraupp, M.; Freissmuth, M.; Ecker, G. F.; Chiba, P. Molecular dissection of dual pseudosymmetric solute translocation pathways in human p-glycoprotein. *Mol Pharmacol* **79**, 443-52.
31. Martin, C.; Berridge, G.; Higgins, C. F.; Mistry, P.; Charlton, P.; Callaghan, R. Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* **2000**, *58*, 624-32.
32. Sauna, Z. E.; Ambudkar, S. V. Characterization of the catalytic cycle of ATP hydrolysis by human P-glycoprotein. The two ATP hydrolysis events in a single

- catalytic cycle are kinetically similar but affect different functional outcomes. *J Biol Chem* **2001**, 276, 11653-61.
33. Callaghan, R.; Ford, R. C.; Kerr, I. D. The translocation mechanism of P-glycoprotein. *FEBS Lett* **2006**, 580, 1056-63.
34. Sauna, Z. E.; Ambudkar, S. V. About a switch: how P-glycoprotein (ABCB1) harnesses the energy of ATP binding and hydrolysis to do mechanical work. *Molecular Cancer Therapeutics* **2007**, 6, 13-23.
35. Nuti, S. L.; Rao, U. S. Proteolytic Cleavage of the Linker Region of the Human P-glycoprotein Modulates Its ATPase Function. *J Biol Chem* **2002**, 277, 29417-23.
36. Sato, T.; Kodan, A.; Kimura, Y.; Ueda, K.; Nakatsu, T.; Kato, H. Functional role of the linker region in purified human P-glycoprotein. *FEBS J* **2009**, 276, 3504-16.
37. Smith, P. C.; Karpowich, N.; Millen, L.; Moody, J. E.; Rosen, J.; Thomas, P. J.; Hunt, J. F. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell* **2002**, 10, 139-49.
38. Higgins, C. F.; Linton, K. J. The ATP switch model for ABC transporters. *Nat Struct Mol Biol* **2004**, 11, 918-26.
39. Moody, J. E.; Millen, L.; Binns, D.; Hunt, J. F.; Thomas, P. J. Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. *J Biol Chem* **2002**, 277, 21111-4.
40. Martin, C.; Berridge, G.; Mistry, P.; Higgins, C.; Charlton, P.; Callaghan, R. Drug binding sites on P-glycoprotein are altered by ATP binding prior to nucleotide hydrolysis. *Biochemistry* **2000**, 39, 11901-6.
41. Rosenberg, M. F.; Velarde, G.; Ford, R. C.; Martin, C.; Berridge, G.; Kerr, I. D.; Callaghan, R.; Schmidlin, A.; Wooding, C.; Linton, K. J.; Higgins, C. F. Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J* **2001**, 20, 5615-25.
42. Martin, C.; Higgins, C. F.; Callaghan, R. The vinblastine binding site adopts high- and low-affinity conformations during a transport cycle of P-glycoprotein. *Biochemistry* **2001**, 40, 15733-42.
43. Rothnie, A.; Storm, J.; Campbell, J.; Linton, K. J.; Kerr, I. D.; Callaghan, R. The topography of transmembrane segment six is altered during the catalytic cycle of P-glycoprotein. *J Biol Chem* **2004**, 279, 34913-21.
44. Senior, A. E.; al-Shawi, M. K.; Urbatsch, I. L. The catalytic cycle of P-glycoprotein. *FEBS Lett* **1995**, 377, 285-9.
45. Senior, A. E. Catalytic mechanism of P-glycoprotein. *Acta Physiol Scand Suppl* **1998**, 643, 213-8.
46. Siarheyeva, A.; Liu, R.; Sharom, F. J. Characterization of an asymmetric occluded state of P-glycoprotein with two bound nucleotides: implications for catalysis. *J Biol Chem* **2010**, 285, 7575-86.
47. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Human P-glycoprotein is active when the two halves are clamped together in the closed conformation. *Biochem Biophys Res Commun* **2010**, 395, 436-40.
48. Dawson, R. J.; Locher, K. P. Structure of a bacterial multidrug ABC transporter. *Nature* **2006**, 443, 180-5.
49. Gottesman, M. M.; Ambudkar, S. V.; Xia, D. Structure of a multidrug transporter. *Nat Biotechnol* **2009**, 27, 546-7.
50. Kuhnle, M.; Egger, M.; Muller, C.; Mahringer, A.; Bernhardt, G.; Fricker, G.; Konig, B.; Buschauer, A. Potent and selective inhibitors of breast cancer resistance

protein (ABCG2) derived from the p-glycoprotein (ABCB1) modulator tariquidar. *J Med Chem* **2009**, 52, 1190-7.

51. Ford, J. M. Experimental reversal of P-glycoprotein-mediated multidrug resistance by pharmacological chemosensitisers. *Eur J Cancer* **1996**, 32A, 991-1001.

52. Kannan, P.; John, C.; Zoghbi, S. S.; Halldin, C.; Gottesman, M. M.; Innis, R. B.; Hall, M. D. Imaging the function of P-glycoprotein with radiotracers: pharmacokinetics and in vivo applications. *Clin Pharmacol Ther* **2009**, 86, 368-77.

53. Raub, T. J. P-Glycoprotein Recognition of Substrates and Circumvention through Rational Drug Design. *Molecular Pharmaceutics* **2005**, 3, 3-25.

54. Pleban, K.; Ecker, G. F. Inhibitors of p-glycoprotein--lead identification and optimisation. *Mini Rev Med Chem* **2005**, 5, 153-63.

55. Klepsch, F.; Jabeen, I.; Chiba, P.; Ecker, G. F. Pharmacoinformatic approaches to design natural product type ligands of ABC-transporters. *Curr Pharm Des* 16, 1742-52.

56. Ekins, S.; Ecker, G. F.; Chiba, P.; Swaan, P. W. Future directions for drug transporter modelling. *Xenobiotica* **2007**, 37, 1152-70.

57. Tmej, C.; Chiba, P.; Huber, M.; Richter, E.; Hitzler, M.; Schaper, K. J.; Ecker, G. A combined Hansch/Free-Wilson approach as predictive tool in QSAR studies on propafenone-type modulators of multidrug resistance. *Arch Pharm (Weinheim)* **1998**, 331, 233-40.

58. Kaiser, D.; Smiesko, M.; Kopp, S.; Chiba, P.; Ecker, G. F. Interaction field based and hologram based QSAR analysis of propafenone-type modulators of multidrug resistance. *Med Chem* **2005**, 1, 431-44.

59. Tmej, C.; Chiba, P.; Schaper, K. J.; Ecker, G.; Fleischhacker, W. Artificial neural networks as versatile tools for prediction of MDR-modulatory activity. *Adv Exp Med Biol* **1999**, 457, 95-105.

60. Klein, C.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity based SAR (SIBAR) as tool for early ADME profiling. *J Comput Aided Mol Des* **2002**, 16, 785-93.

61. Wiese, M.; Pajeva, I. K. Structure-activity relationships of multidrug resistance reversers. *Curr Med Chem* **2001**, 8, 685-713.

62. Hiessbock, R.; Wolf, C.; Richter, E.; Hitzler, M.; Chiba, P.; Kratzel, M.; Ecker, G. Synthesis and in vitro multidrug resistance modulating activity of a series of dihydrobenzopyrans and tetrahydroquinolines. *J Med Chem* **1999**, 42, 1921-6.

63. Pajeva, I.; Wiese, M. Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: a comparative molecular field analysis study. *J Med Chem* **1998**, 41, 1815-26.

64. Pajeva, I. K.; Wiese, M. A Comparative Molecular Field Analysis of Propafenone-type Modulators of Cancer Multidrug Resistance. *Quantitative Structure-Activity Relationships* **1998**, 17, 301-312.

65. Pleban, K.; Hoffer, C.; Kopp, S.; Peer, M.; Chiba, P.; Ecker, G. F. Intramolecular distribution of hydrophobicity influences pharmacological activity of propafenone-type MDR modulators. *Arch Pharm (Weinheim)* **2004**, 337, 328-34.

66. König, G.; Chiba, P.; Ecker, G. F. Hydrophobic moments as physicochemical descriptors in structure-activity relationship studies of P-glycoprotein inhibitors. *Monatshefte für Chemie / Chemical Monthly* **2008**, 139, 401-405.

67. Wang, R. B.; Kuo, C. L.; Lien, L. L.; Lien, E. J. Structure-activity relationship: analyses of p-glycoprotein substrates and inhibitors. *J Clin Pharm Ther* **2003**, 28, 203-28.

68. Seelig, A. A general pattern for substrate recognition by P-glycoprotein. *Eur J Biochem* **1998**, 251, 252-61.
69. Seelig, A. How does P-glycoprotein recognize its substrates? *Int J Clin Pharmacol Ther* **1998**, 36, 50-4.
70. Ecker, G.; Huber, M.; Schmid, D.; Chiba, P. The importance of a nitrogen atom in modulators of multidrug resistance. *Mol Pharmacol* **1999**, 56, 791-6.
71. Pajeva, I. K.; Wiese, M. Pharmacophore model of drugs involved in P-glycoprotein multidrug resistance: explanation of structural variety (hypothesis). *J Med Chem* **2002**, 45, 5671-86.
72. Pajeva, I. K.; Wiese, M. Structure-activity relationships of tariquidar analogs as multidrug resistance modulators. *AAPS J* **2009**, 11, 435-44.
73. Globisch, C.; Pajeva, I. K.; Wiese, M. Structure-activity relationships of a series of tariquidar analogs as multidrug resistance modulators. *Bioorg Med Chem* **2006**, 14, 1588-98.
74. Roe, M.; Folkes, A.; Ashworth, P.; Brumwell, J.; Chima, L.; Hunjan, S.; Pretswell, I.; Dangerfield, W.; Ryder, H.; Charlton, P. Reversal of P-glycoprotein mediated multidrug resistance by novel anthranilamide derivatives. *Bioorg Med Chem Lett* **1999**, 9, 595-600.
75. Kothandan, G.; Gadhe, C. G.; Madhavan, T.; Choi, C. H.; Cho, S. J. Docking and 3D-QSAR (quantitative structure activity relationship) studies of flavones, the potent inhibitors of p-glycoprotein targeting the nucleotide binding domain. *Eur J Med Chem* **2011**.
76. Jekerle, V.; Klinkhammer, W.; Reilly, R. M.; Piquette-Miller, M.; Wiese, M. Novel tetrahydroisoquinolin-ethyl-phenylamine based multidrug resistance inhibitors with broad-spectrum modulating properties. *Cancer Chemother Pharmacol* **2007**, 59, 61-9.
77. Muller, H.; Pajeva, I. K.; Globisch, C.; Wiese, M. Functional assay and structure-activity relationships of new third-generation P-glycoprotein inhibitors. *Bioorg Med Chem* **2008**, 16, 2448-62.
78. Jilek, R. J.; Cramer, R. D. Topomers: a validated protocol for their self-consistent generation. *J Chem Inf Comput Sci* **2004**, 44, 1221-7.
79. Klinkhammer, W.; Muller, H.; Globisch, C.; Pajeva, I. K.; Wiese, M. Synthesis and biological evaluation of a small molecule library of 3rd generation multidrug resistance modulators. *Bioorg Med Chem* **2009**, 17, 2524-35.
80. Gadhe, C. G.; Madhavan, T.; Kothandan, G.; Cho, S. J. In silico quantitative structure-activity relationship studies on P-gp modulators of tetrahydroisoquinoline-ethyl-phenylamine series. *BMC Struct Biol* 11, 5.
81. Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Three-dimensional quantitative structure-activity relationships of inhibitors of P-glycoprotein. *Mol Pharmacol* **2002**, 61, 964-73.
82. Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of three-dimensional quantitative structure-activity relationships of P-glycoprotein inhibitors and substrates. *Mol Pharmacol* **2002**, 61, 974-81.
83. Langer, T.; Eder, M.; Hoffmann, R. D.; Chiba, P.; Ecker, G. F. Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model. *Arch Pharm (Weinheim)* **2004**, 337, 317-27.

84. Chang, C.; Bahadduri, P. M.; Polli, J. E.; Swaan, P. W.; Ekins, S. Rapid identification of P-glycoprotein substrates and inhibitors. *Drug Metab Dispos* **2006**, *34*, 1976-84.
85. Pearce, H. L.; Safa, A. R.; Bach, N. J.; Winter, M. A.; Cirtain, M. C.; Beck, W. T. Essential features of the P-glycoprotein pharmacophore as defined by a series of reserpine analogs that modulate multidrug resistance. *Proc Natl Acad Sci U S A* **1989**, *86*, 5128-32.
86. Pearce, H. L.; Winter, M. A.; Beck, W. T. Structural characteristics of compounds that modulate P-glycoprotein-associated multidrug resistance. *Adv Enzyme Regul* **1990**, *30*, 357-73.
87. Chang, C.; Swaan, P. W. Computational approaches to modeling drug transporters. *Eur J Pharm Sci* **2006**, *27*, 411-24.
88. Cianchetta, G.; Singleton, R. W.; Zhang, M.; Wildgoose, M.; Giesing, D.; Fravolini, A.; Cruciani, G.; Vaz, R. J. A pharmacophore hypothesis for P-glycoprotein substrate recognition using GRIND-based 3D-QSAR. *J Med Chem* **2005**, *48*, 2927-35.
89. Fontaine, F.; Pastor, M.; Zamora, I.; Sanz, F. Anchor-GRIND: filling the gap between standard 3D QSAR and the GRid-INdependent descriptors. *J Med Chem* **2005**, *48*, 2687-94.
90. Cruciani, G.; Pastor, M.; Guba, W. VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. *Eur J Pharm Sci* **2000**, *11* Suppl 2, S29-39.
91. Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J Med Chem* **2005**, *48*, 6970-9.
92. Boccard, J.; Bajot, F.; Di Pietro, A.; Rudaz, S.; Boumendjel, A.; Nicolle, E.; Carrupt, P. A. A 3D linear solvation energy model to quantify the affinity of flavonoid derivatives toward P-glycoprotein. *Eur J Pharm Sci* **2009**, *36*, 254-64.
93. Boumendjel, A.; Beney, C.; Deka, N.; Mariotte, A. M.; Lawson, M. A.; Trompier, D.; Baubichon-Cortay, H.; Di Pietro, A. 4-Hydroxy-6-methoxyaurones with high-affinity binding to cytosolic domain of P-glycoprotein. *Chem Pharm Bull (Tokyo)* **2002**, *50*, 854-6.
94. Boumendjel, A.; Di Pietro, A.; Dumontet, C.; Barron, D. Recent advances in the discovery of flavonoids and analogs with high-affinity binding to P-glycoprotein responsible for cancer cell multidrug resistance. *Med Res Rev* **2002**, *22*, 512-29.
95. Taft, R. W.; Abraham, M. H.; Famini, G. R.; Doherty, R. M.; Abboud, J. L.; Kamlet, M. J. Solubility properties in polymers and biological media 5: an analysis of the physicochemical properties which influence octanol-water partition coefficients of aliphatic and aromatic solutes. *J Pharm Sci* **1985**, *74*, 807-14.
96. Crivori, P.; Reinach, B.; Pezzetta, D.; Poggesi, I. Computational models for identifying potential P-glycoprotein substrates and inhibitors. *Mol Pharm* **2006**, *3*, 33-44.
97. Zdrzil, B.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity-Based Descriptors (SIBAR) as Tool for QSAR Studies on P-Glycoprotein Inhibitors: Influence of the Reference Set. *QSAR & Combinatorial Science* **2007**, *26*, 669-678.
98. Gombar, V. K.; Polli, J. W.; Humphreys, J. E.; Wring, S. A.; Serabjit-Singh, C. S. Predicting P-glycoprotein substrates by a quantitative structure-activity relationship model. *J Pharm Sci* **2004**, *93*, 957-68.
99. Didziapetris, R.; Japertas, P.; Avdeef, A.; Petrauskas, A. Classification analysis of P-glycoprotein substrate specificity. *J Drug Target* **2003**, *11*, 391-406.

100. Gleeson, M. P. Generation of a set of simple, interpretable ADMET rules of thumb. *J Med Chem* **2008**, 51, 817-34.
101. Estrada, E.; Molina, E.; Nodarse, D.; Uriarte, E. Structural contributions of substrates to their binding to P-Glycoprotein. A TOPS-MODE approach. *Curr Pharm Des* **2010**, 16, 2676-709.
102. Kaiser, D.; Terfloth, L.; Kopp, S.; Schulz, J.; de Laet, R.; Chiba, P.; Ecker, G. F.; Gasteiger, J. Self-organizing maps for identification of new inhibitors of P-glycoprotein. *J Med Chem* **2007**, 50, 1698-702.
103. Huang, J.; Ma, G.; Muhammad, I.; Cheng, Y. Identifying P-glycoprotein substrates using a support vector machine optimized by a particle swarm. *J Chem Inf Model* **2007**, 47, 1638-47.
104. Demel, M. A.; Kraemer, O.; Etmayer, P.; Haaksma, E.; Ecker, G. F. Ensemble Rule-Based Classification of Substrates of the Human ABC-Transporter ABCB1 Using Simple Physicochemical Descriptors. *Molecular Informatics* **29**, 233-242.
105. Friedman, J. H.; Popescu, B. E. *Predictive Learning via Rule Ensembles*; 2005.
106. Ecker, G. F.; Stockner, T.; Chiba, P. Computational models for prediction of interactions with ABC-transporters. *Drug Discov Today* **2008**, 13, 311-7.
107. Klepsch, F.; Stockner, T.; Erker, T.; Muller, M.; Chiba, P.; Ecker, G. F. Using structural and mechanistic information to design novel inhibitors/substrates of P-glycoprotein. *Curr Top Med Chem* **2010**, 10, 1769-74.
108. Klepsch, F.; Ecker, G. F. Impact of the Recent Mouse P-Glycoprotein Structure for Structure-Based Ligand Design. *Molecular Informatics* **29**, 276-286.
109. Kerr, I. D.; Jones, P. M.; George, A. M. Multidrug efflux pumps: the structures of prokaryotic ATP-binding cassette transporter efflux pumps and implications for our understanding of eukaryotic P-glycoproteins and homologues. *FEBS J* **2009**, 277, 550-63.
110. Chang, G.; Roth, C. B. Structure of MsbA from E. coli: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science* **2001**, 293, 1793-800.
111. Stenham, D. R.; Campbell, J. D.; Sansom, M. S.; Higgins, C. F.; Kerr, I. D.; Linton, K. J. An atomic detail model for the human ATP binding cassette transporter P-glycoprotein derived from disulfide cross-linking and homology modeling. *FASEB J* **2003**, 17, 2287-9.
112. Seigneuret, M.; Garnier-Suillerot, A. A structural model for the open conformation of the mdr1 P-glycoprotein based on the MsbA crystal structure. *J Biol Chem* **2003**, 278, 30115-24.
113. Vandevuer, S.; Van Bambeke, F.; Tulkens, P. M.; Prevost, M. Predicting the three-dimensional structure of human P-glycoprotein in absence of ATP by computational techniques embodying crosslinking data: insight into the mechanism of ligand migration and binding sites. *Proteins* **2006**, 63, 466-78.
114. Chang, G. Structure of MsbA from *Vibrio cholera*: a multidrug resistance ABC transporter homolog in a closed conformation. *J Mol Biol* **2003**, 330, 419-30.
115. Reyes, C. L.; Chang, G. Structure of the ABC transporter MsbA in complex with ADP.vanadate and lipopolysaccharide. *Science* **2005**, 308, 1028-31.
116. Ravna, A. W.; Sylte, I.; Sager, G. Molecular model of the outward facing state of the human P-glycoprotein (ABCB1), and comparison to a model of the human MRP5 (ABCC5). *Theor Biol Med Model* **2007**, 4, 33.

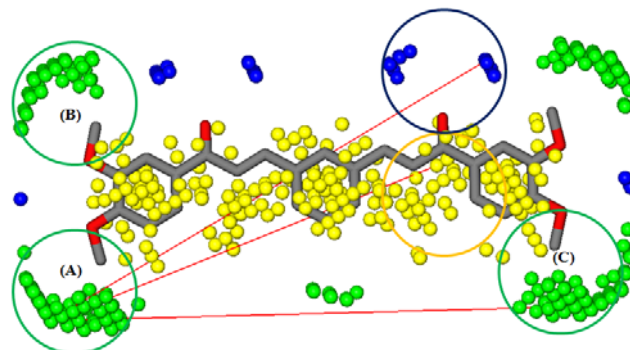
117. O'Mara, M. L.; Tieleman, D. P. P-glycoprotein models of the apo and ATP-bound states based on homology with Sav1866 and MalK. *FEBS Lett* **2007**, 581, 4217-22.
118. Becker, J. P.; Depret, G.; Van Bambeke, F.; Tulkens, P. M.; Prevost, M. Molecular models of human P-glycoprotein in two different catalytic states. *BMC Struct Biol* **2009**, 9, 3.
119. Stockner, T.; de Vries, S. J.; Bonvin, A. M.; Ecker, G. F.; Chiba, P. Data-driven homology modelling of P-glycoprotein in the ATP-bound state indicates flexibility of the transmembrane domains. *FEBS J* **2009**, 276, 964-72.
120. Sakurai, A.; Onishi, Y.; Hirano, H.; Seigneuret, M.; Obanayama, K.; Kim, G.; Liew, E. L.; Sakaeda, T.; Yoshiura, K.; Niikawa, N.; Sakurai, M.; Ishikawa, T. Quantitative structure-activity relationship analysis and molecular dynamics simulation to functionally validate nonsynonymous polymorphisms of human ABC transporter ABCB1 (P-glycoprotein/MDR1). *Biochemistry* **2007**, 46, 7678-93.
121. Globisch, C.; Pajeva, I. K.; Wiese, M. Identification of putative binding sites of P-glycoprotein based on its homology model. *ChemMedChem* **2008**, 3, 280-95.
122. Klepsch, F.; Chiba, P.; Ecker, G. F. Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Comput Biol* **2011**, 7, e1002036.
123. Ward, A.; Reyes, C. L.; Yu, J.; Roth, C. B.; Chang, G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A* **2007**, 104, 19005-10.
124. Ravna, A. W.; Sylte, I.; Sager, G. Binding site of ABC transporter homology models confirmed by ABCB1 crystal structure. *Theor Biol Med Model* **2009**, 6, 20.
125. Pajeva, I. K.; Globisch, C.; Wiese, M. Comparison of the inward- and outward-open homology models and ligand binding of human P-glycoprotein. *FEBS J* **2009**, 276, 7016-26.
126. Locher, K. P.; Lee, A. T.; Rees, D. C. The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* **2002**, 296, 1091-8.
127. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Disulfide cross-linking analysis shows that transmembrane segments 5 and 8 of human P-glycoprotein are close together on the cytoplasmic side of the membrane. *J Biol Chem* **2004**, 279, 7692-7.
128. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Val133 and Cys137 in transmembrane segment 2 are close to Arg935 and Gly939 in transmembrane segment 11 of human P-glycoprotein. *J Biol Chem* **2004**, 279, 18232-8.
129. Lee, J. Y.; Urbatsch, I. L.; Senior, A. E.; Wilkens, S. Projection structure of P-glycoprotein by electron microscopy. Evidence for a closed conformation of the nucleotide binding domains. *J Biol Chem* **2002**, 277, 40125-31.
130. Buchaklian, A. H.; Funk, A. L.; Klug, C. S. Resting state conformation of the MsbA homodimer as studied by site-directed spin labeling. *Biochemistry* **2004**, 43, 8600-6.
131. Chen, J.; Lu, G.; Lin, J.; Davidson, A. L.; Quioco, F. A. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol Cell* **2003**, 12, 651-61.
132. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Arginines in the first transmembrane segment promote maturation of a P-glycoprotein processing mutant by hydrogen bond interactions with tyrosines in transmembrane segment 11. *J Biol Chem* **2008**, 283, 24860-70.



133. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. *Biochem J* **2006**, 396, 537-45.
134. Loo, T. W.; Bartlett, M. C.; Clarke, D. M. Identification of residues in the drug translocation pathway of the human multidrug resistance P-glycoprotein by arginine mutagenesis. *J Biol Chem* **2009**, 284, 24074-87.
135. Dey, S.; Hafkemeyer, P.; Pastan, I.; Gottesman, M. M. A single amino acid residue contributes to distinct mechanisms of inhibition of the human multidrug transporter by stereoisomers of the dopamine receptor antagonist flupentixol. *Biochemistry* **1999**, 38, 6630-9.
136. Pajeva, I. K.; Globisch, C.; Wiese, M. Combined pharmacophore modeling, docking, and 3D QSAR studies of ABCB1 and ABCC1 transporter inhibitors. *ChemMedChem* **2009**, 4, 1883-96.
137. Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional group contributions to drug-receptor interactions. *J Med Chem* **1984**, 27, 1648-57.
138. Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **2007**, 6, 881-90.
139. *Chemical Computing Group, Inc; Molecular Operating Environment (MOE)*, Quebec, Canada, 2010.
140. Hansch, C.; Fujita, T. r-s-p Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc* **1965**, 86, 1616-1626.
141. Hogg, R. V.; Tanis, E. A. In *Probability and Statistical Inference*, Macmillan Publishing: New York, 1993.
142. Jabeen, I.; Wetwitayaklung, P.; Klepsch, F.; Parveen, Z.; Chiba, P.; Ecker, G. F. Probing the stereoselectivity of P-glycoprotein-synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-b][1,4]oxazines. *Chem Commun (Camb)* **2011**, 47, 2586-8.
143. Durán, Á.; Pastor, M. *Pentacle An advanced tool for computing and handling GRid- INdependent Descriptors. User manual version 0.9*.
144. Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J Med Chem* **2000**, 43, 3233-43.
145. Holtt, V.; Kouba, M.; Dietel, M.; Vogt, G. Stereoisomers of calcium antagonists which differ markedly in their potencies as calcium blockers are equally effective in modulating drug transport by P-glycoprotein. *Biochem Pharmacol* **1992**, 43, 2601-8.
146. Neuhoff, S.; Langguth, P.; Dressler, C.; Andersson, T. B.; Regardh, C. G.; Spahn-Langguth, H. Affinities at the verapamil binding site of MDR1-encoded P-glycoprotein: drugs and analogs, stereoisomers and metabolites. *Int J Clin Pharmacol Ther* **2000**, 38, 168-79.
147. Bhatia, P.; Kolinski, M.; Moaddel, R.; Jozwiak, K.; Wainer, I. W. Determination and modelling of stereoselective interactions of ligands with drug transporters: a key dimension in the understanding of drug disposition. *Xenobiotica* **2008**, 38, 656-75.
148. Ryckmans, T.; Edwards, M. P.; Horne, V. A.; Correia, A. M.; Owen, D. R.; Thompson, L. R.; Tran, I.; Tutt, M. F.; Young, T. Rapid assessment of a novel series of selective CB(2) agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis. *Bioorg Med Chem Lett* **2009**, 19, 4406-9.
149. Keseru, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat Rev Drug Discov* **2009**, 8, 203-12.
150. Mortenson, P. N.; Murray, C. W. Assessing the lipophilicity of fragments and early hits. *J Comput Aided Mol Des* **2011**.

151. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. *Proc Natl Acad Sci U S A* **1999**, 96, 9997-10002.
152. Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **2004**, 9, 430-1.
153. Abad-Zapatero, C. Ligand efficiency indices for effective drug discovery. *Expert Opinion on Drug Discovery* **2007**, 2, 469-488.
154. Abad-Zapatero, C.; Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, 10, 464-469.
155. Reynolds, C. H.; Bembenek, S. D.; Tounge, B. A. The role of molecular size in ligand efficiency. *Bioorg Med Chem Lett* **2007**, 17, 4258-61.
156. Verdonk, M. L.; Rees, D. C. Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem* **2008**, 3, 1179-80.
157. Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand binding efficiency: trends, physical basis, and implications. *J Med Chem* **2008**, 51, 2432-8.
158. Ecker, G. F.; Pleban, K.; Kopp, S.; Csaszar, E.; Poelarends, G. J.; Putman, M.; Kaiser, D.; Konings, W. N.; Chiba, P. A three-dimensional model for the substrate binding domain of the multidrug ATP binding cassette transporter LmrA. *Mol Pharmacol* **2004**, 66, 1169-79.
159. Parveen, Z.; Stockner, T.; Bentele, C.; Pferschy, S.; Kraupp, M.; Freissmuth, M.; Ecker, G. F.; Chiba, P. Molecular Dissection of Dual Pseudosymmetric Solute Translocation Pathways in Human P-Glycoprotein. *Mol Pharmacol* **2011**.
160. <https://www.ebi.ac.uk/chembl/db/>.
161. Thai, K. M.; Ecker, G. F. A binary QSAR model for classification of hERG potassium channel blockers. *Bioorg Med Chem* **2008**, 16, 4107-19.
162. Mattson, R. J.; Catt, J. D.; Denhart, D. J.; Deskus, J. A.; Ditta, J. L.; Higgins, M. A.; Marcin, L. R.; Sloan, C. P.; Beno, B. R.; Gao, Q.; Cunningham, M. A.; Mattson, G. K.; Molski, T. F.; Taber, M. T.; Lodge, N. J. Conformationally restricted homotryptamines. 2. Indole cyclopropylmethylamines as selective serotonin reuptake inhibitors. *J Med Chem* **2005**, 48, 6023-34.
163. Jarkas, N.; Voll, R. J.; Williams, L.; Votaw, J. R.; Owens, M.; Goodman, M. M. Synthesis and in vivo evaluation of halogenated N,N-dimethyl-2-(2'-amino-4'-hydroxymethylphenylthio)benzylamine derivatives as PET serotonin transporter ligands. *J Med Chem* **2008**, 51, 271-81.
164. Parhi, A. K.; Wang, J. L.; Oya, S.; Choi, S. R.; Kung, M. P.; Kung, H. F. 2-(2'-((dimethylamino)methyl)-4'-(fluoroalkoxy)-phenylthio)benzenamine derivatives as serotonin transporter imaging agents. *J Med Chem* **2007**, 50, 6673-84.
165. Boot, J. R.; Boulet, S. L.; Clark, B. P.; Cases-Thomas, M. J.; Delhaye, L.; Diker, K.; Fairhurst, J.; Findlay, J.; Gallagher, P. T.; Gilmore, J.; Harris, J. R.; Masters, J. J.; Mitchell, S. N.; Naik, M.; Simmonds, R. G.; Smith, S. M.; Richards, S. J.; Timms, G. H.; Whatton, M. A.; Wolfe, C. N.; Wood, V. A. N-Alkyl-N-arylmethylpiperidin-4-amines: novel dual inhibitors of serotonin and norepinephrine reuptake. *Bioorg Med Chem Lett* **2006**, 16, 2714-8.
166. Lucas, M. C.; Carter, D. S.; Cai, H. Y.; Lee, E. K.; Schoenfeld, R. C.; Steiner, S.; Villa, M.; Weikert, R. J.; Iyer, P. S. Novel, achiral aminoheterocycles as selective monoamine reuptake inhibitors. *Bioorg Med Chem Lett* **2009**, 19, 4630-3.
167. Zeng, F.; Stehouwer, J. S.; Jarkas, N.; Voll, R. J.; Williams, L.; Camp, V. M.; Votaw, J. R.; Owens, M. J.; Kilts, C. D.; Nemeroff, C. B.; Goodman, M. M. Synthesis and biological evaluation of 2beta,3alpha-(substituted phenyl)nortropanes as potential norepinephrine transporter imaging agents. *Bioorg Med Chem Lett* **2007**, 17, 3044-7.

168. Voelker, T.; Xia, H.; Fandrick, K.; Johnson, R.; Janowsky, A.; Cashman, J. R. 2,5-Disubstituted tetrahydrofurans as selective serotonin re-uptake inhibitors. *Bioorg Med Chem* **2009**, *17*, 2047-68.
169. Cashman, J. R.; Voelker, T.; Zhang, H. T.; O'Donnell, J. M. Dual inhibitors of phosphodiesterase-4 and serotonin reuptake. *J Med Chem* **2009**, *52*, 1530-9.
170. Wellfelt, K.; Skold, A. C.; Wallin, A.; Danielsson, B. R. Teratogenicity of the class III antiarrhythmic drug almokalant. Role of hypoxia and reactive oxygen species. *Reprod Toxicol* **1999**, *13*, 93-101.
171. Houltz, B.; Darpo, B.; Swedberg, K.; Blomstrom, P.; Brachmann, J.; Crijns, H. J.; Jensen, S. M.; Svernhage, E.; Vallin, H.; Edvardsson, N. Effects of the Ikr-blocker almokalant and predictors of conversion of chronic atrial tachyarrhythmias to sinus rhythm. A prospective study. *Cardiovasc Drugs Ther* **1999**, *13*, 329-38.



In this part both 2D- and 3D-QSAR models were derived by using general physicochemical and GRID-Independent Molecular Descriptors (GRIND) for prediction of chalcone -ABCB1 interaction.

## Contents

### Introduction

### Results and discussion

#### Chemistry

#### Structure Activity Relationships

#### 2D-QSAR

#### 3D-QSAR

### Conclusion

### Experimental Section

#### Chemistry

#### Biology

### Computational Methods

### Acknowledgements

### References

① **Information:** This chapter was a pre-submission phase of a manuscript by Brunhofer Gerda, **Jabeen Ishrat**, Parveen Zahida, Berner Heinz, Manuel Pastor, Chiba Peter, Erker Thomas and Ecker Gerhard F\*

☞ **Appendix available: Page 170-172**

# Synthesis, ABCB1 inhibitory activity and 3D- QSAR studies of a series of new chalcone derivatives

Brunhofer Gerda<sup>1#</sup>, Jabeen Ishrat<sup>1#</sup>, Parveen Zahida<sup>2#</sup>, Berner Heinz<sup>1</sup>, Manuel Pastor,  
Chiba Peter<sup>2</sup>, Erker Thomas, and Ecker Gerhard F<sup>1\*</sup>

<sup>1</sup>University of Vienna, Department of Medicinal Chemistry, Althanstrasse 14, 1090  
Wien, Austria <sup>2</sup>Medical University of Vienna, Institute of Medical Chemistry,  
Währinger Straße 10, 1090 Wien, Austria

\*Send correspondence to: [gerhard.f.ecker@univie.ac.at](mailto:gerhard.f.ecker@univie.ac.at)

<sup>#</sup>These authors contributed equally to this study.

**Abstract**

ABCB1/P-glycoprotein (P-gp) is an ATP dependent efflux transporter often linked to multidrug resistance in tumors and to modulation of ADME-tox properties of drugs candidates. After withdrawal of several modulators of P-gp from phase II and phase III clinical trials, development of new MDR modulators with better potencies and reduced toxicity is still of an open challenge. Extensive SAR- and QSAR-studies revealed the importance of distinct pharmacophoric features, such as H-bond acceptors, aromatic rings and a basic nitrogen atom. In an attempt to further enhance the chemical space of P-gp inhibitors and to challenge the necessity of a basic nitrogen atom, a set of chalcone- and indanone-type compounds has been synthesised and biologically tested in a daunomycin efflux assay. Several compounds showed promising activity in the low nM range, with the indanone derivative (**10**) being the most active analog (42.5 nM). In order to elucidate the main pharmacophoric features influencing P-gp inhibitory activity, GRID-independent molecular descriptor (GRIND) and general physicochemical descriptors were computed. The obtained 3D-QSAR model identifies specific distance ranges of significant pharmacophoric features, such as hydrogen bond acceptors, hydrophobic groups, as well as steric hot spots. 2D-QSAR studies identified the hydrophobic Van der Waals surface area as well as the number of rotatable bonds as main properties influencing activity.

**Keywords:** ABCB1, P-glycoprotein, Chalcone, GRIND

**Introduction**

Human P-glycoprotein (ABCB1, P-gp) is a multispecific drug efflux transporter which mediates resistance of cancer cells to cytotoxic drugs.<sup>1</sup> It belongs to the family of ATP-binding cassette transport proteins that protect cells from the toxic effect of compounds through active outward transfer.<sup>2</sup> In fact it has been demonstrated in animal models, that even low levels of expression can cause resistance of tumors to cytotoxic drugs. Early on, the concept of simultaneous administration of anticancer drugs with inhibitors of ABCB1 has been advocated as a concept for evading resistance. However, clinical studies have not lived up to the high expectations, and several phase II and phase III clinical studies have been terminated prematurely because of severe P-gp inhibition related side effects. Nevertheless, a proof-of-concept study illustrated that co-

administration of the anticancer agent topotecan and elacridar, an inhibitor of both ABCB1 and ABCG2 (breast cancer resistance protein, BCRP), significantly increased the bioavailability of the anticancer drug and reduced inter-patient variability.<sup>3</sup> In order to find more potent ABC transporter inhibitors, natural products often provide interesting scaffolds and serve as lead structures, as recently reviewed by Klepsch *et al.*<sup>4</sup> Flavonoids are polyphenolic compounds found to be ubiquitous in the plant kingdom. Naturally occurring flavonoids possess inhibitory activity towards ABCG2.<sup>5</sup> But also ABCB1 is influenced by flavonoids as was shown by Zhang and colleagues. They reported that quercetin, genistein and morin had inhibitory activity towards ABCB1-mediated daunomycin transport.<sup>6</sup> Recently, a QSAR model based on naturally occurring flavonoids was published describing their modulating effect on ABCB1.<sup>7</sup>

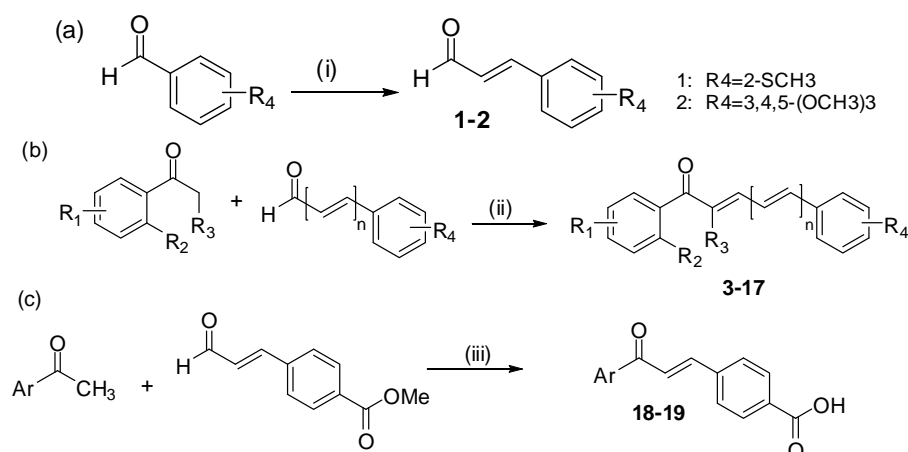
For chalcones, which are biosynthetic precursors of flavonoids, detailed structure activity relationship information is still missing. Similar to the structurally related flavonoids, chalcones are described to possess both ABCB1 and ABCG2 inhibitory properties. Recent studies showed that basic as well as non-basic chalcone derivatives modulate ABC transporters.<sup>8-10</sup> These data together with those presented herein suggest that chalcones offer a new compound class which can be used to restore sensitivity of cells to chemotherapeutic agents (chemo-sensitizers).

### **Results and discussion**

**Chemistry.** Compounds **3-24** were synthesized as outlined in schemes 1 and 2. All compounds are based on the chalcone scaffold, 1,3-diphenyl-2-propen-1-one, which was used as core structure. Substituents on both aromatic rings are mostly methoxy groups, which represent a pharmacophoric substructure for ABCB1 interaction.<sup>11</sup> But also derivatives bearing a combination of methoxy together with thiomethyl and tertiary amino groups showed high ABCB1 inhibitory activity. Further chemical variations concerned the length and type of the spacer linking the two phenyl rings (vinylogous and “bridged” chalcones). The substances were obtained by base-catalyzed reaction of a corresponding acetophenone derivative with an appropriate cinnam- or benzaldehyde (Claisen-Schmidt condensation). Dimerising the core structure resulted in compounds **20-23**, which were received by the same procedure. Compound **19** is an example of a bioisosteric replacement of the phenyl part of compound **18** against a ferrocene feature

which is a strategy often used in medicinal chemistry.<sup>12,13</sup> In compound **24**, one aromatic ring was exchanged by a cyclohexenol moiety. This compound was obtained by a DBU-methanol-promoted reaction of ethyl 2-cyclohexanone carboxylate and 2-methoxy cinnamaldehyde.<sup>14</sup>

**Scheme 1.** Synthesis of (a) compounds **1-2**, (b) compounds **3-17** and (c) compounds **18** and **19<sup>a</sup>**.

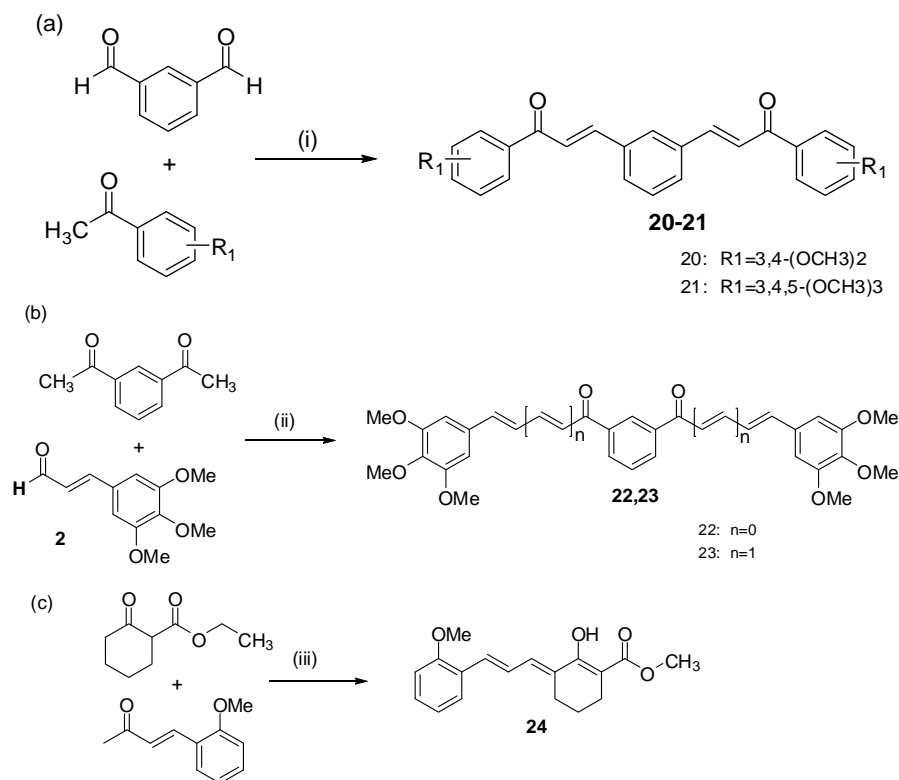


- 3:  $\text{R}_1=2\text{-OCH}_3, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2,4,6\text{-(CH}_3)_3, n=0$   
 4:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2,3\text{-(-CH=CH-CH=CH-), } 4\text{-OCH}_3, n=0$   
 5:  $\text{R}_1=3,4,5\text{-(OCH}_3)_3, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2\text{-OCH}_3, n=1$   
 6:  $\text{R}_1=3,5\text{-(OCH}_3)_2, 4\text{-OH}, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2\text{-OCH}_3, n=1$   
 7:  $\text{R}_1=3,4,5\text{-(OCH}_3)_3, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2\text{-SCH}_3, n=0$   
 8:  $\text{R}_1=3,4,5\text{-(OCH}_3)_3, \text{R}_2=\text{R}_3=\text{H}, \text{R}_4=2\text{-SCH}_3, n=1$   
 9:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-}, \text{R}_4=4\text{-N(CH}_3)_2, n=0$   
 10:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2\text{-R}_3=\text{CH}_2, \text{R}_4=4\text{-N(CH}_3)_2, n=1$   
 11:  $\text{R}_1=4\text{-OCH}_3, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-CH}_2\text{-}, \text{R}_4=3,4\text{-(-CH=CH-CH=CH-), } n=0$   
 12:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-CH}_2\text{-}, \text{R}_4=3,4,5\text{-(OCH}_3)_3, n=1$   
 13:  $\text{R}_1=4\text{-OCH}_3, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-CH}_2\text{-}, \text{R}_4=3,4,5\text{-(OCH}_3)_3, n=1$   
 14:  $\text{R}_1=3\text{-OCH}_3, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-}, \text{R}_4=3,4,5\text{-(OCH}_3)_3, n=1$   
 15:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-}, \text{R}_4=3,4,5\text{-(OCH}_3)_3, n=1$   
 16:  $\text{R}_1=3,4\text{-(OCH}_3)_2, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-}, \text{R}_4=3\text{-OCH}_3, 4\text{-OH}, n=1$   
 17:  $\text{R}_1=4\text{-OCH}_3, \text{R}_2\text{-R}_3=\text{-CH}_2\text{-}, \text{R}_4=2\text{-OCH}_3, n=1$   
 18:  $\text{Ar}=3,4,5\text{-(OCH}_3)_3\text{C}_6\text{H}_2$   
 19:  $\text{Ar}=\text{ferrocenyl}$

<sup>a</sup>Reagents and conditions: (i) (1) (1,3-dioxolan-2-yl-methyl)triphenylphosphonium bromide, NaH, 18-crown-6, THF, room temperature (2) 10% HCl, room temperature; (ii) 50% NaOH, EtOH, room temperature; (iii) (1) 50% NaOH, EtOH, room temperature, (2) 10% HCl, room temperature.

**Scheme 2.** Synthesis of (a) compounds **20-21**, (b) compound **22-23** and (c) compound **24<sup>a</sup>**





<sup>a</sup> Reagents and conditions: (i), (ii) 50% NaOH, EtOH, room temperature; (iii) DBU, MeOH, argon, room temperature.

**Structure Activity Relationships.** The whole data set covers a range of more than four order of magnitude difference in biological activity, with indanone derivative **10** possessing highest biological activity ( $IC_{50}$ : 0.04  $\mu$ M). Compound **9** and **10**, both containing an aniline nitrogen atom, only differ in one rotatable bond showed a difference of 2-fold in their biological activity. A similar trend has been observed for compounds **20-23** that are among the most active chalcone derivatives containing similar substitution pattern on both sides of central benzene ring of each compound, two carbonyl groups and a different number of rotatable bonds. Almost no difference in biological activity exists between compounds **20-22**. However, **23** ( $IC_{50}$ : 0.06 $\mu$ M), which possess two additional number of rotatable bonds, is about one order of magnitude more active than **20-22** ( $IC_{50}$ ; 0.14: 0.12: 0.14  $\mu$ M). We also synthesised indanone **14-17** and chalcone derivatives **3-8** and **11-13**, containing different numbers and pattern of methoxy groups. Almost no difference in biological activity of **15** and **16** has been identified. Overall, a minor difference ( $\sim$  factor of 1-2) exists in the indanones **14-17**. A similar trend has been observed for chalcone derivatives, which indicates that

a different number and pattern of methoxy groups on the indanone and chalcone scaffold and at the benzene ring on the opposite side of the molecules is not making any significant difference for biological activity. **18** and **19** are acidic analogs and comprise the least active compounds ( $IC_{50}$ : 141.30: 107.20),

**Table 1.** Biological activity ( $\mu\text{M}$ ),  $pIC_{50}$  values, hydrophobic surface area ( $V_{sa\_hyd}$ ) and the number of rotatable bonds ( $b\_rotN$ ) for compounds **3-24**.

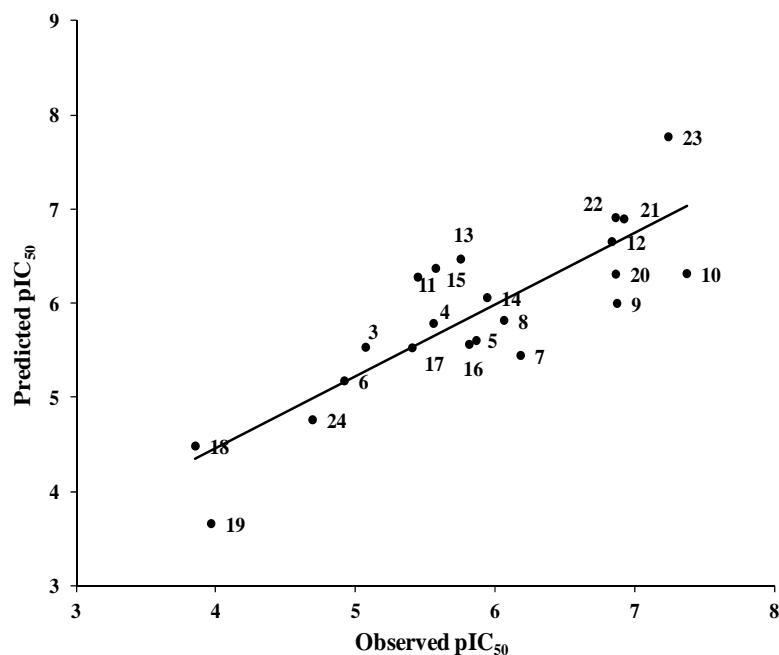
| Codes     | $IC_{50}$ ( $\mu\text{M}$ ) $\pm$ SD <sup>a</sup> | $pIC_{50}$ | $V_{sa\_hyd}$ | $b\_rotN$ |
|-----------|---|------------|---------------|-----------|
| <b>3</b>  | 8.51 $\pm$ 1.50                                   | 5.07       | 269.85        | 4.00      |
| <b>4</b>  | 2.75 $\pm$ 0.10                                   | 5.56       | 315.08        | 6.00      |
| <b>5</b>  | 1.38 $\pm$ 0.59                                   | 5.86       | 338.97        | 8.00      |
| <b>6</b>  | 12.0 $\pm$ 0.70                                   | 4.92       | 303.31        | 7.00      |
| <b>7</b>  | 0.65 $\pm$ 0.15                                   | 6.19       | 318.56        | 7.00      |
| <b>8</b>  | 0.87 $\pm$ 0.03                                   | 6.06       | 350.18        | 8.00      |
| <b>9</b>  | 0.13 $\pm$ 0.07                                   | 6.87       | 298.03        | 4.00      |
| <b>10</b> | 0.04 $\pm$ 0.01                                   | 7.37       | 329.65        | 5.00      |
| <b>11</b> | 3.55 $\pm$ 0.33                                   | 5.45       | 262.88        | 2.00      |
| <b>12</b> | 0.14 $\pm$ 0.04                                   | 6.84       | 371.69        | 7.00      |
| <b>13</b> | 1.74 $\pm$ 0.38                                   | 5.76       | 343.39        | 6.00      |
| <b>14</b> | 1.15 $\pm$ 0.03                                   | 5.94       | 327.58        | 6.00      |
| <b>15</b> | 2.63 $\pm$ 2.25                                   | 5.58       | 355.89        | 7.00      |
| <b>16</b> | 2.33 $\pm$ 0.34                                   | 5.63       | 291.92        | 5.00      |
| <b>17</b> | 3.89 $\pm$ 0.40                                   | 5.41       | 270.96        | 4.00      |
| <b>18</b> | 141.30 $\pm$ 5.60                                 | 3.85       | 265.44        | 7.00      |
| <b>19</b> | 107.20 $\pm$ 8.40                                 | 3.97       | 192.15        | 4.00      |
| <b>20</b> | 0.14 $\pm$ 0.06                                   | 6.86       | 406.62        | 10.00     |
| <b>21</b> | 0.12 $\pm$ 0.07                                   | 6.92       | 463.23        | 12.00     |
| <b>22</b> | 0.14 $\pm$ 0.02                                   | 6.86       | 463.24        | 12.00     |
| <b>23</b> | 0.06 $\pm$ 0.01                                   | 7.24       | 526.47        | 14.00     |
| <b>24</b> | 20.4 $\pm$ 0.20                                   | 4.69       | 252.94        | 5.00      |

<sup>a</sup> Each  $IC_{50}$  determination was performed with eight concentrations, and each assay point was determined in triplicate

**2D-QSAR.** In order to identify the most relevant physicochemical features important for high ABCB1 inhibitory activity of chalcone analogs multiple linear regression analysis was performed. Use of MOE's contingency<sup>15</sup> analysis tool for identification of the most important descriptors revealed the following equation (Equ. 1).

$$\text{Log}(1/IC_{50}) = 0.02 (vsa\_hyd) - 0.36 (b\_rotN) + 0.93 \quad (\text{Equ.1})$$

$$n = 22, R^2 = 0.79, q^2 (\text{LOO}) = 0.71, \text{RMSE} = 0.51$$



**Figure 1.** Plot of observed vs. predicted MDR-modulating activity of compounds 3-24, predicted values were obtained by leave-one-out cross validation.

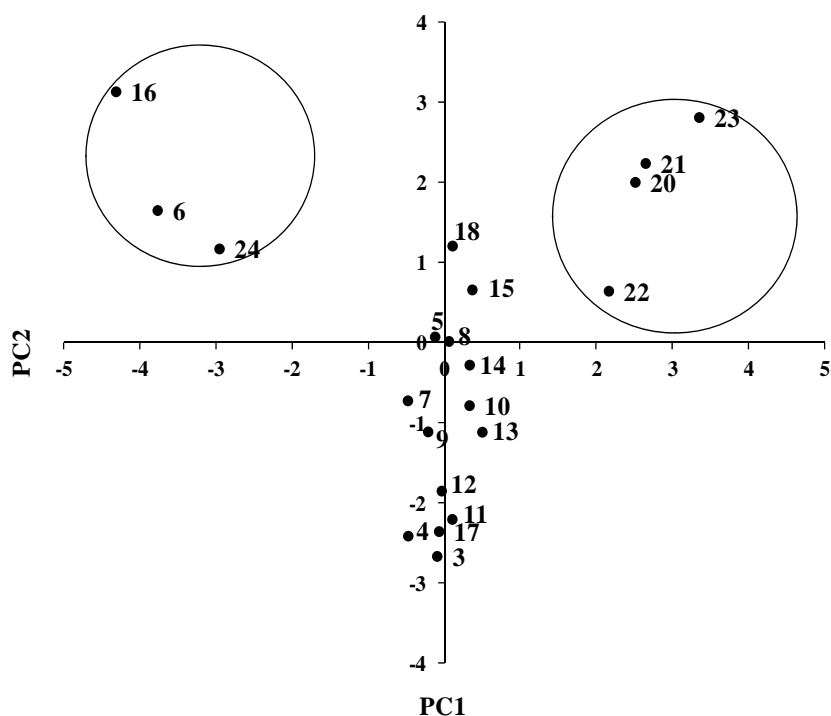
Figure 5 shows a plot of observed versus biological activity predicted by QSAR equation 1. An excellent QSAR model was obtained with all predicted values within one order of magnitude from the measured ones, no outlier was identified (residual value < one log unit). Descriptors contributing most to the variance in the biological activity comprised  $vsa\_hyd$  and  $b\_rotN$  (Equ. 1). This indicates that within this data set the number of rotatable bonds and the hydrophobic surface area are the most important structural attributes for high biological activity. This is in line with previous findings by Wang and colleagues, who showed that hydrophobic distribution within the molecules along with molecular weight, number of rotatable bonds and energy of highest occupied orbital  $E_{homo}$  are important descriptors for P-gp inhibitory potency.<sup>16</sup> However, in our case the QSAR equation reveals a negative contribution of the number of rotatable bonds, which points towards an unfavorable entropic contribution.

Extensive QSAR studies on a large set of propafenone analogs revealed the importance of hydrogen bond acceptors and their strength, the distance between aromatic moieties and H-bond acceptors as well as the influence of global physicochemical parameters, such as lipophilicity and molar refractivity.<sup>17-19</sup> However, for the present data set of chalcone derivatives we identified only a poor correlation ( $r^2 = 0.18$ ) between  $clogP$  and biological activity (Suppl Figure 1). This indicates that the

variance in the biological activity of chalcone derivatives is mainly driven by the concrete pattern of hydrophobicity distribution within the molecules, as represented by *vsa\_hyd*, rather than by their ability to penetrate in the membrane bilayer. This is also in line with the findings of Pleban *et al.*,<sup>20</sup> on the importance of the distribution of hydrophobicity within ABCB1 inhibitors. Later on this was further confirmed by König *et al.*, by using hydrophobic moments as QSAR descriptors.<sup>21</sup> Finally, this is additionally supported by the recent X-ray structure of mouse P-gp, which shows a large inner cavity exhibiting several hydrophobic patches for space directed hydrophobic interactions.<sup>22</sup>

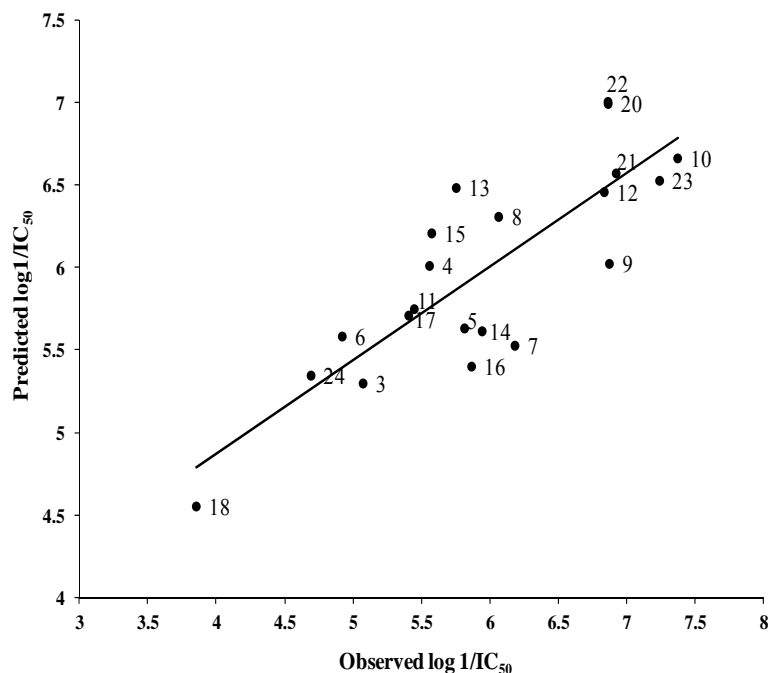
**3D QSAR.** Due to its rigid scaffold, chalcones represent versatile tools for 3D-QSAR studies. In recent years especially GRID-independent descriptors (GRIND) gained a lot of attraction in the field. GRIND descriptors are based on molecular interaction field (MIF) calculations and are alignment-independent, thus allowing the analysis of structurally diverse data series.<sup>23,23</sup> 3D conformations of the molecules in the data set were obtained from their 2D co-ordinates by using program CORINA.<sup>24</sup> GRIND descriptors were derived by computing molecular interaction fields (MIF) and by identifying the regions with maximum field intensity at relative distances by using the AMANDA algorithm implemented in software Pentacle version 1.06.<sup>25</sup> The Consistently Large Auto and Cross Correlation (CLACC) algorithm was used for encoding the prefiltered nodes into GRIND. The values obtained from the analysis were represented directly in correlograms plots, where the product of node- node energies versus distance separating the nodes is reported.

First, the data set was examined using principal component analysis (PCA). The first two components explained 43% of the variance in the GRIND descriptors and separate the data mainly on basis of their pharmacophoric pattern (Figure 2). Thus, compounds **6**, **16**, **24** in the upper left cluster share a similar pharmacophore having one hydrogen bond donor (OH). Compounds **20**, **21**, **22**, **23** which contain two carbonyl groups and show higher flexibility, are located in the upper right hand side of the plot. Interestingly, acid **18**, which exhibits pharmacophoric features from both clusters (two C=O, one OH), is located almost in the middle of these two clusters. Compounds with negative PC values are relatively small (mostly indanones and tetralones) and contain one carbonyl group, 1-5 methoxy groups, but no hydrogen bond donor.



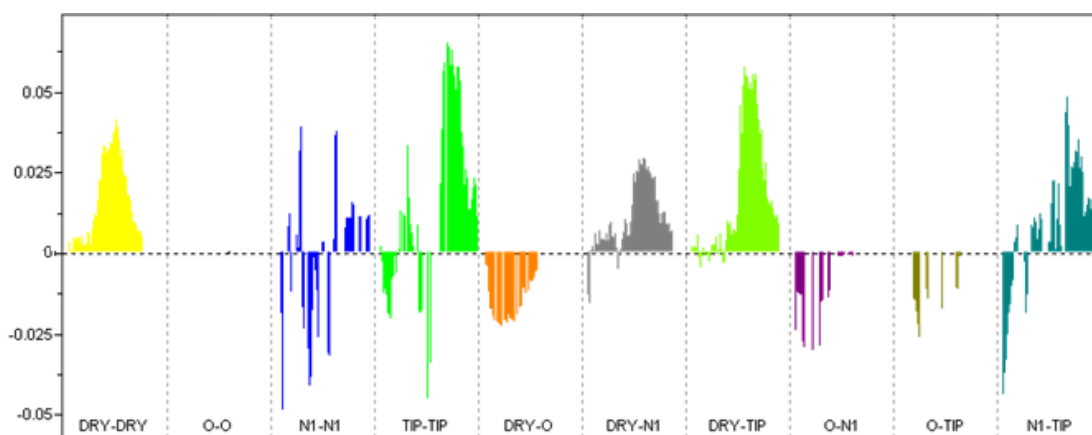
**Figure 2.** PCA score plot showing that the whole data set is divided into three groups; the upper right part contains flexible and large compounds, the lower portion contains less flexible and small compounds (mostly indanones and tetralones); the left part of the diagram contains all compounds having a hydrogen bond donor group.

In order to analyze the underlying important pharmacophoric patterns, PLS multivariate data analysis correlating the biological activity with the complete set of GRIND variables (790) was carried out using the AMANDA algorithm.<sup>31</sup> The PLS analysis resulted in a two-latent variable model with an  $r^2 = 0.85$ . However, the cross-validation of the model yielded  $q^2_{\text{LOO}}$  values of 0.26, which is quite unsatisfying. Thus, to reduce the high number of variables, a variable selection was applied by using FFD factorial selection implemented in Pentacle.<sup>25</sup> The resulting number of active variables decreased from 523 to 422 which relatively improved the quality of the model ( $r^2 = 0.98$ ,  $q^2_{\text{LOO}} = 0.66$ ). Figure 3 shows the plot of the experimental versus predicted biological activities.



**Figure 3.** Plot of observed vs. predicted MDR activity (expressed as  $\log(1/IC_{50})$  values) using the GRIND model

Analysis of the PLS coefficients profile of the 3<sup>rd</sup> Latent Variable (LV) of the PLS model illustrates the identification of key descriptors for high biological activity (Figure 4). Activity increases strongly with high value of the descriptors DRY-DRY, TIP-TIP, Dry-TIP, and N1-TIP (Table 2).



**Figure 4.** PLS Coefficient correlograms showing the descriptors which are directly (positive values) or inversely (negative values) correlated to the biological activity. The activity particularly increases with the increase in (DRY-DRY), (DRY-TIP) and (TIP-TIP) descriptor values.

**Table 2.** GRID- independent descriptors that are highly correlated to biological activity of chalcone derivatives **3-24**

| Variable | Distance      | Correlogram | Comment  |
|----------|---------------|-------------|--|
| 40       | 16.00-16.40 Å | DRY-DRY     | Represents two large hydrophobic groups, remains highly consistent throughout the length of the correlogram            |
| 292      | 22.00-22.40 Å | TIP-TIP     | Distance between two steric hot spots of the molecule  |
| 518      | 17.60-18.00 Å | DRY-TIP     | Distance of a hydrophobic group to one particular steric hot spot  |
| 715      | 1.60-2.00 Å   | N1-TIP      | Distance between a hydrogen bond acceptor and a steric hot spot , showing negative contribution to biological activity |
| 763      | 20.80-21.20 Å | N1-TIP      | Distance between a hydrogen bond acceptor and a steric hot spot, showing beneficial contribution to potency            |
| 442      | 18.80-19.20 Å | DRY-N1      | Distance between one of the two hydrophobic moieties to a hydrogen bond acceptor                                       |

The GRIND model indicates the presence of two hydrophobic moieties, which are localized at two of the three steric hot spots identified, molecular boundaries (Figure 5 a, b) in the most active ( $IC_{50} > 1 \mu M$ ) ABCB1 inhibitors. Most of the QSAR studies in the past two decades pointed towards the importance of hydrophobic substructures for high ABCB1 inhibitory potency.<sup>17-19</sup> Furthermore, by using a MIF-based pharmacophore model Broccatelli *et al*, recently provided also evidence for the importance of a distinct three dimensional shape of inhibitors of ABCB1.<sup>26</sup>

Also our model elucidates the importance of an optimal shape of the ligands and identifies three important steric hot spots (“edges”) A, B and C. In the molecules where hot spot (A) is 10.00-10.40 Å apart from (B), this represents two edges related to two methoxy substitutions at positions 6 and 7 of indanone and tetralone derivatives (Figure 5 b). However, in compounds **20-23**, which contain 3,4,5-tri methoxy groups, it represent the distance between the 3- and 4-annelated methoxy groups. The 3<sup>rd</sup> steric hot spot (C) represents the substituted benzene ring on the opposite side of the indanone and tetralone scaffold, which is at a distance of 22.00-22.40 Å from edge (A) in most of the

active ( $IC_{50} < 1\mu M$ ) compounds. This further confirms the importance of distinct 3D shape requirements for inhibitors of ABCB1.

It seems that out of three identified steric hot spots, (A) represent the most favorable one as it serve as an anchor to measure the distances to a hydrophobic feature (Figure 5c). Analyzing the most active compounds ( $IC_{50} < 1\mu M$ ) reveals the presence of a hydrophobic region around substituted benzene ring at the opposite side at a distance of 17.60-18.00 Å from edge (A). Optimal shape and hydrophobicity was also identified in other studies as major physicochemical parameters responsible for high affinity of flavonoid derivatives.<sup>27,28,29</sup> Furthermore, Cianchetta and co-workers identified the same features at a distance of 20.5 Å apart from each other in selected substrates of ABCB1.<sup>30</sup>

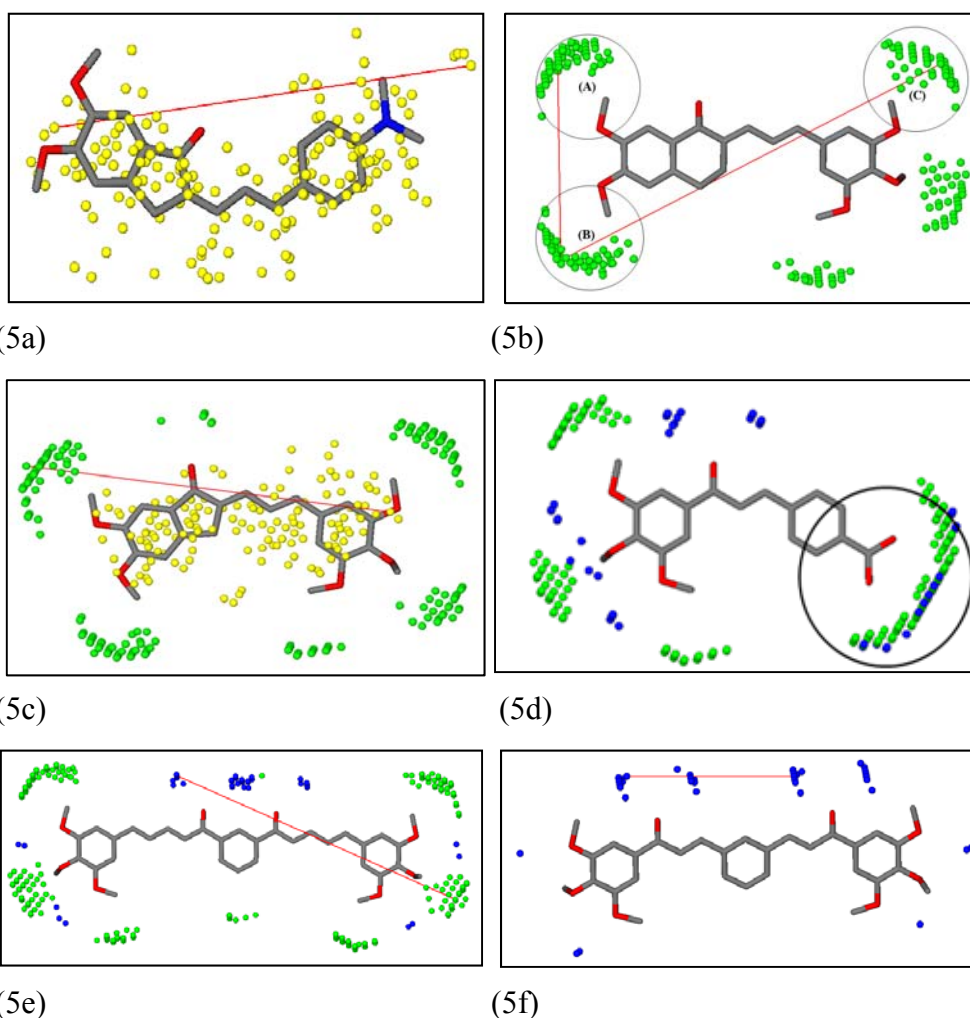
In order to further explore the hydrogen bonding related properties, a distance matrix of hydrogen bond acceptors from all three steric hot spots as well as their mutual distances were computed by GRIND descriptors. In some of the compounds edge (A) again represents an anchor point for corresponding distance calculations. Interestingly, two different distances ranges having opposite behavior have been identified. First, compounds with low activity values show a hydrogen bond acceptor at a distance of 1.60-2.00 Å from a steric hot spot. (Figure 5d). In contrast, both features far apart (20.80-21.20 Å) from each other is seen in the most active compounds ( $IC_{50} < 1\mu M$ ) (Figure 5e). This indicates that potent ABCB1 inhibitors show an elongated structure and have a hydrogen bond acceptor far from the edges of the molecules.

The number and pattern of H-bond acceptor groups is a subject of various publications. Seelig defined two patterns of H-bond acceptors and proposed that P-gp ligands may contain two or more H-bond acceptors which are separated either  $2.5 \pm 0.3$  Å and/or  $4.6 \pm 0.6$  Å apart from each other.<sup>31,32</sup> Interestingly, no consistency has been observed in the distance profile between different pairs of H-bond acceptors in chalcone derivatives. A distance of 9.20-9.60 Å has been identified between two carbonyl groups of **20-23** and between one carbonyl group and a region between the 6- and 7-methoxy group of the highly active indanone derivative **10** ( $IC_{50}$ :  $0.04\mu M$ ) (Figure 4f). However, this distance is not consistently present in all compounds having  $IC_{50}$  values  $< 1\mu M$ . Interestingly, a similar distance (8.80-9.20Å) has been found to be important for activity of conformationally rigid benzopyrano[3,4b][1,4]oxazine-type inhibitors of ABCB1.<sup>33</sup> In another study, performed on flavonoid derivatives, a distance of 8.00 Å between two



H-bond acceptors has been linked to high P-gp inhibitory activity.<sup>34</sup> A slightly larger distance (11.5-15 Å) has been identified by Cianchetta and co-workers,<sup>30</sup> for substrates of ABCB1. Finally, a similar distance range between two H-bond acceptors has been proposed by Pajeva *et al*, for a diverse data set of P-gp substrates/inhibitors, which are supposed to interact with the verapamil binding site.<sup>35</sup>

Although some similarities in mutual distances between two H-bond acceptors seem to exist, there is still an inconsistent picture and no clear threshold for efficient separation of more potent ABCB1 inhibitors from least active ones. This most probably reflects the notion that the large binding site in P-gp offers numerous possibilities to contribute to hydrogen bond driven interactions and thus allows a series of distinct, but different binding modes.



**Figure 5.** (a) DRY-DRY hot spots (yellow color) which represents two hydrophobic regions 16.00-16.40 Å apart, present in most active compounds. (b) Shows steric hot spots (green color) which makes three important boundaries (A, B and C) for most

potent inhibitors of ABCB1 where A-B: 10.00-10.40 Å and A-C: 22.00-22.40 Å. (c) Represent the distance range of a hydrophobic substructure (yellow hot spot) from molecular extreme (A) (17.60-18.00 Å) (green hot spot). (d) Shows a carboxylic acid group (N1: blue hot spot) present very close (1.60-2.00 Å) to one of the molecular boundary (encircle), having negative effect to biological activity, while (e) represent the same pharmacophoric features at longer distance (20.80-21.20 Å) showing positive contribution towards biological activity. (f) N1-N1 hot spots (blue color) representing two H-bond acceptors at a distance of 9.20-9.60 Å which is favorable for biological activity of most of the compounds.

### Conclusion

With the current study we present a series of novel chalcone derivatives which in part show ABCB1 inhibitory activity in the nanomolar range. Based on a set of 22 compounds covering different modifications of the chalcone scaffold, two predictive QSAR models were established in order to elucidate the molecular features responsible for high biological activity. 2D-QSAR analysis revealed the importance of the hydrophobic surface area and the number of rotatable bonds. Interestingly, in contrast to several other compound classes, there was only a poor correlation between overall lipophilicity of the compounds and their ABCB1 inhibitory activity. This indicates that for chalcones hydrophobic areas directly contribute to ligand binding. This is further exemplified by the GRIND analysis, which identified three hydrophobic hot spots in the molecules. Furthermore, distinct distances between these hydrophobic features and H-bond acceptors have been exemplified. Remarkably, these compounds do not contain a basic nitrogen atom. Furthermore, they exhibit a quite rigid and planar structure, which renders them quite unique in the chemical space of ABCB1 inhibitors. Thus, chalcones represent an interesting new class of ABCB1 ligands which will deserve further investigation.

### Experimental Section

**Chemistry.** Unless otherwise stated, all chemicals were obtained from Sigma-Aldrich or TCI Europe and were of analytical grade. Melting points were determined on a Kofler hot stage apparatus and are uncorrected. The  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker Avance DPx200 (200 and 50 MHz). Chemical shifts are reported in  $\delta$  units (ppm) relative to  $\text{Me}_4\text{Si}$  line as internal standard and  $J$  values are reported in Hertz. Mass spectra were obtained by a Hewlett Packard (GC: 5890; MS: 5970)

spectrometer. The purity of the synthesized compounds was established by combustion analysis with a Perkin-Elmer 2400 CHN elemental analyzer and was within  $\pm 0.4\%$ . Solutions in organic solvents were dried over anhydrous sodium sulphate.

**General Synthesis Procedure for Compounds 1 and 2.** To a suspension of 5 mmol of the corresponding benzaldehyde, 7.5 mmol (3.220 g) (1,3-dioxolan-2-ylmethyl)triphenylphosphonium bromide and 0.005 g 18-crown-6 in anhydrous THF and under argon atmosphere 20.8 mmol (0.499 g) NaH were added carefully. The reaction mixture was stirred till the reaction was completed (monitoring by TLC). Then, the mixture was cooled to  $0^{\circ}\text{C}$  and first water and then 10% HCl were carefully added. After 60 minutes stirring at room temperature, the mixture was extracted with ethyl acetate, 10% HCl and water. The combined organic phase was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was removed *in vacuo*. The so-obtained crude product was purified by flash chromatography.

**(E)-2-Methylthiocinnamaldehyde (1) and 3,4,5-Trimethoxycinnamaldehyde (2).**<sup>36</sup> Detailed description of the compounds is available in the supplementary data.

**General Synthesis Procedure for Compounds 3-23.** A solution of the 2.5 mmol of the appropriate acetophenone, indanone, tetralone derivative or 1,3-diacetylbenzene and 2 mL 50% NaOH in 10 mL ethanol was stirred at room temperature for 30 minutes. Then, 2.5 mmol (or 5 mmol with 1,3-diacetylbenzene) of the corresponding benzaldehyde or cinnamaldehyde derivative, dissolved in 1 mL ethanol, were added and stirred at room temperature. After conversion of the starting compounds was completed as monitored by TLC, the reaction mixture was poured into ice water and acidified with 10% HCl to pH 6. The so-formed solid was filtered off and the crude product was further purified by recrystallization in ethanol.

**(E)-3-(2,4,6-Trimethylphenyl)-1-(2'-methoxyphenyl)-2-propen-1-one (3).**<sup>37</sup> Detailed description of the compounds is available in the supplementary data.

**(E)-1-(3,4-Dimethoxyphenyl)-3-(4-methoxy-1-naphthyl)-2-propen-1-one (4).** Yield: 0.174 g (20%) yellow solid; mp:  $112-116^{\circ}\text{C}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  8.62 (AB-system,  $J=15.3$ , 1H), 8.38-8.21 (m, 2H), 7.89 (d,  $J=8.2\text{Hz}$ , 1H), 7.77-7.49 (m, 5H), 7.00-6.84 (m, 2H), 4.05 (s, 3H), 3.99 (s, 3H), 3.97 (s, 3H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  188.6, 157.6, 153.1, 149.2, 141.0, 132.8, 131.6, 127.5, 126.0, 125.6, 124.8, 123.2, 122.9, 122.6, 121.9, 110.8, 110.0, 103.7, 56.1, 56.0, 55.7. MS  $m/z$ : 348 (72%,  $\text{M}^+$ ), 165

(100%), 158 (40%), 139 (91%), 79 (55%). Anal. Calcd for  $C_{22}H_{20}O_4 \cdot xH_2O$ : C 75.45; H 5.81. Found: C 75.42; H 5.90.

**(2E,4E)-5-(2-Methoxyphenyl)-1-(3',4',5'-trimethoxyphenyl)-2,4-pentadien-1-one (5).** Yield: 1.240 g (70%) of yellow crystals; mp: 111 °C.  $^1H$  NMR ( $CDCl_3$ ):  $\delta$  7.72-7.53 (m, 2H), 7.52-7.23 (m, 4H), 7.17-6.89 (m, 4H), 3.95 (s, 6H), 3.93 (s, 3H), 3.90 (s, 3H).  $^{13}C$  NMR ( $CDCl_3$ ):  $\delta$  189.2, 157.6, 153.0, 145.9, 142.1, 137.3, 133.7, 130.4, 127.4, 125.0, 124.3, 120.7, 111.1, 105.8, 60.9, 56.3, 55.5. MS  $m/z$ : 354 (100%,  $M^+$ ), 339 (73%), 195 (28%), 140 (25%), 91 (54%). Anal. Calcd for  $C_{21}H_{22}O_5 \cdot 0.3H_2O$ : C 71.16; H 6.26. Found: C 70.91; H 6.47.

**5-(2-Methoxyphenyl)-1-(4'-hydroxy-3',5'-dimethoxyphenyl)-2,4-pentadien-1-one (6).** Yield: 0.240 g (28%) yellow solid; mp: 44-51 °C.  $^1H$  NMR ( $CDCl_3$ ):  $\delta$  7.78-6.87 (m, 10H), 5.99 (s, broad, 1H), 3.98 (s, 6H), 3.90 (s, 3H).  $^{13}C$  NMR ( $CDCl_3$ ):  $\delta$  188.6, 157.5, 146.8 (2C), 145.4, 139.3, 137.0, 130.3, 129.9, 127.5, 127.4, 125.1, 124.2, 120.7, 111.1, 105.7(2C), 56.5 (2C), 55.5. MS  $m/z$ : 340 (100%,  $M^+$ ), 309 (8%), 246 (6%), 181 (35%), 115 (63%). Anal. Calcd for  $C_{20}H_{20}O_5 \cdot 0.1H_2O$ : C 70.20; H 5.95. Found: C 70.13; H 5.71.

**(E)-3-(2-Methylthiophenyl)-1-(3',4',5'-trimethoxyphenyl)-2-propen-1-one (7).** Yield: 0.496 g (64%) yellow crystals; mp: 83-86 °C.  $^1H$  NMR ( $CDCl_3$ ):  $\delta$  8.22 (AB-system,  $J=15.6$  Hz, 1H), 7.71-7.59 (m, 1H), 7.48-7.27 (m, 6H), 3.95 (s, 6H), 3.94 (s, 3H), 2.50 (s, 3H).  $^{13}C$  NMR ( $CDCl_3$ ):  $\delta$  189.8, 153.1 (2C), 142.3, 141.7, 139.9, 134.1, 133.3, 130.4, 127.3, 127.0, 125.5, 124.3, 106.2 (2C), 60.9, 56.3, 16.5. MS  $m/z$ : 344 (1%,  $M^+$ ), 297 (100%), 236 (5%), 195 (13%), 149 (88%). Anal. Calcd for  $C_{19}H_{20}O_4S$ : C 66.26; H 5.85. Found: C 66.22; H 5.91.

**5-(2-Methylthiophenyl)-1-(3',4',5'-trimethoxyphenyl)-2,4-pentadiene-1-one (8).** Yield: 0.509 g (55%) yellow crystals. mp: 114.5-115.5 °C.  $^1H$  NMR ( $CDCl_3$ ):  $\delta$  7.76-7.42 (m, 3 H), 7.40-6.86 (m, 7 H), 3.95 (s, 6H), 3.94 (s, 3H), 2.48 (s, 3H).  $^{13}C$  NMR ( $CDCl_3$ ):  $\delta$  189.0, 153.1 (2C), 144.8, 142.3, 138.6, 138.3, 135.1, 133.4, 129.4, 128.4, 127.3, 126.2, 125.6, 125.2, 105.8 (2C), 60.9, 56.3 (2C), 16.5. MS:  $m/z$  370 ( $M^+$ , 65%), 355 (46%), 297 (60%), 187 (100%), 149 (90%). Anal. Calcd for  $C_{21}H_{22}O_4S$ : C 68.08; H 5.99. Found: C 67.74; H 6.01.

**2-[1-(4-Dimethylaminophenyl)-methylidene]-5,6-dimethoxyindan-1-one (9).** Yield: 0.457 g (57%) yellow crystals. mp: 196-203 °C.  $^1H$  NMR ( $CDCl_3$ ):  $\delta$  7.68-7.49

(m, 3H), 7.34 (s, 1H), 6.97 (s, 1H), 6.73 (d,  $J=9.0$  Hz, 2H), 3.99 (s, 3H), 3.95 (s, 3H), 3.90 (s, 2H), 3.04 (s, 6H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  193.3, 154.7, 150.9, 149.4, 144.3, 133.3, 132.4 (2C), 131.7, 130.6, 123.4, 111.9 (2C), 107.2, 104.9, 56.2, 56.1, 40.7 (2C), 32.4. MS:  $m/z$  323 ( $\text{M}^+$ , 100%), 280 (13%), 237 (14%), 161 (15%), 121 (20%). Anal. Calcd for  $\text{C}_{20}\text{H}_{21}\text{NO}_3$ : C 74.28; H 6.55; N 4.33. Found: C 74.07; H 6.50; N 4.25.

**2-[3-(4-Dimethylaminophenyl)-prop-2-en-yliden]-5,6-dimethoxyindan-1-one**

**(10).** Yield: 0.422 g (48%) orange crystals. mp: 210-213°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.50-7.29 (m, 4H), 7.02-6.61 (m, 5H), 3.99 (s, 3H), 3.93 (s, 3H), 3.72 (s, 2H), 3.01 (s, 6H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  192.5, 154.8, 150.9, 149.4, 143.8, 142.2, 134.1, 133.3, 132.7, 128.7 (2C), 124.6, 119.9, 112.0 (2C), 107.2, 104.8, 56.2, 56.1, 40.2 (2C), 30.2. MS:  $m/z$  349 ( $\text{M}^+$ , 100%), 334 (17%), 175 (19%), 144 (25%), 117 (27%). Anal. Calcd for  $\text{C}_{22}\text{H}_{23}\text{NO}_3$ : C 75.62; H 6.64; N 4.01. Found: C 75.35; H 6.60; N 3.96.

**6-Methoxy-2-[1-naphthalen-2-yl-methylidene]-3,4-dihydro-2H-naphthalen-1-one (11).**<sup>38</sup> Detailed description of the compounds is available in the supplementary data.

**6,7-Dimethoxy-2-(3-(3,4,5-trimethoxyphenyl)-2-propenylidene)-1-tetralone (12).**

Yield: 0.282 g (27%) yellow solid. mp: 89-94°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.62 (s, 1H), 7.51 (AB-system,  $J=10.4$  Hz, 1H) 7.17-6.87 (m, 2H), 6.82-6.64 (m, 3H), 3.96 (s, 3H), 3.95 (s, 3H), 3.88 (s, 6H), 3.16-2.88 (m, 4H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  186.1, 153.4, 153.3, 148.2, 140.4, 138.3, 135.1, 134.1, 132.4, 126.9, 122.9, 109.9, 109.5, 104.2 (2C), 61.0, 56.2 (2C), 56.0 (2C), 28.5, 26.3. MS:  $m/z$  410 ( $\text{M}^+$ , 100%), 395 (37%), 379 (40%), 231 (35%), 175 (41%). Anal. Calcd for  $\text{C}_{24}\text{H}_{26}\text{O}_6$ : C 70.23; H 6.38. Found: C 70.01; H 6.39.

**6-Methoxy-2-(3-(3,4,5-trimethoxyphenyl)-2-propenylidene)-1-tetralone (13).**

Yield: 0.399 g (42%) yellow crystals. mp: 152-156°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  8.10 (d,  $J=8.7$  Hz, 1H), 7.51 (AB-system,  $J=10.2$  Hz, 1H), 7.15-6.84 (m, 3H), 6.78-6.68 (m, 3H), 3.92 (s, 6H), 3.88 (s, 3H), 3.87 (s, 3H), 3.00 (s, 4H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  186.2, 163.4, 153.4, 145.8, 140.5, 135.2, 134.3, 132.4, 130.6, 127.3, 122.9, 113.2, 112.4, 104.2 (2C), 61.0, 56.2 (2C), 55.4, 29.2, 26.1. MS:  $m/z$  380 ( $\text{M}^+$ , 100%), 365 (46%), 349 (47%), 161 (69%), 115 (33%). Anal. Calcd for  $\text{C}_{23}\text{H}_{24}\text{O}_5$ : C 72.61; H 6.36. Found: C 72.36; H 6.31.

**6-Methoxy-2-(3-(3,4,5-trimethoxyphenyl)-2-propenylidene)-1-indanone (14).**

Yield: 0.445 g (49%) yellow crystals. mp: 182-186°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.51-7.29 (m, 3H), 7.19 (dd,  $J=8.3$  Hz,  $J=2.5$  Hz, 1H), 7.02-6.82 (m, 2H), 6.74 (s, 2H), 3.93 (s,

6H), 3.89 (s, 3H), 3.86 (s, 3H), 3.84-3.78 (m, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  193.5, 159.5, 153.4, 141.9, 141.6, 140.5, 136.7, 133.2, 131.9, 126.9, 123.7, 123.7, 105.6, 104.4 (2C), 61.0, 56.2 (2C), 55.6, 29.8. MS:  $m/z$  366 ( $\text{M}^+$ , 100%), 351 (44%), 335 (46%), 165 (24%), 82 (24%). Anal. Calcd for  $\text{C}_{22}\text{H}_{22}\text{O}_5$ : C 72.12; H 6.05. Found: C 71.85; H 6.04.

**6,7-Dimethoxy-2-(3-(3,4,5-trimethoxyphenyl)-2-propenylidene)-1-indanone (15).** Yield: 0.709g (72%) yellow crystals. mp: 193-195°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.39-7.29 (m, 2H), 6.99-6.87 (m, 3H), 6.73 (s, 2H), 3.99 (s, 3H), 3.93 (s, 9H), 3.89 (s, 3H), 3.78 (s, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  182.4, 155.2, 153.4, 149.5, 144.0, 141.1, 136.5, 132.4, 132.1, 131.7, 123.8, 107.2, 104.9, 104.3 (2C), 61.0, 56.2, 56.2 (2C), 56.1, 30.1. MS:  $m/z$  396 ( $\text{M}^+$ , 100%), 381 (36%), 365 (52%), 175 (22%), 165 (25%). Anal. Calcd for  $\text{C}_{23}\text{H}_{24}\text{O}_6$ : C 69.68; H 6.10. Found: C 69.44; H 6.05.

**6,7-Dimethoxy-2-(3-(4-hydroxy-3-methoxyphenyl)-2-propenylidene)-1-indanone (16).** Yield: 0.559g (64%) yellow crystals. mp: 114-120°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.41-7.28 (m, 2H), 7.19-7.05 (m, 1H), 7.03-6.69 (m, 5H), 5.87 (s, 1H), 3.99 (s, 3H), 3.96 (s, 3H), 3.94 (s, 3H), 3.76 (s, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  192.5, 155.1, 149.5, 146.9, 146.7, 144.0, 141.4, 135.6, 132.4, 132.4, 129.1, 122.2, 121.3, 114.8, 109.1, 107.2, 104.9, 56.2, 56.1, 55.9, 30.1. MS:  $m/z$  352 ( $\text{M}^+$ , 100%), 337 (34%), 335 (25%), 176 (16%), 101 (21%). Anal. Calcd for  $\text{C}_{21}\text{H}_{20}\text{O}_5 \times 0.2\text{H}_2\text{O}$ : C 70.85; H 5.78. Found: C 70.67; H 5.71.

**5-Methoxy-2-(3-(2-methoxyphenyl)-2-propenylidene)-1-indanone (17).** Yield: 0.490 g (64%) orange crystals. mp: 151-154°C.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.80 (d,  $J=8.5\text{Hz}$ , 1H), 7.56 (dd,  $J=7.8\text{Hz}$ ,  $J=1.5\text{Hz}$ , 1H), 7.44-7.25 (m, 3H), 7.15-6.85 (m, 5H), 3.89 (s, 6H), 3.78 (s, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  192.2, 164.9, 157.4, 151.8, 136.5, 135.9, 133.1, 131.8, 130.2, 127.3, 125.8, 125.4, 124.9, 120.7, 115.0, 111.1, 109.7, 55.6, 55.5, 30.5. MS:  $m/z$  306 ( $\text{M}^+$ , 100%), 199 (37%), 145 (30%), 115 (26%), 101 (22%). Anal. Calcd for  $\text{C}_{20}\text{H}_{18}\text{O}_3$ : C 78.41; H 5.29. Found: C 78.14; H 5.98.

**4-[(E)-3-Oxo-3-(3',4',5'-trimethoxyphenyl)-1-propenyl]benzoic acid (18).** Yield: 0.445 g (52 %) yellow crystals. mp: 223°C.  $^1\text{H}$  NMR ( $d_6$ -DMSO):  $\delta$  13.1 (s, 1H), 8.08 (AB-system  $J=15.7\text{ Hz}$ , 1H), 8.01 (s, 4H), 7.78 (AB-system  $J=15.7\text{ Hz}$ , 1H), 7.45 (s, 2H), 3.90 (s, 6H), 3.77 (s, 3H).  $^{13}\text{C}$  NMR ( $d_6$ -DMSO):  $\delta$  187.3, 166.4, 152.5, 142.0, 141.7, 138.3, 132.2, 131.6, 129.2, 128.5, 123.5, 105.8, 59.7, 55.7. MS:  $m/z$  342 ( $\text{M}^+$ , 100%), 327 (29%), 299 (24%), 195 (47%), 103 (26%). Anal. Calcd for  $\text{C}_{19}\text{H}_{18}\text{O}_6$ : C 66.66; H 5.30. Found: C 66.52; H 5.28.

**4-[(E)-3-Oxo-3-ferrocenyl-1-propenyl]benzoic acid (19).** Yield: 0.405 g (45 %) red crystals. mp: >350°C. <sup>1</sup>H NMR (*d*<sub>6</sub>-DMSO): δ 13.12 (s, broad, 1H), 8.30-7.82 (m, 4H), 7.67 (AB-system, *J*=15.6Hz, 1H), 7.53 (AB-system, *J*=15.6Hz, 1H), 5.07 (s, 2H), 4.68 (s, 2H), 4.22 (s, 5H). <sup>13</sup>C NMR (*d*<sub>6</sub>-DMSO): δ 192.1, 158.4, 139.3, 138.6 (2C), 131.9, 128.9, 126.0 (2C), 80.7, 73.2 (2C), 70.0 (5C), 69.9 (2C). MS: *m/z* 360 (M<sup>+</sup>, 100%), 165 (76%), 121 (33%), 102 (15%), 56 (28%). Anal. Calcd for C<sub>20</sub>H<sub>16</sub>O<sub>3</sub>Fe: C 66.69; H 4.48. Found: C 66.43; H 4.23.

**1-(3,4-Dimethoxyphenyl)-3-[3-(3,4-dimethoxyphenyl)-3-oxo-1-propenyl]phenyl)-2-propen-1-one (20).** Yield: 0.161 g (14%) pale yellow solid. mp: 121-125°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 7.95-7.79 (m, 3H), 7.76-7.57 (m, 8H), 7.56-7.44 (m, 1H), 6.95 (d, *J*=8.3Hz, 2H), 3.99 (s, 12H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 188.2 (2C), 153.4 (2C), 149.3 (2C), 143.0 (2C), 135.8 (2C), 131.1 (2C), 129.8 (2C), 129.5, 128.3, 123.1 (2C), 122.5 (2C), 110.7 (2C), 110.0 (2C), 56.1 (2C), 56.1 (2C). MS: *m/z* 458 (M<sup>+</sup>, 54%), 293 (51%), 165 (100%), 137 (17%), 77 (26%). Anal. Calcd for C<sub>28</sub>H<sub>26</sub>O<sub>6</sub>·0.5H<sub>2</sub>O: C 71.93; H 5.82. Found: C 71.82; H 6.05.

**3-[3-[3-Oxo-3-(3,4,5-trimethoxyphenyl)-1-propenyl]phenyl]-1-(3,4,5-trimethoxyphenyl)-2-propen-1-one (21).** Yield: 0.519 g (40%) white solid. mp: 152-154°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 7.94-7.80 (m, 3H), 7.72 (dd, *J*=7.6Hz, *J*=1.3Hz, 2H), 7.61-7.47 (m, 3H), 7.30 (s, 3H), 3.97 (s, 12H), 3.95 (s, 6H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 188.9 (2C), 153.2 (4C), 143.6 (2C), 135.7 (4C), 133.2 (2C), 129.8 (2C), 129.6, 128.7, 122.6 (2C), 106.1 (4C), 61.0 (2C), 56.4 (4C). MS: *m/z* 518 (M<sup>+</sup>, 100%), 323 (36%), 195 (95%), 152 (24%), 136 (23%). Anal. Calcd for C<sub>30</sub>H<sub>30</sub>O<sub>8</sub>: C 69.49; H 5.83. Found: C 69.19; H 5.85.

**(E)-3-(3,4,5-Trimethoxyphenyl)-1-{3-[(E)-3-(3,4,5-trimethoxyphenyl)-2-propenoyl]phenyl}-2-propen-1-one (22).** Yield: 0.052 g (4%) pale yellow solid. mp: 145-148°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 8.63 (s, 1H), 8.23 (dd, *J*=7.7Hz, *J*=1.6Hz, 2H), 7.78 (AB-system, *J*=15.5Hz, 2H), 7.67 (t, *J*=7.7Hz, 2H), 7.46 (AB-system, *J*=15.5Hz, 2H), 6.90 (s, 4H), 3.94 (s, 12H), 3.92 (s, 6H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 189.8 (2C), 153.5 (4C), 146.0 (2C), 138.7 (2C), 132.3 (2C), 130.0 (2C), 129.0, 128.2, 120.8 (2C), 105.8 (4C), 61.0 (2C), 56.1 (4C). MS: *m/z* 518 (M<sup>+</sup>, 100%), 487 (38%), 221 (33%), 206 (20%), 193 (15%). Anal. Calcd for C<sub>30</sub>H<sub>30</sub>O<sub>8</sub>·0.1H<sub>2</sub>O: C 69.25; H 5.85. Found: C 69.06; H 5.92.

**5-(3,4,5-Trimethoxyphenyl)-1-{3-[5-(3,4,5-trimethoxyphenyl)-2,4-pentadienoyl]phenyl}-2,4-pentadien-1-one (23).** Yield: 0.271 g (19%) yellow solid. mp: 142-145°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 8.57 (s, 1H), 8.18 (dd, *J*=7.8Hz, *J*=1.5Hz, 2H), 7.78-7.54 (m, 3H), 7.17 (AB-system, *J*=14.8Hz, 2H), 7.06-6.91 (m, 4H), 6.75 (s, 4H), 3.92 (s, 12H), 3.89 (s, 6H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 189.5 (2C), 153.4 (4C), 145.4 (2C), 142.5 (2C), 139.5 (2C), 138.5 (2C), 132.2 (2C), 131.6 (2C), 129.0, 128.1, 126.3 (2C), 124.6 (2C), 104.5 (4C), 61.0 (2C), 56.1 (4C). MS: *m/z* 570 (M<sup>+</sup>, 4%), 218 (100%), 204 (19%), 188 (55%), 117 (27%). Anal. Calcd for C<sub>34</sub>H<sub>34</sub>O<sub>8</sub>·0.3H<sub>2</sub>O: C 70.89; H 6.08. Found: C 70.80; H 6.00.

**Methyl 2-hydroxy-3-[3-(2-methoxyphenyl)-prop-2-enylidene]cyclohexene-1-carboxylate (24).** The compound was synthesized according to the procedure described previously:<sup>14</sup> to a solution of 10 mmol (1.702 g, 1.6 mL) ethyl 2-cyclohexanone carboxylate and 10 mmol (1.522 g, 1.5 ml) diazabicycloundecene (DBU) in 20 mL dry methanol 10 mmol (1.622 g) 2-methoxycinnamaldehyde, dissolved in 10 mL dry methanol, added. The reaction mixture stirred under argon atmosphere at room temperature for 40 hours. Then, it was cooled to 0°C, the obtained precipitate was filtered off and recrystallized in ethanol. Yield: 0.599g (20%) yellow crystals. mp: 125-128°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 12.22 (s, 1H), 7.52 (dd, *J*=7.7Hz, *J*=1.5Hz, 1H), 7.36-7.09 (m, 4H), 7.05-6.81 (m, 2H), 3.87 (s, 3H), 3.79 (s, 3H), 2.70-2.52 (m, 2H), 2.50-2.35 (m, 2H), 1.88-1.67 (m, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 173.2, 165.1, 157.1, 132.1, 130.4, 130.0, 129.2, 126.9, 126.2, 124.7, 120.7, 111.0, 99.6, 55.5, 51.6, 25.7, 23.2, 22.1. MS: *m/z* 300 (M<sup>+</sup>, 57%), 268 (100%), 209 (13%), 134 (60%), 91 (47%). Anal. Calcd for C<sub>18</sub>H<sub>20</sub>O<sub>4</sub>: C 71.98; H 6.71. Found: C 71.96; H 6.72.

### Biology

**Cell lines.** The resistant CCRF vcr1000 cell line was maintained in RPMI 1640 medium containing 10% fetal calf serum (FCS) and 1000ng/mL vincristine. The selecting agent was washed out 1 week before the experiments. This cell line was selected due to its distinct P-gp expression.

**Inhibition of daunorubicin efflux.** IC<sub>50</sub> values for daunorubicin efflux inhibition were determined as reported (Chiba *et al*, 1996). Briefly, cells were sedimented, the supernatant was removed by aspiration, and the cells were resuspended at a density of 1



$\times 10^6$ /mL in RPMI 1640 medium containing daunorubicin (Sigma Chemical Co., St. Louis, MO) at a final concentration of 3  $\mu$ mol/l. Cell suspensions were incubated at 37°C for 30 min. Tubes were chilled on ice and centrifuged at 500 g in an Eppendorf 5403 centrifuge (Eppendorf, Hamburg, Germany). Supernatants were removed, and the cell pellet was resuspended in medium pre-warmed to 37°C containing either no inhibitor or compounds at various concentrations ranging from 20  $\mu$ M to 200  $\mu$ M, depending on the solubility and expected potency of the inhibitor. Eight concentrations (serial 1:3 dilution) were tested for each inhibitor. After 60, 120, 180 and 240 seconds, aliquots of the incubation mixture were transferred to tubes containing an equal volume of ice-cold stop solution (RPMI medium containing GPV31 at a final concentration of 5  $\mu$ mol/l). Zero time points were determined by immediately pipetting daunorubicin-preloaded cells into ice cold stop solution. Samples drawn at the respective time points were kept in an ice water bath and measured within 1h on a Becton Dickinson FACSCalibur flow cytometer (Becton Dickinson, Vienna, Austria). Viable cells were selected by setting appropriate gates for forward and side scatter. The excitation and emission wavelengths were 482 nm and 558 nm, respectively. Five thousand gated events were accumulated for the determination of mean fluorescence values.

**Mathematical model for determination of inhibitor IC<sub>50</sub> values.** Initial efflux rates were calculated from the time dependent linear decrease in mean fluorescence through the use of linear regression analysis. IC<sub>50</sub> values were determined from dose response curves of inhibitor concentration versus initial efflux rates. Data points were fitted according to the equation  $V = V_0 - (V_0 - V_{inf}) \times [I]^n \div (IC_{50}^n + [I]^n)$  where V is the velocity of transport, V<sub>0</sub> is the transport velocity in the absence of inhibitor, V<sub>inf</sub> is the efflux velocity at infinite concentration of inhibitor (which is equal to simple diffusion), [I] is the inhibitor concentration, IC<sub>50</sub> is the 50% inhibitory concentration (50% occupancy value) and n is the Hill coefficient.

### **Computational Models**

**GRID-Independent Molecular Descriptor Analysis.** Molecular discovery software Pentacle version 1.06 was used for computing alignment-independent 3D-descriptors.<sup>39</sup> This so called GRIND approach aims to extract the information enclosed in the molecular interaction fields (MIFs) and compress it into new types of variables whose values are independent of the spatial position of the molecule studied. Most relevant

regions are extracted from the MIF by an optimization algorithm that uses the intensity of the field at a node and the mutual node-node distances between the chosen nodes as a scoring function. At each point, the interaction energy ( $E_{xyz}$ ) is calculated as a sum of Lennard-Jones energy ( $E_{lj}$ ), Hydrogen bond ( $E_{hb}$ ) and Electrostatic ( $E_{al}$ ) interactions.

$$E_{xyz} = \sum E_{lj} + \sum E_{al} + \sum E_{hb}$$

Default values for Grid Step (0.5 Å) and probes (DRY representing Hydrophobic interaction, O (Carbonyl Oxygen) representing hydrogen bond acceptor groups, N1 (Amide Nitrogen) representing H-bond donor groups and TIP representing a shape descriptor) were used for computation of the MIF. MIF discretization was performed by the AMANDA algorithm using default values for probe cutoff (DRY= -0.5, O= -2.6, N1= -4.2, TIP= -0.74).<sup>40</sup> Nodes with an energy value below this cutoff were discarded. Large Auto and Cross Correlation (CLACC) algorithm was used for encoding the prefiltered nodes into GRIND thus producing most consistent variables as compared to MACC.<sup>41</sup>

**QSAR.** Molecular structures were built with the builder function of MOE<sup>42</sup> version 2010 and energy minimised, partial charges were assigned by MMFF94 force field. 2D molecular descriptors, including atom and bond counts, connectivity indices, partial charge descriptors, pharmacophore feature descriptors and general physicochemical descriptors were calculated by using software MOE version 2010. PLS analysis was performed with the MOE QSAR model tool and the predictive ability of the models was determined by leave one out cross validation (LOO).

**Acknowledgment.** Ishrat Jabeen acknowledges financial support provided by the Higher Education Commission of Pakistan (HEC). GE, ZP and Pc acknowledge financial support provided by the Austrian Science Fund (SFB 35)

## References

1. Loo, T. W.; Clarke, D. M. Mutational analysis of ABC proteins. *Archives of Biochemistry and Biophysics* **2008**, *476*, 51-64.
2. Gottesman, M. M.; Hrycyna, C. A.; Schoenlein, P. V.; Germann, U. A.; Pastan, I. Genetic Analysis of the Multidrug Transporter. *Annual Review of Genetics* **1995**, *29*, 607-649.
3. Kruijtzter, C. M. F.; Beijnen, J. H.; Rosing, H.; ten Bokkel Huinink, W. W.; Schot, M.; Jewell, R. C.; Paul, E. M.; Schellens, J. H. M. Increased Oral Bioavailability of Topotecan in Combination With the Breast Cancer Resistance Protein and P-Glycoprotein Inhibitor GF120918. *Journal of Clinical Oncology* **2002**, *20*, 2943-2950.

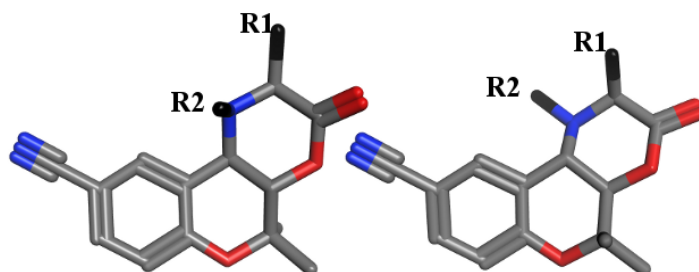
4. Klepsch, F.; Jabeen, I.; Chiba, P.; Ecker, G. F. Pharmacoinformatic Approaches to Design Natural Product Type Ligands of ABC-Transporters. *Current Pharmaceutical Design* **2010**, *16*, 1742-1752.
5. Zhang, S.; Yang, X.; Coburn, R. A.; Morris, M. E. Structure activity relationships and quantitative structure activity relationships for the flavonoid-mediated inhibition of breast cancer resistance protein. *Biochemical Pharmacology* **2005**, *70*, 627-639.
6. Zhang, S.; Morris, M. E. Effects of the Flavonoids Biochanin A, Morin, Phloretin, and Silymarin on P-Glycoprotein-Mediated Transport. *Journal of Pharmacology and Experimental Therapeutics* **2003**, *304*, 1258-1267.
7. Sheu, M.-T.; Liou, Y.-B.; Kao, Y.-H.; Lin, Y.-K.; Ho, H.-O. A Quantitative Structure-Activity Relationship for the Modulation Effects of Flavonoids on P-Glycoprotein-Mediated Transport. *Chemical & Pharmaceutical Bulletin* **2010**, *58*, 1187-1194.
8. Boumendjel, A.; McLeer-Florin, A.; Champelovier, P.; Allegro, D.; Muhammad, D.; Souard, F.; Derouazi, M.; Peyrot, V.; Toussaint, B.; Boutonnat, J. A novel chalcone derivative which acts as a microtubule depolymerising agent and an inhibitor of P-gp and BCRP in in-vitro and in-vivo glioblastoma models. *BMC Cancer* **2009**, *9*, 242.
9. Han, Y.; Riwanto, M.; Go, M.-L.; Rachel Ee, P. L. Modulation of breast cancer resistance protein (BCRP/ABCG2) by non-basic chalcone analogues. *European Journal of Pharmaceutical Sciences* **2008**, *35*, 30-41.
10. Liu, X. L.; Tee, H. W.; Go, M. L. Functionalized chalcones as selective inhibitors of P-glycoprotein and breast cancer resistance protein. *Bioorg Med Chem* **2008**, *16*, 171 - 180.
11. Seelig, A. A general pattern for substrate recognition by P-glycoprotein. *European Journal of Biochemistry* **1998**, *251*, 252-261.
12. Biot, C.; Glorian, G.; Maciejewski, L. A.; Brocard, J. S.; Domarle, O.; Blampain, G.; Millet, P.; Georges, A. J.; Abessolo, H.; Dive, D.; Lebibi, J. Synthesis and Antimalarial Activity in Vitro and in Vivo of a New Ferrocene- $\pi$ -Chloroquine Analogue. *Journal of Medicinal Chemistry* **1997**, *40*, 3715-3718.
13. Huber, D.; Hübner, H.; Gmeiner, P. 1,1'-Disubstituted Ferrocenes as Molecular Hinges in Mono- and Bivalent Dopamine Receptor Ligands. *Journal of Medicinal Chemistry* **2009**, *52*, 6860-6870.
14. Charonnet, E.; Filippini, M.-H. I. n.; Rodriguez, J. Novel DBU-MeOH-Promoted One-Pot Stereoselective  $\hat{\text{I}}^3$ -Functionalization of 1,3-Dicarbonyls: An Easy Access to  $\hat{\text{I}}^3$ -Arylidene,  $\hat{\text{I}}^3$ -Alkylidene and  $\hat{\text{I}}^3$ -Allylidene  $\hat{\text{I}}^{\pm}$ -Ketoesters and -amides. *Synthesis* **2001**, *2001*, 0788,0804.
15. Hogg, R. V.; Tanis, E. A. In *Probability and Statistical Inference*, Macmillan Publishing: New York, 1993.
16. Wang, R. B.; Kuo, C. L.; Lien, L. L.; Lien, E. J. Structure-activity relationship: analyses of p-glycoprotein substrates and inhibitors. *J Clin Pharm Ther* **2003**, *28*, 203-28.
17. Wiese, M.; Pajeva, I. K. Structure-activity relationships of multidrug resistance reversers. *Curr Med Chem* **2001**, *8*, 685-713.
18. Chiba, P.; Holzer, W.; Landau, M.; Bechmann, G.; Lorenz, K.; Plagens, B.; Hitzler, M.; Richter, E.; Ecker, G. Substituted 4-acylpyrazoles and 4-acylpyrazolones: synthesis and multidrug resistance-modulating activity. *J Med Chem* **1998**, *41*, 4001-11.

19. Hiessbock, R.; Wolf, C.; Richter, E.; Hitzler, M.; Chiba, P.; Kratzel, M.; Ecker, G. Synthesis and in vitro multidrug resistance modulating activity of a series of dihydrobenzopyrans and tetrahydroquinolines. *J Med Chem* **1999**, *42*, 1921-6.
20. Pleban, K.; Hoffer, C.; Kopp, S.; Peer, M.; Chiba, P.; Ecker, G. F. Intramolecular distribution of hydrophobicity influences pharmacological activity of propafenone-type MDR modulators. *Arch Pharm (Weinheim)* **2004**, *337*, 328-34.
21. König, G.; Chiba, P.; Ecker, G. F. Hydrophobic moments as physicochemical descriptors in structure-activity relationship studies of P-glycoprotein inhibitors. *Monatshefte für Chemie / Chemical Monthly* **2008**, *139*, 401-405.
22. Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L.; Chang, G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **2009**, *323*, 1718-22.
23. Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRIND-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J Med Chem* **2000**, *43*, 3233-43.
24. Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Computer Methodology* **1990**, *3*, 537-547.
25. Durán, Á.; Pastor, M. An advanced tool for computing and handling GRIND-INdependent. Descriptors. User Manual Version 1.06. **2011**.
26. Broccatelli, F.; Carosati, E.; Neri, A.; Frosini, M.; Goracci, L.; Oprea, T. I.; Cruciani, G. A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields. *J Med Chem* **2011**.
27. Boccard, J.; Bajot, F.; Di Pietro, A.; Rudaz, S.; Boumendjel, A.; Nicolle, E.; Carrupt, P. A. A 3D linear solvation energy model to quantify the affinity of flavonoid derivatives toward P-glycoprotein. *Eur J Pharm Sci* **2009**, *36*, 254-64.
28. Boumendjel, A.; Beney, C.; Deka, N.; Mariotte, A. M.; Lawson, M. A.; Trompier, D.; Baubichon-Cortay, H.; Di Pietro, A. 4-Hydroxy-6-methoxyaurones with high-affinity binding to cytosolic domain of P-glycoprotein. *Chem Pharm Bull (Tokyo)* **2002**, *50*, 854-6.
29. Boumendjel, A.; Di Pietro, A.; Dumontet, C.; Barron, D. Recent advances in the discovery of flavonoids and analogs with high-affinity binding to P-glycoprotein responsible for cancer cell multidrug resistance. *Med Res Rev* **2002**, *22*, 512-29.
30. Cianchetta, G.; Singleton, R. W.; Zhang, M.; Wildgoose, M.; Giesing, D.; Fravolini, A.; Cruciani, G.; Vaz, R. J. A pharmacophore hypothesis for P-glycoprotein substrate recognition using GRIND-based 3D-QSAR. *J Med Chem* **2005**, *48*, 2927-35.
31. Seelig, A. A general pattern for substrate recognition by P-glycoprotein. *Eur J Biochem* **1998**, *251*, 252-61.
32. Seelig, A. How does P-glycoprotein recognize its substrates? *Int J Clin Pharmacol Ther* **1998**, *36*, 50-4.
33. Jabeen, I.; Wetwitayaklung, P.; Chiba, P.; Pastor, M.; Ecker, G. F. Synthesis, Biological Activity and Quantitative Structure-Activity Relationship Studies of a Series of Benzopyranes and Benzopyrano[3,4-b][1,4]oxazines as Inhibitors of the Multidrug Transporter P-glycoprotein. *European Journal of Medicinal Chemistry* **2011**, submitted article.
34. Crivori, P.; Reinach, B.; Pezzetta, D.; Poggesi, I. Computational models for identifying potential P-glycoprotein substrates and inhibitors. *Mol Pharm* **2006**, *3*, 33-44.

35. Pajeva, I. K.; Wiese, M. Pharmacophore model of drugs involved in P-glycoprotein multidrug resistance: explanation of structural variety (hypothesis). *J Med Chem* **2002**, *45*, 5671-86.
36. Joshi, B. P.; Sharma, A.; Sinha, A. K. Ultrasound-assisted convenient synthesis of hypolipidemic active natural methoxylated (E)-arylalkenes and arylalkanones. *Tetrahedron* **2005**, *61*, 3075-3080.
37. Barnes, R. P.; Cochrane, C. C. The Properties of o-Methoxybenzoylmesitylmethane. *Journal of the American Chemical Society* **1942**, *64*, 2262-2262.
38. Xu, W.-Z.; Huang, Z.-T.; Zheng, Q.-Y. Synthesis of Benzo[c]xanthenes from 2-Benzylidene-1-tetralones by the Ultraviolet Radiation-Mediated Tandem Reaction. *The Journal of Organic Chemistry* **2008**, *73*, 5606-5608.
39. Durán, Á.; Pastor, M. Pentacle An advanced tool for computing and handling GRid- INdependent Descriptors. User manual version 0.9. In *User manual version 0.9*.
40. Duran, A.; Martinez, G. C.; Pastor, M. Development and validation of AMANDA, a new algorithm for selecting highly relevant regions in Molecular Interaction Fields. *J Chem Inf Model* **2008**, *48*, 1813-23.
41. Clementi, M.; Clementi, S.; Cruciani, G.; Pastor, M. Chemometric Detection of Binding Sites of 7TM Receptors. In *Molecular Modelling and Prediction of Bioreactivity*, Gundertofte, K., Jorgensen, F. S, Ed. Kluwer Academic/Plenum Publishers: New York, 2000; pp 207-212.
42. *Chemical Computing Group, Inc; Molecular Operating Environment (MOE)*, Quebec, Canada, 2010.

Synthesis, Biological Activity and Quantitative Structure-Activity Relationship Studies of a Series of Benzopyranes and Benzopyrano[3,4-b][1,4]oxazines as Inhibitors of the Multidrug Transporter P-glycoprotein

Page 70-103



In this part a data set of enantiomerically pure benzopyrano[3,4-b][1,4]oxazines were used to establish predictive models for P-glycoprotein inhibitors. This includes 2D- and 3D-QSAR models, using simple physicochemical as well as GRIND molecular descriptors.

#### Contents:

1. Introduction
2. Chemistry
3. Pharmacology
4. Results and Discussion
  - 4.1 Structure Activity Relationships (SAR)
  - 4.2 Hansch Analysis
  - 4.3 GRID-Independent Molecular Descriptor Analysis
5. Conclusions
6. Experimental Section
  - 6.1 Chemistry
  - 6.2 Biological Assay
  - 6.3 Computational Methods

① **Information:** This chapter was submitted to *European Journal of Medicinal Chemistry*, 2011 by **Ishrat Jabeen**, Penpun Wetwitayaklung, Peter Chiba, Manuel Pastor and Gerhard F. Ecker.

**Synthesis, Biological Activity and Quantitative Structure-Activity Relationship Studies of a Series of Benzopyranes and Benzopyrano[3,4-b][1,4]oxazines as Inhibitors of the Multidrug Transporter P-glycoprotein**

**Ishrat Jabeen<sup>a</sup>, Penpun Wetwitayaklung<sup>a</sup>, Peter Chiba<sup>b</sup>, Manuel Pastor<sup>c</sup> and Gerhard F. Ecker<sup>a,\*</sup>**

<sup>a</sup> *University of Vienna, Department of Medicinal Chemistry, Althanstrasse 14, 1090 Vienna, Austria*

<sup>b</sup> *Medical University of Vienna, Institute of Medical Chemistry, Weahringerstrasse 10, 1090, Vienna, Austria*

<sup>c</sup> *Research Unit on Biomedical Informatics (GRIB), IMIM/Universitat Pompeu Fabra, Dr. Aiguader 88, E-08003 Barcelona, Spain*

\*Author to whom correspondence should be addressed; e-mail:

[Gerhard.f.ecker@univie.ac.at](mailto:Gerhard.f.ecker@univie.ac.at), Phone: +43-1-4277-55110; fax: +43-1-4277-9551

**Abstract**

The ATP-binding cassette efflux transporter P-glycoprotein (P-gp) is notorious for contributing to multidrug resistance (MDR) in antitumor therapy. Due to its expression in many blood-organ barriers, it also influences the pharmacokinetics of drugs and drug candidates and is involved in drug/drug- and drug/nutrient interactions. However, due to lack of structural information the molecular basis of ligand/transporter interaction still needs to be elucidated. Towards this goal a data set of enantiomerically pure benzopyrano[3,4-*b*][1,4]oxazine analogs have been synthesized and pharmacologically tested. Both QSAR models using simple physicochemical and novel GRID-independent molecular descriptors (GRIND) were computed. The results from 2D-QSAR showed a linear correlation of vdw surface areas ( $\text{\AA}^2$ ) of hydrophobic atoms with pharmacological activity. GRIND studies allowed to identify important mutual distances between pharmacophoric features, which include one H-bond donor, two H-bond acceptors and two hydrophobic groups as well as their distances from different steric hot spots of the molecules. Activity of the compounds particularly increases with increase of the distance of an H-bond donor or a hydrophobic feature from a particular steric hot spot of the benzopyran analogs.

**1. Introduction**

Development of multidrug resistance (MDR) is one of the major challenges in cancer chemotherapy, as it limits the effectiveness of many clinically important agents [1]. One of the basic underlying mechanisms is over expression of the *mdr1* gene product, P-glycoprotein (P-gp) [2], which belongs to the ATP-binding cassette (ABC) family of transporters [3]. It is highly promiscuous in its ligand recognition profile and thus transports a large variety of structurally and functionally diverse compounds out of tumor cells [4]. Despite its role in tumor cells it is expressed at the epithelial cells of liver, kidney, intestine, and colon, as well as at the blood brain barrier. Thus, apart from playing an important role in maintaining a concentration gradient of toxic compounds at these physiological barriers, P-gp also modulates the pharmacokinetics of drugs that are recognized as P-gp substrates.

Within the past decade numerous inhibitors of P-gp mediated drug efflux have been identified [3]. Several compounds entered even phase III clinical studies, such as MS-

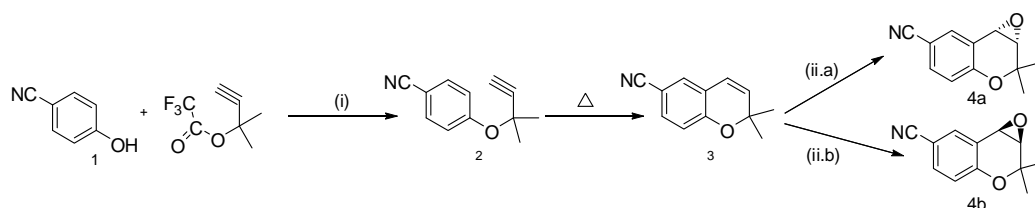


209 (dofequidar fumarate), tariquidar, valsopodar and elacridar [5, 6]. However, none made it to the market so far, mainly because of lack of efficacy or severe side effects.

In light of our extensive SAR and QSAR studies of propafenones, benzophenones and dihydrobenzopyranes, [7-9] a new class of conformationally restricted benzopyrano[3,4-b][1,4]oxazines have been synthesized and biologically tested with respect to their ability to block P-gp mediated daunomycin efflux. These new P-gp inhibitors offer the advantage of remarkably reduced conformational flexibility and thus might be versatile molecular tools for probing stereoselective differences of drug/P-gp interaction [10], as well as for 3D-QSAR studies. These might be performed by utilising alignment-dependent approaches, such as CoMFA and CoMSIA, or by alignment independent methods using descriptors derived from Molecular Interaction Fields (MIFs), as exemplified with GRIND [11]. The latter thus also allow the analysis of structurally diverse data series. GRID MIFs have been applied to many areas of computational drug discovery, including 3D-QSAR [12], high-throughput virtual screening [13], ADME profiling, kinetic [14, 15] and metabolism [16] prediction of early drug candidates. Within this manuscript we explore the capability of the GRIND approach to derive predictive 3D-QSAR models for a set of diastereomeric benzopyrano[3,4-b][1,4]oxazines.

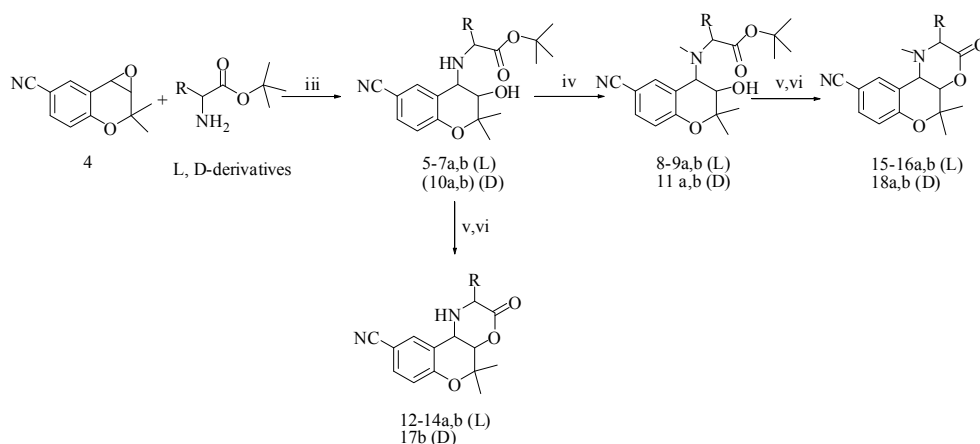
## 2. Chemistry

Synthesis of the benzopyran common scaffold was achieved in analogy to the procedure reported by Godfrey *et al* [17] and following our strategy outlined recently [10]. Briefly, O-alkylation of 4-hydroxybenzoxonitril (**1**) with 3-trifluoroacetyl-3-methyl-but-1-yne followed by thermal cyclization gave 6-cyano-2,2-dimethyl-2H-1-benzopyran **3**. Enantioselective epoxidation was performed in analogy to the method published by Lee *et al* [18]. The enantioselective epoxidation of benzopyranes with Jacobsons Mn(III) Salen catalyst and commercial household bleach (sodium hypochlorite) as a stoichiometric oxygen source yielded (S,S)- and (R,R)-epoxides **4a** and **4b**, respectively (Scheme 1). Enantiomeric purity of both epoxides was confirmed by HPLC analysis, using a LiChroCART (R,R)-Whelk-01 column (25x0,4 cm) and n-hexane/isopropanol (95:5) as eluent.



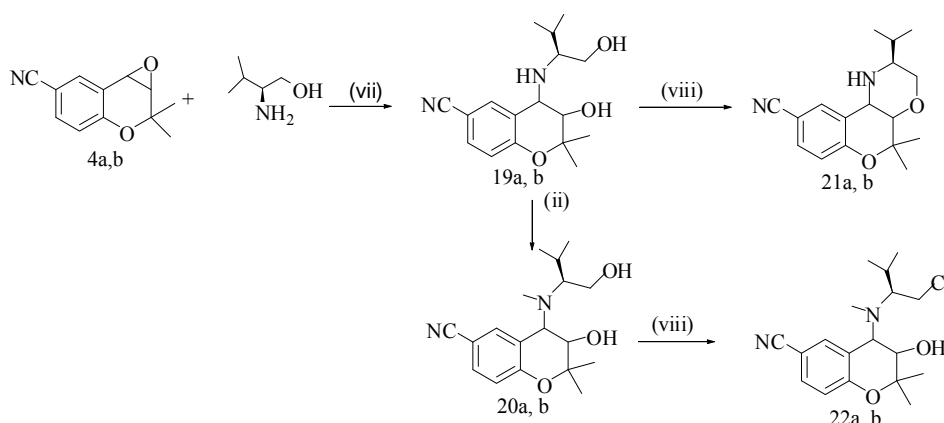
**Scheme 1.** Synthesis of the benzopyran ring system and enantiomeric pure (S,S)- and (R,R)-epoxide **4a,b**; **(i)** DBU, CuCl<sub>2</sub>, -4 °C, Ar atmosphere; **(ii.a)** (S,S)-Mn (III) Salen NaOCl solution, buffer to pH 11.3, 0 °C; **(ii.b)** (R,R)-Mn (III) Salen NaOCl solution, buffer to pH 11.3, 0 °C

Nucleophilic ring opening of these epoxides with L- and D-amino acid t-butyl esters is regioselective and stereoselective, thus giving optically pure *trans* 3,4-disubstituted benzopyranes. t-butyl esters of L-alanine, L- and D-valine as well as L-phenylalanine were reacted with each epoxid-enantiomer to give the diastereomeric esters **5a-7b**, and **10a,b**, respectively. **5a-6b** and **10a,b** were further N-methylated to yield **8a-9b** (derived from L-alanine and -valine) and **11a,b** (derived from D-valine). All tert-butyl-esters (**5a-11b**) were hydrolysed with 70% HClO<sub>4</sub> [19] to yield the corresponding acids which were subsequently cyclized without further purification using bis-(2-oxo-3-oxazolidinyl) phosphinic chloride, 4-dimethylaminopyridine and triethylamine to yield the target compounds **12a-18b** (Scheme 2). In the subsequent sections, compounds derived from epoxide enantiomers **4a** and **4b** are classified as series (a) and series (b), respectively.



**Scheme 2.** Synthesis of target compounds **12a-18b**; **(iii)** 96% ethanol, reflux; **(iv)** Acidic menthol, paraformaldehyde, sodium cyanoborohydride **(v)** 70% HClO<sub>4</sub>; **(vi)** 4-dimethylaminopyridine, bis-(2-oxo-3-oxazolidinyl)-phosphinic chloride, triethylamine, dichloromethane.

For means of comparison, also a set of corresponding ethers were synthesized. The solution of both enantiomers of epoxides **4a,b** in 96% ethanol were reacted with L-valinol to yield amino alcohol substituted 2H-1-benzopyran-3-ols **19a,b**. N-methylation as described before gave the tertiary amines **20a,b**. Valinol analogs **19a,b** were successfully cyclised by mesylation followed by intramolecular O-alkylation to yield **21a,b**. Surprisingly, in case of the tertiary amine (**20a,b**), the cyclisation failed and the respective chloro derivatives **22a,b** were obtained (Scheme 3).



**Scheme 3.** Synthesis of target compounds **19a-22b**; (vii) 96% ethanol at 65°C reflux for 5 days; (viii) Trimethylamine, triethylamine hydrochlorid, and solution of methane sulfonyl chloride in small amount of toluene at 0°C.

### 3. Pharmacology

Biological activity of target compounds **5a-22b** was assessed using the daunorubicin efflux protocol as described previously [20]. Briefly, multidrug resistant CCRF-CEM vcr 1000 cells were incubated with daunorubicin and the decrease in mean cellular fluorescence in dependence of time was measured in presence of various concentrations of the modulator. IC<sub>50</sub> values were calculated from the concentration-response curve of efflux  $V_{max}/K_m$  vs concentration of the modulator. Thus, the effect of different modulators on the transport rate is measured in a direct functional assay. Values are given in Table 1 and are the mean of at least three independently performed experiments. Generally, inter experimental variation was below 20%.

**Table 1.** Enantiomerically pure benzopyrano[3,4-b][1,4]oxazines (**5a-22b**) and their IC<sub>50</sub> values.

| #   | Scaffold | Stereo Chemistry | R1   | R2              | IC <sub>50</sub> μM | logP(o/w) |
|-----|----------|------------------|--|-----------------|---------------------|-----------|
| 5a  | (A)      | (3S,4R)          | (L) CH <sub>3</sub>                                  | H               | 29.85               | 2.84      |
| 5b  | (A)      | (3R,4S)          | (L) CH <sub>3</sub>                                  | H               | 14.55               | 2.84      |
| 6a  | (A)      | (3S,4R)          | (L) CH(CH <sub>3</sub> ) <sub>2</sub>                | H               | 2.40                | 3.82      |
| 6b  | (A)      | (3R,4S)          | (L) CH(CH <sub>3</sub> ) <sub>2</sub>                | H               | 2.70                | 3.82      |
| 7a  | (A)      | (3S,4R)          | (L) CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | H               | 0.55                | 4.38      |
| 7b  | (A)      | (3R,4S)          | (L) CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | H               | 0.77                | 4.38      |
| 8a  | (A)      | (3S,4R)          | (L) CH <sub>3</sub>                                  | CH <sub>3</sub> | 3.96                | 3.11      |
| 8b  | (A)      | (3R,4S)          | (L) CH <sub>3</sub>                                  | CH <sub>3</sub> | 3.72                | 3.11      |
| 9a  | (A)      | (3S,4R)          | (L) CH(CH <sub>3</sub> ) <sub>2</sub>                | CH <sub>3</sub> | 0.96                | 4.08      |
| 9b  | (A)      | (3R,4S)          | (L) CH(CH <sub>3</sub> ) <sub>2</sub>                | CH <sub>3</sub> | 1.35                | 4.08      |
| 10a | (A)      | (3S,4R)          | (D) CH(CH <sub>3</sub> ) <sub>2</sub>                | H               | 4.62                | 3.81      |
| 10b | (A)      | (3R,4S)          | (D) CH(CH <sub>3</sub> ) <sub>2</sub>                | H               | 1.34                | 3.81      |
| 11a | (A)      | (3S,4R)          | (D) CH(CH <sub>3</sub> ) <sub>2</sub>                | CH <sub>3</sub> | 1.01                | 4.08      |
| 11b | (A)      | (3R,4S)          | (D) CH(CH <sub>3</sub> ) <sub>2</sub>                | CH <sub>3</sub> | 1.00                | 4.08      |
| 12a | (B)      | (2S,4aS,10bR)    | CH <sub>3</sub>                                      | H               | 1241.65             | 1.98      |
| 12b | (B)      | (2S,4aR,10bS)    | CH <sub>3</sub>                                      | H               | 76.89               | 1.98      |
| 13a | (B)      | (2S,4aS,10bR)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | H               | 15.32               | 2.94      |
| 13b | (B)      | (2S,4aR,10bS)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | H               | 59.33               | 2.94      |
| 14a | (B)      | (2S,4aS,10bR)    | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> )     | H               | 2.68                | 3.51      |
| 14b | (B)      | (2S,4aR,10bS)    | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> )     | H               | 259.78              | 3.51      |
| 15a | (B)      | (2S,4aS,10bR)    | CH <sub>3</sub>                                      | CH <sub>3</sub> | 47.83               | 2.24      |
| 15b | (B)      | (2S,4aR,10bS)    | CH <sub>3</sub>                                      | CH <sub>3</sub> | 28.93               | 2.24      |
| 16a | (B)      | (2S,4aS,10bR)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | CH <sub>3</sub> | 47.51               | 3.21      |
| 16b | (B)      | (2S,4aR,10bS)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | CH <sub>3</sub> | 16.70               | 3.21      |
| 17b | (B)      | (2R,4aR,10bS)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | H               | 9.63                | 2.95      |
| 18a | (B)      | (2R,4aS,10bR)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | CH <sub>3</sub> | 79.27               | 3.22      |
| 18b | (B)      | (2R,4aR,10bS)    | CH(CH <sub>3</sub> ) <sub>2</sub>                    | CH <sub>3</sub> | 27.84               | 3.22      |
| 19a | (C)      | (2S,3S,4R)       | ---  | H               | 54.05               | 2.14      |
| 19b | (C)      | (2S,3R,4S)       | ---  | H               | 102.64              | 2.14      |
| 20a | (C)      | (2S,3S,4R)       | ---  | CH <sub>3</sub> | 5.46                | 2.14      |
| 20b | (C)      | (2S,3R,4S)       | ---  | CH <sub>3</sub> | 6.84                | 2.14      |
| 21a | (D)      | (2S,4aS,10bR)    | ---  | H               | 48.80               | 3.10      |
| 21b | (D)      | (2S,4aR,10bS)    | ---  | H               | 44.00               | 3.10      |
| 22a | (E)      | (2S,3S,4R)       | ---  | CH <sub>3</sub> | 35.22               | 3.66      |
| 22b | (E)      | (2S,3R,4S)       | ---  | CH <sub>3</sub> | 45.17               | 3.66      |

## 4. Results and Discussion

### 4.1 Structure Activity Relationships (SAR)

Biological activity values of the data series cover a range of more than three orders of magnitude (Table 1) with the two phenylalanine esters **7a** and **7b** being the most

active compounds (**7a**: 0.55  $\mu\text{M}$ ; **7b**: 0.77  $\mu\text{M}$ ), followed by N-methylated L-valine analogues **9a** (0.96  $\mu\text{M}$ ) and **9b** (1.35  $\mu\text{M}$ ), which are by a factor of 2 more active than the corresponding unsubstituted analogs **6a** (2.40  $\mu\text{M}$ ) and **6b** (2.70  $\mu\text{M}$ ). The same trend could be observed for the respective D-valine derivatives. This observation is even more pronounced for the alanine derivatives (compare methylated analogs **8a** (3.96  $\mu\text{M}$ ) and **8b** (3.72  $\mu\text{M}$ ) vs respective secondary amines **5a** (29.85  $\mu\text{M}$ ) and **5b** (14.55  $\mu\text{M}$ )). This most probably is due to a logP effect with more lipophilic compounds showing higher biological activity, which has been shown for numerous classes of P-gp inhibitors [21].

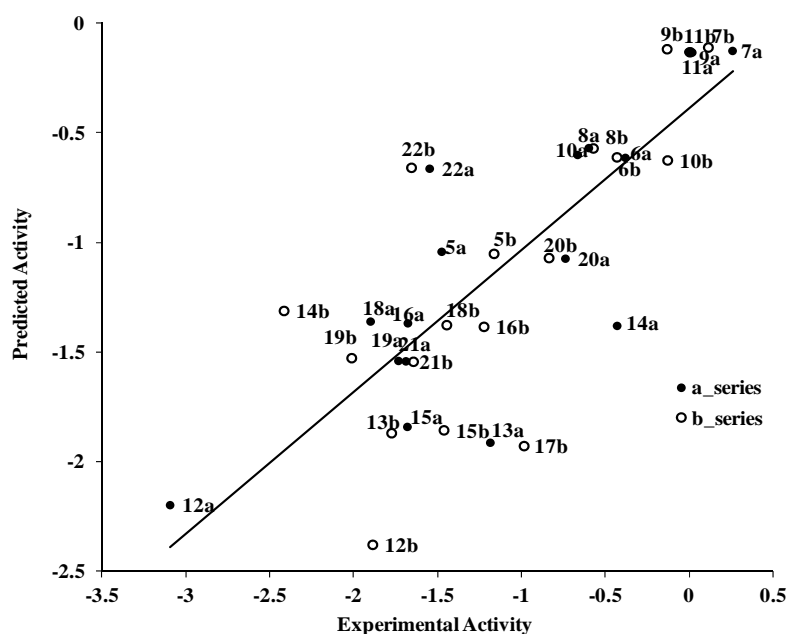
It has to be noted that for all seven diastereoisomeric pairs showing a bicyclic scaffold almost no differences in biological activity exist. However, this pattern changes remarkably upon ring closure to the tricyclic benzopyrano[3,4-b][1,4]oxazines. While all stereoisomers containing a valine moiety (**13a,b**; **16a,b**, **18a,b**, **19a,b-22a,b**) are still within one order of magnitude, both the alanine and phenylalanine derivatives exhibit remarkable differences in their activity values. Interestingly, in case of alanine, the 4aS,10bR-isomer **12a** is by a factor of 15 less active than the diastereomeric 4aR,10bS analogue **12b**, whereas in case of the phenylalanine derivatives this behavior reverses with the 4aS,10bR-isomer **14a** being by two orders of magnitude more active than **14b**. This difference in their activities might be due to difference in mode of interaction of diastereoisomeric pairs as has been indicated in a preceding publication [10].

## 4.2 Hansch Analysis

3D structures of all diastereoisomers were built with the builder function of MOE 2009-10 and energy minimised using the MMFF94 force field which uses a bond charge increment method to set the electrostatic partial charges [22]. In order to determine the influence of physicochemical properties of the compounds on their biological activity, QSAR analyses were performed by using the software package MOE version 2009-10. Multiple linear regression analysis using MOE's contingency analysis tool for identification of the most important descriptors revealed an equation solely based on the hydrophobic Van der Waals surface area ( $V_{\text{sa\_hyd}}$ ) (Equ. 1). Interestingly, descriptors related to electrostatic properties, such as topological polar surface area and molar refractivity, did not show significant contributions to the model.

$$\text{Log}(1/IC_{50}) = 0.01 (V_{\text{sa\_hyd}}) - 4.74 \quad (\text{Equ. 1})$$

$n = 35$ ,  $R^2 = 0.67$ ,  $q^2(\text{LOO}) = 0.63$ ,  $\text{RMSE} = 0.48$

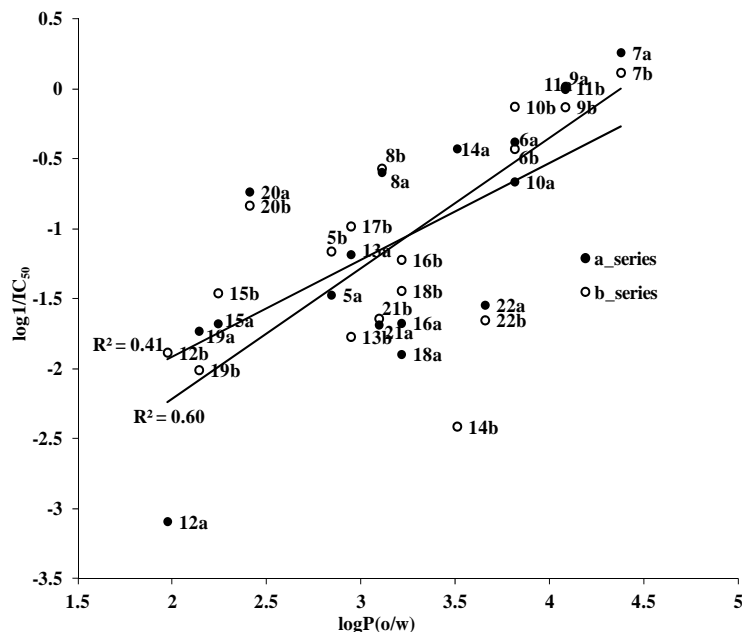


**Figure 1.** Plot of observed vs. predicted MDR-modulating activity, predicted values were obtained with leave-one-out cross validation procedure.

Figure 1 shows a plot of observed vs biological activity predicted by QSAR equation 1, which demonstrates the validity of the QSAR model. However, compounds **14b** and **22b** show outlier behaviour (residual value above one log unit). Upon removal of these two compounds, the  $q^2$  value improves to 0.70. Interestingly, both compounds belong to the (b) series of diastereoisomers, which indicates that for this series additional factors other than lipophilicity might play a role.

$V_{\text{sa\_hyd}}$  describes the sum of VDW surface areas of hydrophobic atoms ( $\text{\AA}^2$ ). This is perfectly in line with previous studies which showed that distribution of hydrophobicity within the molecules influences their mode of interaction with P-gp [23] and lipophilicity needs to be considered as a space directed property [24, 25]. This space-directedness might be indicative for different orientations of molecules within the binding area of P-gp, which is mainly hydrophobic [26]. In addition, overall lipophilicity ( $\log P(\text{o/w})$ ) of the compounds (i.e. to enrich in biological membranes) also plays an important role, as stressed out in numerous publications [7, 27, 28]. A separate  $\log P(\text{o/w})$  analysis of the two diastereoisomeric series reveals the same picture as already outlined previously [10]. Also for this extended data set the (a) series of compounds shows a better correlation with  $\log P$  values ( $R^2 = 0.60$ ) than the respective

diastereoisomers of the (b) series ( $R^2 = 0.41$ ) (Figure 2). However, within the latter the phenylalanine **14b** seems to be an outlier. Upon removal of **14b** from the data set, the  $R^2$  value improves up to 0.59. This once more strengthens the hypothesis of steric constrains caused by the benzyl moiety in **14b** which might lead to a different binding mode [10].

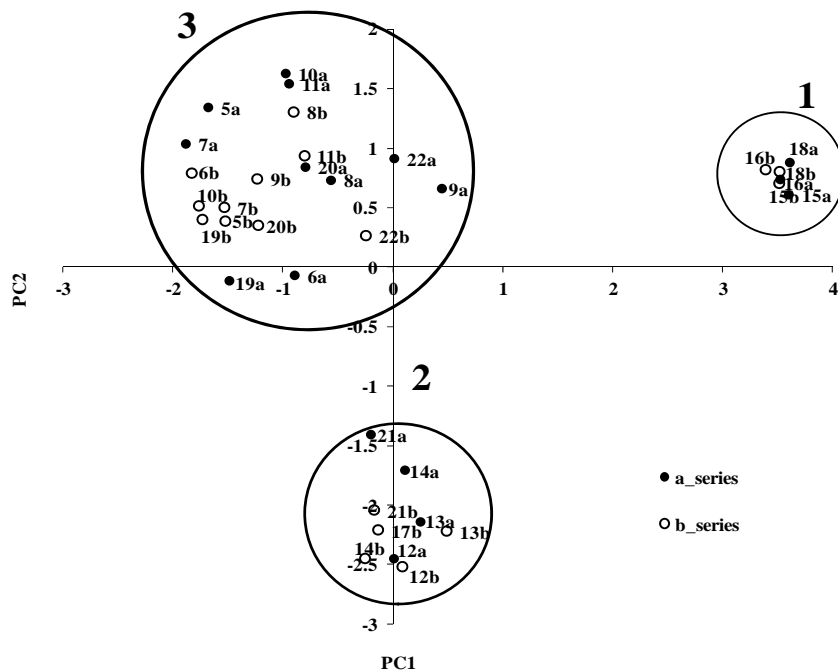


**Figure 2.** Correlation of MDR-modulating activity of compounds vs.  $\log P(o/w)$  values.

#### 4.3. GRID Independent Molecular Descriptor (GRIND) Analysis

Molecules along with their activity values (expressed as  $\log 1/IC_{50}$ ) were loaded into the software package Pentacle (v 1.06) [29] to derive 3D-QSAR model using GRIND descriptors (see methods). According to previous findings for propafenone analogs, all compounds were modeled in their neutral form [30]. Structural variance of the data was analyzed with principal component analysis (PCA) performed on the complete set of GRIND descriptors. The first two principal components explain about 32% of the descriptor variance in the data set. Principal component analysis (PCA) on the data matrix showed that the whole data set splits up into three different clusters (Figure 3). The 1<sup>st</sup> PC separates the compounds on basis of shape and H-bond donors in the molecules. Molecules in the cluster on the right hand side (cluster 1) do not contain any H-bond donor, while those located on the upper left hand side (cluster 2) contain one H-bond donor group. The 3<sup>rd</sup> cluster (cluster 3) exhibits compounds containing two H-

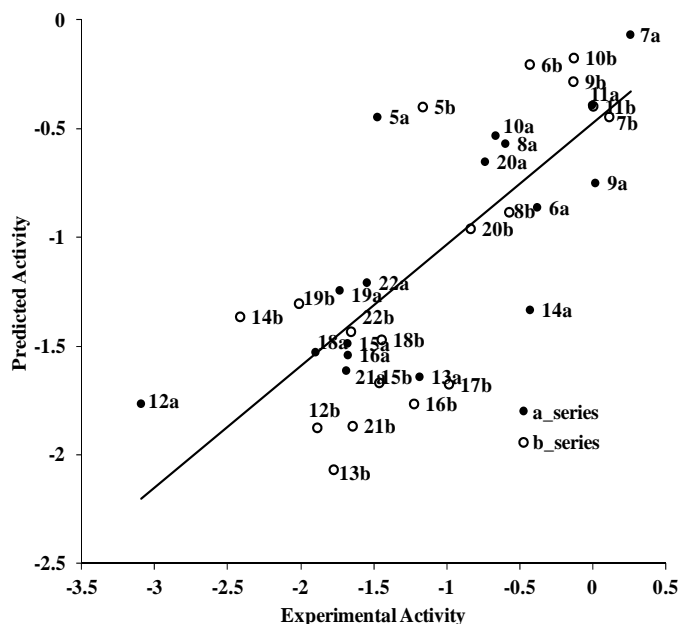
bond donor groups in their structures. The 2<sup>nd</sup> PC separates rigid and smaller compounds (cluster 1 and cluster 2) from the flexible ones (cluster 3).



**Figure 3.** PCA Score plot shows the whole data set consists of three different types of inhibitors of P-gp overall no outlier has been observed in the dataset.

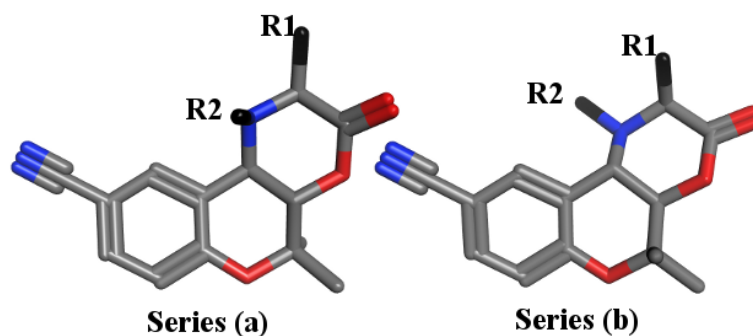
To analyze the pharmacophoric aspect of ligand-protein interaction, PLS analysis correlating the activity with the complete set of variables (450) was carried out using the AMANDA algorithm implemented in Pentacle (v 1.06). This resulted in a one-latent variable model with an  $r^2 = 0.51$  and a cross-validated (LOO)  $q^2$  value of 0.27, which was quite unsatisfactory. Thus, variable selection was applied to reduce the variable number using FFD factorial selection [31] implemented in Pentacle. This resulted in decrease of active variables from 335 to 196 and an increase of model performance ( $r^2$  of 0.72,  $q^2 = 0.58$ , standard error of prediction 0.52 (Figure 4).





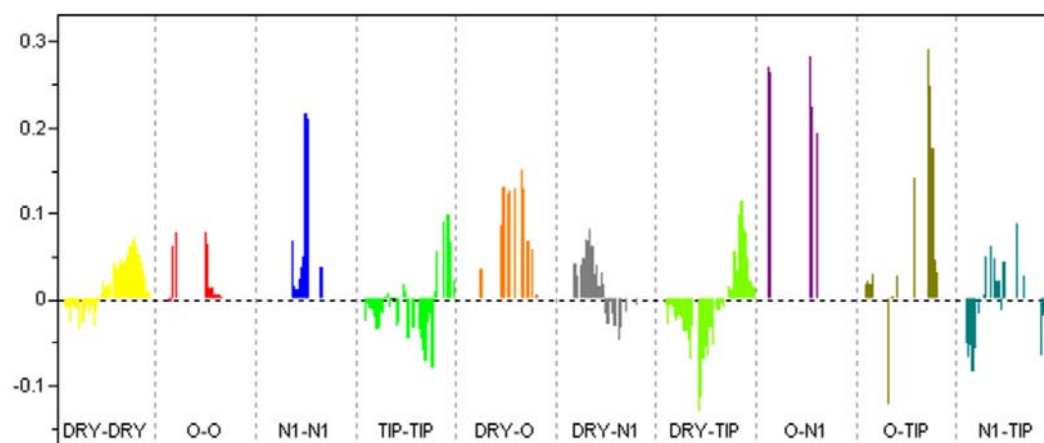
**Figure 4.** Plot of observed vs. predicted (LOO) MDR-modulating activity ( $\log 1/IC_{50}$ ), of inhibitors of P-gp obtained with the GRIND algorithm

With the exception of compounds **14b**, **5a**, and **12a**, all compounds were within one order of magnitude with their predicted values (**14b**: obs, 23.21; pred, 259.77; **5a**: obs, 29.84; pred, 2.80; **12a**, obs, 1241.65; pred, 58.30) (Figure 4). The outlier behavior of these three compounds might be due to potential different interaction behavior of the two diastereomeric series as reported by Jabeen *et al* [10]. However, building two separate QSAR models composed of compounds of series (a) and series (b) in two separate training sets showed an analogous picture and did not improve the results (data not shown). Thus, although GRIND descriptors are able to capture different configurations, they were not able to extract the stereochemical differences of the two series. This might be due to the fact that the molecules are quite compact (Figure 5).



**Figure 5.** 3D representatives of series (a) having 2S,4aS,10bR-configuration and series (b) having 2S,4aR,10bS-configuration.

Analysis of the PLS coefficients profile of the 1<sup>st</sup> Latent Variable of the PLS model allows to identify those descriptors which exhibit strongest contribution to the model. According to the correlogram plot given in figure 6, certain distances of the N1-N1, O-N1, and O-TIP probes are participating most in explaining the variance in the activity values.



**Figure 6.** PLS Coefficient correlograms showing the descriptors which are directly (positive value) or inversely (negative values) correlated to activity. Activity particularly increases with the increase in (N1-N1), (O-N1) and (O-TIP) descriptor value.

**Table 2.** Summary of GRIND variables and their corresponding distances that are identified as being highly correlated to biological activity of compounds **5a-22b**

| Variable | Correlogrm | Distance      | Comment   |
|----------|------------|---------------|---|
| 33       | DRY-DRY    | 13.20-13.60 Å | Optimal distance separating two hydrophobic groups. More pronounced in phenylalanine derivatives  |
| 112      | N1-N1      | 8.80-9.20 Å   | Related to two hydrogen bond acceptor atoms in the molecules. This is mainly associated to the carbonyl group and the hydroxyl groups in tertiary butyl esters. |
| 321      | O-N1       | 2.40-2.80 Å   | Well pronounced in tert-butyl esters with IC <sub>50</sub> ~1µM. Positive contribution towards biological activity.   |
| 339      | O-N1       | 9.60-10.00 Å  | Complements N1-N1, contributing directly to the biological activity   |
| 392      | O-TIP      | 12.80-13.20 Å | H-bond donor present far away from a steric hot spot; positive contribution   |
| 374      | O-TIP      | 5.60-6.00 Å   | H-bond donor present quite near to a steric hot spot; contributing negatively   |
| 308      | DRY-TIP    | 15.20-15.60 Å | Complements to DRY-DRY correlogram; positive contribution to biological activity  |

The sum of the VDW surface areas of hydrophobic atoms ( $vsa_{hyd}$ ) has emerged as an important determinant for high biological activity of benzopyran-type P-gp inhibitors (Equ. 1). The 3D-QSAR model using GRIND descriptors further refines this general property and identified two hydrophobic regions (DRY-DRY) separated by a certain distance range in all active compounds. These represent the aromatic ring of the benzopyran ring system and R1. In the most active phenylalanine derivatives (**7a,b** and **14a,b**) the two regions are separated by a distance of 13.20-13.60 Å, which is considered optimal according to the GRIND model. Thus, adding a large hydrophobic group (large  $vsa_{hyd}$ ) at the position of R1 might lead to a further increase of the biological activity.

Previous QSAR studies on propafenone derivatives have demonstrated the importance of H-bond acceptors and their distance from the central aromatic ring [32, 33]. Furthermore, Seelig [34, 35] more explicitly defined two patterns of H-bond acceptor groups and their fixed spatial distance observed in ligands of P-gp. Pattern I contains two H-bond acceptors separated by a distance of  $2.51 \pm 0.30$  Å, while pattern II comprises two or three H-bond acceptor groups at a distance of  $4.60 \pm 0.60$  Å apart. Interestingly, the 3D-QSAR model based on benzopyrano[3,4-b][1,4]oxazines identified an optimal distance of 8.80-9.20 Å between two H-bond acceptor groups (N1-N1) in all compounds exhibiting  $IC_{50} \sim 1$  μM. The N1-N1 probe is mainly associated to the carbonyl group and the hydroxy group in tertiary butyl esters **7a-11b**. For tricyclic compounds (**15a-16b** and **18a,b**) it is associated to the distance of the carbonyl group and the tertiary nitrogen atom. Finally, for amino alcohols **19a-20b** this descriptor refers to the two hydroxy groups in the molecules. This indicates that two H-bond acceptors are important attributes for biological activity of P-gp inhibitors if they are separated by a distance of  $\sim 8.80-9.20$  Å (Figure 7a), which is in line with several other studies. Crivori *et al.*, used GRIND descriptors to identify 3D pharmacophoric features which differentiate P-gp inhibitors from substrates. They reported two H-bond acceptors at a distance of 8.00 Å apart from each other in P-gp inhibitors [36], whereas a distance of 11.50 Å between two H-bond acceptors, along with the importance of shape descriptors, have been reported by Cianchetta *et al.* for substrates of P-gp [37].

However, despite of similarities in the number of H-bond acceptors necessary for good biological activity, a direct comparison of distance matrices thereof across

different chemical scaffolds reveals some differences. This most probably is due to the fact that the large binding site of P-gp has multiple spots able to contribute to H-bond interactions and that different chemical series most probably utilize different H-bond interaction patterns.

Apart a certain number of H-bond acceptors, also one H-bond donor along with hydrophobicity have been identified as important pharmacophoric features of P-gp inhibitors/substrates [9, 33]. It is worth noting that a very similar MIF based pharmacophore of P-gp inhibitors was recently published by Broccatelli *et al* [38]. They identified one H-bond acceptor and two large hydrophobic regions, together with an optimal molecular shape, as being important for high activity, and successfully used their model for virtual screening to identify new P-gp inhibitors. The results are further in line with Boccard *et al* [39] outlining an optimal shape and hydrophobicity as major physicochemical parameters responsible for the affinity of flavonoid derivatives for P-gp [40, 41].

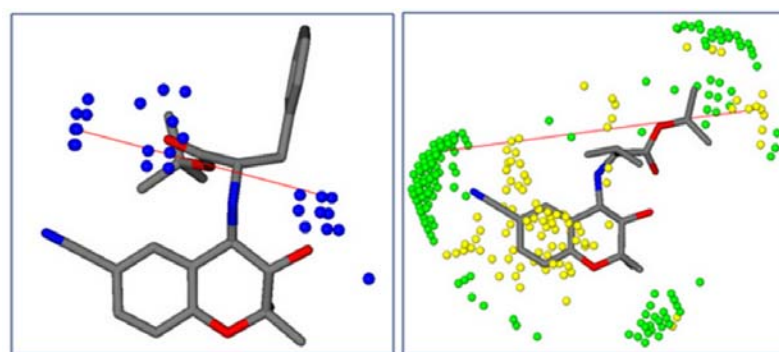
Also in our GRIND model shape based probes (TIP) defining steric hot spots exhibit a significant contribution. Especially the 9-carbonitrile group in the benzopyran scaffold encodes an important molecular boundary (steric hot spot) and serves as anchor for defining optimal distance ranges to an H-bond donor (O-TIP correlogram) as well as to a hydrophobic feature (DRY-TIP correlogram). The O-TIP combination of probes encodes the shape of the molecules (steric hot spots) together with an H-bond donor group. Interestingly, O-TIP coefficients are negative for a distance between 5.60-6.00 Å, but become positive for larger distances (12.80 to 13.20 Å). These distances (12.80 to 13.20 Å) are present in benzopyranes bearing tertiary butyl esters (**5a-11b**) and amino alcohol derivatives (**19a-20b** and **22a,b**) as shown in figure 7c. In tricyclic diastereoisomers (**12a-14b** and **17b**) these descriptors are linked to shorter distances and mark (-NH) as an H-bond donor at a distance of 5.60-6.00 Å apart from the cyano group, which is the main group contributing to the TIP MIF. This is related to a negative influence towards biological activity (Figure 7d). This indicates that more potent P-gp inhibitors show extended conformations and have an H-bond donor group far from regions with a strong TIP probe related field.

Analyzing the Dry-TIP correlogram it becomes evident that a hydrophobic group in a distance of 15.20-15.60 Å from one of the “edges” of the molecule (steric hot spot,

cyano group) positively contributes to biological activity. In tert-butyl esters (**5a-11b**) and **14a,b** these two probes map the distance between a hydrophobic group ( $R_1$ ) (**14a,b**) or tert-butyl group in **5a-11b** from the cyano group (Figure 7b). In analogy to the O-TIP correlogram, DRY-TIP shows a negative contribution towards biological activity for shorter distances (7.60-8.00 Å) of these probes.

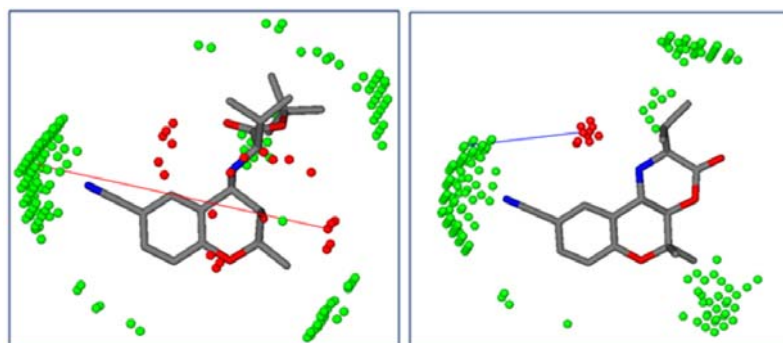
Finally, the O-N1 correlogram (H-bond donor – H-bond acceptor) points towards two positive contributions at a distance of 9.60-10.00 Å and 2.40-2.80 Å, respectively (Figure 7e, f). The first distance is linked to the hydroxyl and carbonyl group in **5a-11b** and is complementary to the N1-N1 correlogram as already discussed. The second distance refers to the –NH and carbonyl group. O-N1 probes at both distance ranges are well pronounced in t-butyl esters (**5a-11b**) as well as in amino alcohol substituted derivatives (**19a-20 b** and **22a,b**). However, in all tricyclic compounds (**12a-18b**) the two probes do not fit either of the distance ranges.

To summarize, two H-bond-acceptor groups and one H-bond donor at a particular distance from each other and from a particular “edge” or steric hot spot of the molecule plays a major role in the interaction of benzopyran-type P-gp inhibitors (Figure 7).



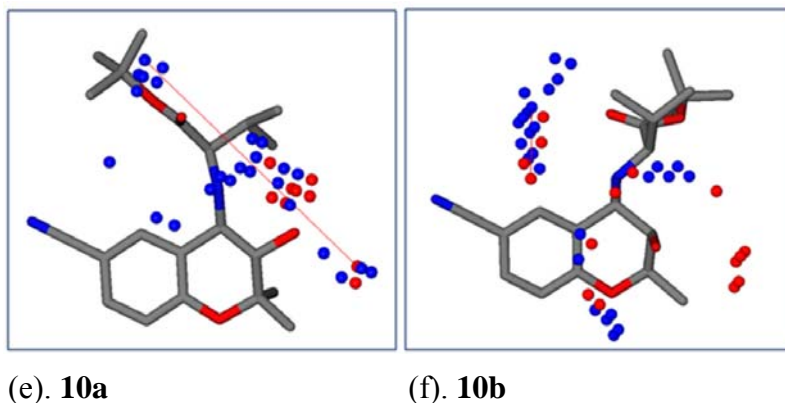
(a). 7a

(b). 6b



(c). 10b

(d). 20b



**Figure 7.** (a). Represents two H-bond acceptors (N1-N1: blue hot spots) at a distance of 8.80-9.20 Å (b). Dry-TIP represents a hydrophobic probe (DRY: yellow hot spots) at a distance of 15.20-15.60 Å from a steric hot spot (TIP: green region) (c) O-TIP outline an H-bond donor (OH) (O: red hot spot) at a distance of 12.80-13.20 Å from the 9-carbonitril edge” of the molecule. (d) Marks an H-bond donor (-NH) at a distance of 5.60-6.00 Å from the 9-carbonitril edge of the molecule (O-TIP) (e) Representing an H-bond donor (OH) at a distance of 9.60-10.00 Å from an H-bond acceptor (C=O), present only in esters (O-N1). (f) Representing, H-bond donor (-NH) at a distance of 2.40-2.80 Å from an H-bond acceptor (C=O) (O-N1).

## 5. Conclusions

Benzopyrano-[3,4-b][1,4]oxazines are versatile molecular tools to probe the stereoselectivity of P-glycoprotein. For a distinct substitution pattern, different pairs of diastereoisomers exhibit a large difference in their potency to inhibit P-gp mediated drug efflux pump. However, GRIND-based 3D-QSAR models emphasise could not link these differences to concrete differences of distances of pharmacophoric hot spots. Nevertheless, GRIND analysis provided a reasonably well performing 3D-QSAR model outlining a set of important distances of pharmacophoric features. Two H-bond-acceptor groups, one H-bond donor at a particular distance from each other as well as distinct distances of these probes to steric hot spots play a major role in the interaction of benzopyran-type P-gp inhibitors. Activity particularly increases with increase of the distance of an H-bond donor or a hydrophobic feature from a particular steric hot spot of the benzopyran analogs. This not only further strengthens the importance of H-bonding, but also indicates that a certain shape/configuration of the molecules is important for high activity. Further analyses will focus on a generalisation of this concept.

## 6. Experimental Section.

### 6.1. Chemistry

Melting points all compounds were determined on Kofler hot plate apparatus and are uncorrected. Infrared spectra were recorded on a Perkin Elmer 298 spectrometer, a Bruker Avance Dpx 200 spectrometer, a Varian Unity plus 300 spectrometer  $^1\text{H}$  spectra were referenced to tetramethylsilane internal standard ( $\delta$  0.0);  $^{13}\text{C}$  spectroscopy  $\text{CDCl}_3$  served as the internal standard ( $\delta$  77.0) and on a Bruker AM 360 L spectrometer. An asterisk indicates peaks of double intensity. GC-MS spectrum was recorded on a HP 5890 A gas chromatography / HP 5970 mass detector. Optical rotations were measured on a Perkin Elmer 241 polarimeter in a standardised cuvette. Flash chromatography was carried out on MERCK silica gel 60, TLC on plastic sheets (Merck silica gel 60 F254).

**6.1.1. General procedure for the enantiomerically pure (S,S)- (4a) and (R,R)- epoxide (4b).** Commercial household bleach (DanKlorix<sup>®</sup>) was buffered to pH 11.3 with 0.05 N  $\text{Na}_2\text{HPO}_4$  and 1N NaOH and then cooled to 0°C. To 1000 mL of this solution a solution of **3** (75.58 mmol) and Mn(III) Salen catalyst ( $2.74 \times 10^{-3}$  mmol) in 76 mL of  $\text{CH}_2\text{Cl}_2$  was added, stirred at 0°C for 5 hr and then at room temperature overnight. The mixture was filtered through Celite and the organic phase was separated, brined once, dried ( $\text{Na}_2\text{SO}_4$ ) and brought to dryness. Purification by flash chromatography (petroleum ether-ethylacetate; 8:2) yield 76.9% of (S,S)-**4a** and 78.9% of (R,R)-**4b** as colourless crystals; mp 133-135°C;  $^1\text{H}$  NMR (200MHz;  $\text{CDCl}_3$ ):  $\delta$  = 1.28 (s, 3H,  $\text{CH}_3$ ), 1.57(s, 3H,  $\text{CH}_3$ ), 3.52 (d, 1H,  $J$  = 4.52 Hz, 3-H/ 4-H), 3.89 (d, 1H,  $J$  = 4.52 Hz, 3-H/ 4-H), 6.84 (d, 1H,  $J$  = 8.53 Hz, 8-H), 7.51 (dd, 1H,  $J$  = 2.00 Hz,  $J$  = 8.41 Hz, 7-H), 7.63 (d, 1H,  $J$  = 2.01 Hz, 5-H);  $^{13}\text{C}$  NMR (200MHz,  $\text{CDCl}_3$ ):  $\delta$  = 22.99 ( $\text{CH}_3$ ), 25.46 ( $\text{CH}_3$ ), 49.34 (3-C), 62.27 (4-C), 74.64 (2-C), 104.27 (6-C), 118.70 (CN), 119.00 (8-C), 121.67 (4a-C), 133.77, 134.38 (5-C, 7-C), 156.45 (8a-C); IR (KBr): 2227 (CN)  $\text{cm}^{-1}$ , 1280 (epoxide)  $\text{cm}^{-1}$ .

**6.1.2. General procedure for amino acid-tert-butyl-ester (5a-7b, 10a,b).** A solution of enantiomeric pure epoxide **4a** or **4b** (4.97 mmol) and corresponding D and L-amino acid-tert-butyl ester (5.47 mmol) in 50 mL 96% ethanol was stirred at 80°C for 5 days, then evaporated in vacuo. Purification by flash chromatography (petroleum ether/ethylacetate = 8/2) yield respective amino acid t-butyl ester (**5a-11b**). Compounds **5a-7b** and **12a-14b** are already published by Jabeen *et al.*[10]

**6.1.2.1. (3S,4R)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-D-valine-tert-butyl-ester (10a).** (S,S) epoxide 4a and D-valine tert-butyl-ester gave 10a, yield (59%) as yellowish oil;  $[\alpha]^{20} = (+) 182.38$  ( $c = 0.105$ , in  $\text{CH}_2\text{Cl}_2$ );  $^1\text{H NMR}$  (200MHz,  $\text{CDCl}_3$ ):  $\delta = 0.89, 0.99$  (each d, each 3H, each  $J = 7.03$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.17 (s, 3H, 2- $\text{CH}_3$ ), 1.44 (s, 12H, 2- $\text{CH}_3$ ,  $\text{C}(\text{CH}_3)_3$ ), 1.86-2.05 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.38 (br, 1H, NH), 2.76 (d, 1H,  $J = 4.27$  Hz, N-CH-CO), 3.14 (dd, 1H,  $J = 9.54$  Hz, 3.77 Hz, 3-H), 3.78 (d, 1H,  $J = 9.54$  Hz, 4-H), 3.84 (d, 1H,  $J = 3.77$  Hz, OH), 6.80 (d, 1H,  $J = 8.53$  Hz, 8-H), 7.39 (dd, 1H,  $J = 8.56$  Hz, 1.76 Hz, 7-H), 7.94 (d, 1H,  $J = 1.76$  Hz, 5H);  $^{13}\text{C NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta = 17.67$  (CH- $\text{CH}_3$ ), 19.06 (2 $\text{CH}_3$ ), 19.77 (CH- $\text{CH}_3$ ), 27.04 (2- $\text{CH}_3$ ), 27.84 ( $\text{C}(\text{CH}_3)_3$ ), 32.65 ( $\text{CH}(\text{CH}_3)_2$ ), 56.54 (4-C), 60.53 (N-CH-CO), 71.01 (3-C), 79.60 ( $\text{C}(\text{CH}_3)_3$ ), 82.33 (2-C), 103.58 (6-C), 118.19 (8-C), 119.34 (CN), 123.82 (4a-C), 132.42, 133.40 (5-C, 7C), 157.56 (8a-C), 177.98 (C=O); MS  $m/z$  375 (0.10%, M+), 273 (36.9%), 160 (22.7%); Anal. Calcd for  $\text{C}_{21}\text{H}_{30}\text{N}_2\text{O}_4$  %: C 67.35, H 8.07, N 7.48; found: C 67.62, H 8.27, N 7.21.

**6.1.2.2 (3R,4S)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-D-valine-tert-butyl-ester (10b).** (R,R) epoxide 4b and D-valine tert-butyl-ester gave 10b, yield (57%) as yellowish oil;  $[\alpha]^{20} = (+) 27.43$  ( $c = 0.113$ , in  $\text{CH}_2\text{Cl}_2$ );  $^1\text{H NMR}$  (200MHz,  $\text{CDCl}_3$ ):  $\delta = 0.92, 1.06$  (each d, each 3H, each  $J = 6.78$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.13 (s, 3H, 2- $\text{CH}_3$ ), 1.43 (s, 12H, 2- $\text{CH}_3$ ,  $\text{C}(\text{CH}_3)_3$ ), 2.00-2.19 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 1.86 (br, 1H, NH), 3.37 (d, 1H,  $J = 4.26$  Hz, N-CH-CO), 3.43 (br, 2H, 3H, 4H), 4.17 (br, 1H, OH), 6.74 (d, 1H,  $J = 8.54$  Hz, 8-H), 7.34 (dd, 1H,  $J = 8.53$  Hz, 1.75 Hz, 7H), 7.69 (d, 1H,  $J = 1.75$  Hz, 5H);  $^{13}\text{C NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta = 17.69$  (CH- $\text{CH}_3$ ), 18.81 (2- $\text{CH}_3$ ), 19.49 (CH- $\text{CH}_3$ ), 26.70 (2- $\text{CH}_3$ ), 27.92 ( $\text{C}(\text{CH}_3)_3$ ), 32.55 ( $\text{CH}(\text{CH}_3)_2$ ), 56.68 (4-C), 66.54 (N-CH-CO), 74.24 (3-C), 79.72 ( $\text{C}(\text{CH}_3)_3$ ), 82.13 (2-C), 103.33(6-C), 117.91 (8-C), 119.28 (CN), 126.38 (4a-C), 132.30 (5-C, 7-C), 156.46 (8a-C), 175.98 (C=O); MS  $m/z$  375 (0.18, M+), 273 (28.7%), 160 (32%); Anal. Calcd for  $\text{C}_{21}\text{H}_{30}\text{N}_2\text{O}_4$  %: C 67.35, H 8.07, N 7.48; found: C 67.65, H 8.36, N 7.38.

**6.1.3. General procedure Synthesis of N-methyl derivatives of tert-butyl esters (8a–9b, 11ab).** The suspension of corresponding amino acid-ter-butyl ester (**5a–6b**, **10a,b**) with paraformaldehyde and sodium cyanoborohydride in acidic menthol was stirred and refluxed at room temperature, 4 days, then evaporated in vacuo. Flash chromatography (petroleum ether: ethyl acetate/ 6:1:2)



**6.1.3.1. (3S,4R)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-L-alanine-tert-butyl-ester (8a).** From 5a gave 8a, yield 81% as yellowish oil;  $[\alpha]^{20} = (+) 6.92$  ( $c = 0.325$ , in  $\text{CH}_2\text{Cl}_2$ );  $^1\text{H NMR}$  (200MHz,  $\text{CDCl}_3$ ):  $\delta = 1.21, 1.50$  (each s, each 3H, 2x2- $\text{CH}_3$ ), 1.52 (s, 9H,  $\text{C}(\text{CH}_3)_3$ ), 1.55 (d, 3H,  $J = 5.56$  Hz,  $\text{CH}-\text{CH}_3$ ), 2.17 (s, 3H,  $\text{N}-\text{CH}_3$ ), 3.77 (br, 2H, 3-H, 4-H), 3.84 (q, 1H,  $J = 5.56$  Hz,  $\text{CH}-\text{CH}_3$ ), 5.52 (br, 1H, OH), 6.82 (d, 1H,  $J = 8.46$  Hz, 8-H), 7.39 (dd, 1H,  $J = 2.02$  Hz, 8.46 Hz, 7-H), 7.65 (d, 1H,  $J = 2.02$  Hz, 5-H);  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ ):  $\delta = 17.11, 18.47$  (2x2- $\text{CH}_3$ ), 26.76 ( $\text{CH}-\text{CH}_3$ ), 27.63 ( $\text{C}(\text{CH}_3)_3$ ), 63.42 (4-C), 65.24 ( $\text{N}-\text{CH}-\text{CO}$ ), 68.76 (3-C), 79.66 ( $\text{C}(\text{CH}_3)_3$ ), 82.33 (2-C), 103.12 (6-C), 117.90 (8-C), 119.22 (CN), 123.96 (4a-C), 132.04\*, 132.52\* (5-C, 7-C), 157.54 (8a-C), 175.67 ( $\text{C}=\text{O}$ ); MS  $m/z$  360 (0.35,  $\text{M}^+$ ), 259 (52.7), 160 (23.5); Anal. Calcd for  $\text{C}_{20}\text{H}_{28}\text{N}_2\text{O}_4$  %: C 66.64, H 7.83, N 7.77; found: C 66.38, H 7.69, N 7.73.

**6.1.3.2. (3R,4S)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-L-alanine-tert-butyl-ester (8b).** From 5b gave 8b, yields 78% as colourless oil;  $[\alpha]^{20} = (-)40.31$  ( $c = 0.129$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 2225 (CN)  $\text{cm}^{-1}$ , 1708 (COOR)  $\text{cm}^{-1}$ ;  $^1\text{H NMR}$  (200MHz,  $\text{CDCl}_3$ ):  $\delta = 1.20$  (s, 3H, 2- $\text{CH}_3$ ), 1.42 (d, 3H,  $J = 7.45$  Hz,  $\text{CH}-\text{CH}_3$ ), 1.45 (s, 9H,  $\text{C}(\text{CH}_3)_3$ ), 1.50 (s, 3H, 2- $\text{CH}_3$ ), 2.44 (s, 3H,  $\text{N}-\text{CH}_3$ ), 3.60 (dd, 1H,  $J = 2.15$  Hz, 9.85 Hz, 3-H), 3.91 (d, 1H,  $J = 9.85$ , 4-H), 5.62 (br, 1H, OH), 6.78 (d, 1H,  $J = 8.59$  Hz, 8-H), 7.37 (dd, 1H,  $J = 2.15$  Hz, 8.46 Hz, 7-H), 7.63 (d, 1H,  $J = 2.15$  Hz, 5-H);  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ ):  $\delta = 16.69, 18.86$  (2x2- $\text{CH}_3$ ), 27.13 ( $\text{CH}-\text{CH}_3$ ), 27.91 ( $\text{C}(\text{CH}_3)_3$ ), 34.96 ( $\text{N}-\text{CH}_3$ ), 58.66 (4-C), 59.61 ( $\text{N}-\text{CH}-\text{CO}$ ), 70.05 (3-C), 80.21 ( $\text{C}(\text{CH}_3)_3$ ), 82.45 (2-C), 103.40 (6-C), 118.01 (8-C), 119.63 (CN), 125.11 (4a-C), 132.25\*, 132.83\* (5-C, 7-C), 158.08 (8a-C), 177.79 ( $\text{C}=\text{O}$ ); MS  $m/z$  361 (0.35,  $\text{M}^+$ ), 259 (53.5%); Anal. Calcd for  $\text{C}_{20}\text{H}_{28}\text{N}_2\text{O}_4$  %: C 66.64, H 7.83, N 7.77; found: C 66.76, H 7.55, N 7.65.

**6.1.3.3. (3S,4R)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-L-valine-tert-butyl-ester (9a).** From 6a gave 9a, yield (96%) as colourless oil;  $[\alpha]^{20} = (+) 25$  ( $c = 0.252$ , in  $\text{CH}_2\text{Cl}_2$ );  $^1\text{H NMR}$  (200MHz,  $\text{CDCl}_3$ ):  $\delta = 0.97, 1.09$  (each d, each 3H, each  $J = 6.70$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.12 (s, 3H, 2- $\text{CH}_3$ ), 1.45 (s, 12H, 2- $\text{CH}_3$ ,  $\text{C}(\text{CH}_3)_3$ ), 2.09 (s, 3H,  $\text{N}-\text{CH}_3$ ), 2.11-2.27 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 3.34 (d, 1H,  $J = 7.83$  Hz,  $\text{N}-\text{CH}-\text{CO}$ ), 3.68 (d, 2H,  $J = 9.59$  Hz, 4-H), 3.74 (dd, 1H,  $J = 9.54$  Hz, 2.40 Hz, 3-H), 4.74 (d, 1H,  $J = 2.4$  Hz, OH), 6.76 (d, 1H,  $J = 8.46$  Hz, 8-H), 7.34 (dd, 1H,  $J =$

8.46 Hz, 2.02 Hz, 7-H), 7.82 (d, 1H,  $J = 2.02$  Hz, 5H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta = 18.49$  (CH-CH<sub>3</sub>), 19.49, 19.85 (CH-CH<sub>3</sub>, 2-CH<sub>3</sub>), 26.91 (2-CH<sub>3</sub>), 27.91 (C(CH<sub>3</sub>)<sub>3</sub>, CH(CH<sub>3</sub>)<sub>2</sub>), 29.83 (N-CH<sub>3</sub>), 65.15 (4-C), 68.28 (N-CH-CO), 74.02 (3-C), 79.68 (C(CH<sub>3</sub>)<sub>3</sub>), 82.61 (2-C), 103.44 (6-C), 118.24 (8-C), 119.43 (CN), 124.53 (4a-C), 132.22, 132.84 (5-C, 7C), 157.67 (8a-C), 174.30 (C=O); MS  $m/z$  338 (0.12%, M<sup>+</sup>), 287 (36%), 160 (17%); Anal. Calcd for C<sub>22</sub>H<sub>32</sub>N<sub>2</sub>O<sub>4</sub> %: C 68.01, H 8.30, N 7.21; found: C 68.76, H 6.66, N 6.86.

**6.1.3.4. (3R,4S)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-L-valine-tert-butyl-ester (9b).** From 6b gave 9b, yield 84% as colourless oil;  $[\alpha]^{20} = (-) 20.94$  ( $c = 0.117$ , in CH<sub>2</sub>Cl<sub>2</sub>);  $^1\text{H}$  NMR (200MHz, CDCl<sub>3</sub>):  $\delta = 0.95$ , 1.06 (each d, each 3H, each  $J = 6.57$  Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.17 (s, 3H, 2-CH<sub>3</sub>), 1.46 (s, 12H, 2-CH<sub>3</sub>, C(CH<sub>3</sub>)<sub>3</sub>), 2.28 (s, 3H, N-CH<sub>3</sub>), 2.07-2.22 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 3.14 (d, 1H,  $J = 8.97$  Hz, N-CH-CO), 3.92 (d, 1H,  $J = 9.51$  Hz, 4-H), 3.61 (dd, 1H,  $J = 9.51$  Hz, 2.27 Hz, 3-H), 5.48 (br, 1H, OH), 6.76 (d, 1H,  $J = 8.46$  Hz, 8-H), 7.34 (dd, 1H,  $J = 8.46$  Hz, 2.02 Hz, 7-H), 7.65 (s, 1H, 5H);  $^{13}\text{C}$  NMR (CDCl<sub>3</sub>):  $\delta = 18.66$  (CH-CH<sub>3</sub>), 19.59 (2-CH<sub>3</sub>), 21.12 (CH-CH<sub>3</sub>), 27.05 (2-CH<sub>3</sub>), 27.91 (C(CH<sub>3</sub>)<sub>3</sub>), 28.22 (CH(CH<sub>3</sub>)<sub>2</sub>), 33.89 (N-CH<sub>3</sub>), 58.88 (4-C), 69.20 (N-CH-CO), 72.36 (3-C), 79.94 (C(CH<sub>3</sub>)<sub>3</sub>), 82.64 (2-C), 103.38 (6-C), 118.09 (8-C), 119.46 (CN), 124.53 (4a-C), 132.17, 133.01 (5-C, 7C), 158.11 (8a-C), 177.34 (C=O); MS  $m/z$  338 (0.10%, M<sup>+</sup>), 287 (31%), 160 (15.3%); Anal. Calcd for C<sub>22</sub>H<sub>32</sub>N<sub>2</sub>O<sub>4</sub> %: C 68.01, H 8.30, N 7.21; found: C 68.05, H 8.32, N 6.97.

**6.1.3.5. (3S,4R)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-D-valine-tert-butyl-ester (11a).** From 10a gave 11a, yield 96% as colourless oil;  $[\alpha]^{20} = (+) 19.05$  ( $c = 0.273$ , in CH<sub>2</sub>Cl<sub>2</sub>);  $^1\text{H}$  NMR (200MHz, CDCl<sub>3</sub>):  $\delta = 0.96$ , 1.07 (each d, each 3H, each  $J = 6.63$  Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.17 (s, 3H, 2-CH<sub>3</sub>), 1.47 (s, 12H, 2-CH<sub>3</sub>, C(CH<sub>3</sub>)<sub>3</sub>), 2.28 (s, 3H, N-CH<sub>3</sub>), 2.08-2.22 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 3.16 (d, 1H,  $J = 8.97$  Hz, N-CH-CO), 3.62 (d, 1H,  $J = 9.54$  Hz, 3-H), 3.93 (d, 1H,  $J = 9.54$  Hz, 4-H), 5.54 (br, 1H, OH), 6.77 (d, 1H,  $J = 8.46$  Hz, 8-H), 7.35 (dd, 1H,  $J = 8.46$  Hz, 1.64 Hz, 7-H), 7.66 (s, 1H, 5H);  $^{13}\text{C}$  NMR (CDCl<sub>3</sub>):  $\delta = 18.59$  (CH-CH<sub>3</sub>), 19.59 (2-CH<sub>3</sub>), 21.13 (CH-CH<sub>3</sub>), 27.06 (2-CH<sub>3</sub>), 27.94 (C(CH<sub>3</sub>)<sub>3</sub>), 28.24 (CH(CH<sub>3</sub>)<sub>2</sub>), 33.89 (N-CH<sub>3</sub>), 58.81 (4-C), 69.24 (N-CH-CO), 72.41 (3-C), 79.96 (C(CH<sub>3</sub>)<sub>3</sub>), 82.73 (2-C), 103.30 (6-C), 118.13 (8-C), 119.50 (CN), 124.54 (4a-C), 132.23, 133.04 (5-C, 7-C), 158.16 (8a-

C), 177.42 (C=O); MS m/z 338 (0.14%, M<sup>+</sup>), 287 (56.6%), 160 (22.3%); Anal. Calcd for C<sub>22</sub>H<sub>32</sub>N<sub>2</sub>O<sub>4</sub> %: C 68.01, H 8.30, N 7.21; found: C 68.24, H 8.32, N 7.47.

**6.1.3.6. (3R,4S)-N-methyl-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4-yl)-D-valine-tert-butyl-ester (11b).** From 10b gave 11b, yield 92% as colourless oil;  $[\alpha]^{20} = (-) 25.79$  (c = 0.126, in CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (200MHz, CDCl<sub>3</sub>): δ = 0.98, 1.11 (each d, each 3H, each J = 6.70 Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.14 (s, 3H, 2-CH<sub>3</sub>), 1.46 (s, 12H, 2-CH<sub>3</sub>, C(CH<sub>3</sub>)<sub>3</sub>), 2.10 (s, 3H, N-CH<sub>3</sub>), 2.15-2.26 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 3.36 (d, 1H, J = 7.7 Hz, N-CH-CO), 3.69 (dd, 1H, J = 9.73 Hz, 2.27 Hz, 3-H), 3.75 (d, 1H, J = 9.73 Hz, 4-H), 4.77 (d, 1H, J = 2.27 Hz, OH), 4.77 (d, 1H, J = 8.50 Hz, 8H), 7.35 (dd, 1H, J = 8.50 Hz, 2.00 Hz, 7-H), 7.83 (s, 1H, J = 2.00 Hz, 5H); <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ = 18.54 (CH-CH<sub>3</sub>), 19.54 (2-CH<sub>3</sub>), 19.88 (CH-CH<sub>3</sub>), 26.96 (2-CH<sub>3</sub>), 27.97 (C(CH<sub>3</sub>)<sub>3</sub>), (CH(CH<sub>3</sub>)<sub>2</sub>), 29.89 (N-CH<sub>3</sub>), 65.24 (4-C), 68.34 (N-CH-CO), 74.08 (3-C), 79.73 (C(CH<sub>3</sub>)<sub>3</sub>), 82.71 (2-C), 103.48 (6-C), 118.29 (8-C), 119.50 (CN), 124.56 (4a-C), 132.28, 132.91 (5-C, 7-C), 157.73 (8a-C), 174.39 (C=O); MS m/z 338 (0.17%, M<sup>+</sup>), 287 (46.6%), 160 (22%); Anal. Calcd for C<sub>22</sub>H<sub>32</sub>N<sub>2</sub>O<sub>4</sub> %: C 68.01, H 8.30, N 7.21; found: C 68.24, H 8.55, N 6.99.

**6.1.4. General procedure for cyclisation (12a-18b).** 4.61 mmol of amino acid-tert-butyl ester (**5a,b-11a,b**) was dissolved in a small amount of CH<sub>2</sub>Cl<sub>2</sub>, hydrolysed by 6 mL of 70% HClO<sub>4</sub>, stirred overnight, and 4N NH<sub>4</sub>OH solution was added slowly. The precipitate was dried and used in the next reaction step without further purification. A suspension of precipitates (2.76 mmol), 4-dimethylaminopyridine (0.69 mmol) and bis (2-oxo-3-oxazolidinyl) phosphinic chloride (4.12 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was heated to reflux at 80°C for 10 min, then triethylamine (0.95 mL, 6.85 mmol) was added and the solution was refluxed at 70°C for 4 days. The suspension was filtered and evaporated to dryness. Purification was done by flash chromatography (petroleum ether/ethylacetate; 9:1) to yield the target compounds (**12a-18b**).

**6.1.4.1. (2S,4aS,10bR)-N-methyl-,2,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1] benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (15a).** From 8a gave 15a, 51% as white crystal; mp 120-121°C;  $[\alpha]^{20} = (+)112.40$  (c = 0.121, in CH<sub>2</sub>Cl<sub>2</sub>); IR (KBr) 2231 (CN) cm<sup>-1</sup>, 1745 (lactone) cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ = 1.30 (s, 3H, 5-CH<sub>3</sub>), 1.51 (d, 3H, J = 7.2 Hz, 2-CH<sub>3</sub>), 1.54 (s, 3H, 5-CH<sub>3</sub>), 2.20 (s, 3H, N-CH<sub>3</sub>), 3.95 (q, 1H, J =

7.2 Hz, N-CH-CO), 4.19 (d, 1H, J = 11.24 Hz, 4a-H), 4.51 (d, 1H, J = 11.24 Hz, 10b-H), 6.87 (d, 1H, J = 8.53 Hz, 7H), 7.45 (dd, 1H, J = 1.87 Hz, 8.53 Hz, 7H), 7.75 (d, 1H, J = 1.87 Hz, 10H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  14.93, 19.66 (2x5- $\text{CH}_3$ ), 26.05 (2- $\text{CH}_3$ ), 31.11 (N- $\text{CH}_3$ ), 56.24 (10b-C), 61.14 (2-C), 74.50 (4a-C), 78.19 (5-C), 104.47 (9-C), 118.51 (7-C), 118.88 (CN), 120.08 (10a-C), 131.53\*, 133.19\* (8-C, 10-C), 156.61 (6a-C), 171.49 (C=O); MS  $m/z$  286 (1.00, M+), 185 (22%), 170 (100%); Anal. Calcd for  $\text{C}_{16}\text{H}_{18}\text{N}_2\text{O}_3$  %: C, 67.12, H 6.34, N 9.78: found: C 67.20, H 6.29, N 9.75.

**6.1.4.2. (2S,4aR,10bS)-N-methyl-2,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1] benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (15b).** From 8b gave 15b, yielded 62% as yellow crystal; mp 91.5-92.5 °C;  $[\alpha]^{20} = (-)114.91$  ( $c = 0.108$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr): 2224 (CN)  $\text{cm}^{-1}$ , 1750 (C=O)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta = 1.29$  (s, 3H, 5- $\text{CH}_3$ ), 1.54 (d, 3H, J = 4.8 Hz, 2- $\text{CH}_3$ ), 1.56 (s, 3H, 5- $\text{CH}_3$ ), 2.32 (s, 3H, N $\text{CH}_3$ ), 3.60 (q, 1H J = 7.33 Hz, 2H), 4.11 (d, 1H, J = 10.99 Hz, 4a-H), 4.51 (d, 1H, J = 10.99 Hz, 10b-H), 6.87 (d, 1H, J = 8.56 Hz, 7H), 7.45 (dd, 1H, J = 1.96 Hz, 8.56 Hz, 8H), 7.72 (d, 1H, J = 1.96 Hz, 10-H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta = 18.69$ , 19.43 (2x5- $\text{CH}_3$ ), 26.05 (2- $\text{CH}_3$ ), 37.26 (N- $\text{CH}_3$ ), 51.23 (10b-C), 60.49 (N-CH-CO), 73.92 (4a-C), 78.03 (5-C), 104.50 (9-C), 118.46 (7-C), 118.87 (CN), 120.37 (10a-C), 131.66\*, 133.11\* (8-C, 10-C), 156.69 (6a-C), 171.83 (C=O); MS  $m/z$  286. (1.86, M+), 185 (21%), 170 (100%); Anal. Calcd for  $\text{C}_{16}\text{H}_{18}\text{N}_2\text{O}_3$  %: C 67.12, H 6.34, N 9.78; found: C 67.26, H 6.50, N 9.36.

**6.1.4.3. (2S,4aS,10bR)-2-isopropyl-1,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (16a).** From 9a gave 16a, 57% as yellowish oil;  $[\alpha]^{20} = (+) 119.57$  ( $c = 0.115$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 2225 (CN)  $\text{cm}^{-1}$ , 1757 (lactone)  $\text{cm}^{-1}$ , 1574 (NH);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta = 0.97$ , 1.06 (each d, each 3H, J = 6.57 Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.38 (s, 3H, 5- $\text{CH}_3$ ), 1.54 (s, 3H, 5- $\text{CH}_3$ ), 1.63-1.84 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.71(s, 3H, N- $\text{CH}_3$ ), 2.92 (d, 1H, J = 10.24 Hz, N-CH-CO), 3.61 (d, 1H, J = 11.75 Hz, 4a-H), 4.26 (d, 1H, J = 11.75 Hz, 10b-H), 6.83 (d, 1H, J = 8.56 Hz, 7H), 7.44 (dd, 1H, J = 8.56 Hz, 2.02 Hz, 8H), 7.72 (br, 1H, 10H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta = 19.30$  ( $\text{CH}-\text{CH}_3$ ), 20.40, 21.19 (5- $\text{CH}_3$ ,  $\text{CH}-\text{CH}_3$ ), 27.14 (5- $\text{CH}_3$ ), 32.55 ( $\text{CH}(\text{CH}_3)_2$ ), 47.14 (N- $\text{CH}_3$ ), 57.05 (10b-C), 73.27 (N-CH-CO), 76.13 (4a-C), 78.58 (5-C), 104.17 (9-C), 117.83 (7-C), 119.23 (CN), 123.90 (10a-C), 130.95, 132.97 (8-C, 10-C), 155.59 (6a-C), 170.31 (C=O); MS  $m/z$  314 (0.41%, M+), 185 (38.5%), 170

(92.4%); Anal. Calcd for  $C_{18}H_{22}N_2O_3$  %: C, 68.77, H 7.05, N 8.91: found: C 68.65, H 7.11, N 8.47.

**6.1.4.4. (2S,4aR,10bS)-2-isopropyl-1,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (16b).** From **9b** gave **16b**, 52% as yellowish solid; mp 124-125 °C;  $[\alpha]^{20} = (-) 119.35$  ( $c = 0.124$ , in  $CH_2Cl_2$ ); IR (KBr) 2230 (CN)  $cm^{-1}$ , 1741 (lactone)  $cm^{-1}$ ;  $^1H$  NMR ( $CDCl_3$ ):  $\delta = 1.07, 1.14$  (each d, each 3H,  $J = 6.69$  Hz,  $CH(CH_3)_2$ ), 1.26 (s, 3H, 5CH<sub>3</sub>), 1.55 (s, 3H, 5CH<sub>3</sub>), 2.30 (s, 3H, N-CH<sub>3</sub>), 2.35-2.45 (m, 1H,  $CH(CH_3)_2$ ), 3.17(d, 1H,  $J = 5.94$  Hz, N-CH-CO), 4.07 (d, 1H,  $J = 10.55$  Hz, 4a-H), 4.54 (d, 1H,  $J = 10.55$  Hz, 10b-H), 6.88 (d, 1H,  $J = 8.58$  Hz, 7H), 7.46(dd, 1H,  $J = 8.56$  Hz, 1.64 Hz, 8H), 7.78 (br, 1H, 10H);  $^{13}C$  NMR( $CDCl_3$ ):  $\delta = 19.06$  ( $CH(CH_3)_2$ ), 19.99\*25.99 (2\*5-CH<sub>3</sub>), 31.81 ( $CH(CH_3)_2$ ), 37.93 (N-CH<sub>3</sub>), 53.21 (10b-C), 70.86 (N-CH-CO), 73.63 (4a-C), 78.01 (5-C), 104.62 (9-C), 118.70 (7-C), 118.95 (CN), 120.26 (10a-C), 131.81, 133.10 (8-C, 10-C), 156.73 (6a-C), 170.73 (C=O); MS  $m/z$  314 (7.3%, M+), 243 (31%), 170 (100%); Anal. Calcd for  $C_{18}H_{22}N_2O_3$  %: C 61.62, H 6.61, N, 7.98; found: C 60.01, H 6.11, N 7.47

**6.1.4.5. (2R,4aR,10bS)-2-isopropyl-5,5-dimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (17b).** From **10b** gave **17b**, 45% as pale yellowish crystal; mp 164.5-166°C;  $[\alpha]^{20} = (-) 84.81$  ( $c = 0.158$ , in  $CH_2Cl_2$ ); IR (KBr) 2224 (CN)  $cm^{-1}$ , 1733 (lactone)  $cm^{-1}$ ;  $^1H$  NMR ( $CDCl_3$ ):  $\delta = 1.00, 1.12$  (each d, each 3H, each  $J = 7.08$  Hz,  $CH(CH_3)_2$ ), 1.29 (s, 3H, 5-CH<sub>3</sub>), 1.52 (s, 3H, 5-CH<sub>3</sub>), 2.47-2.62 (m, 1H,  $CH(CH_3)_2$ ), 3.86 (br, 1H, 10b-H), 3.96 (d, 1H,  $J = 5.94$  Hz, N-CH-CO), 4.06 (d, 1H,  $J = 9.98$  Hz, 4a-H), 6.87 (d, 1H,  $J = 8.59$  Hz, 7H), 7.46 (dd, 1H,  $J = 8.59$  Hz, 2.02 Hz, 8H), 7.82 (d, 1H,  $J = 2.02$  Hz, 10H);  $^{13}C$  NMR( $CDCl_3$ ):  $\delta = 17.42$  (CH-CH<sub>3</sub>), 19.01 (5-CH<sub>3</sub>), 25.87 (5-CH<sub>3</sub>), 19.80 (CH-CH<sub>3</sub>), 30.98 ( $CH(CH_3)_2$ ), 49.28 (10b-C), 63.73 (N-CH-CO), 77.95 (5-C), 83.03 (4a-C), 104.23 (9-C), 118.25 (7-C), 118.88 (CN), 121.34 (10a-C), 131.30, 133.37 (8-C, 10-C), 155.96 (6a-C), 169.47 (C=O); MS  $m/z$  300 (0.57%, M+), 185 (38.5%), 170 (100%); Anal. Calcd for  $C_{17}H_{20}N_2O_3$  %: C 67.98, H 6.71, N 9.33; found: C 68.28, H 6.91, N 9.04.

**6.1.4.6. (2R,4aS,10bR)-2-isopropyl-1,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (18a).** From **11a** gave **18a**, 45% yellowish solid; mp: 151-153 °C;  $[\alpha]^{20} = (+) 107.17$  ( $c = 0.30$ , in  $CH_2Cl_2$ ); IR (KBr) 2230 (CN)  $cm^{-1}$ , 1741 (lactone)  $cm^{-1}$ ;  $^1H$  NMR ( $CDCl_3$ ):  $\delta = 1.07,$

1.14 (each d, each 3H,  $J = 6.69$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.26 (s, 3H, 5- $\text{CH}_3$ ), 1.56 (s, 3H, 5- $\text{CH}_3$ ), 2.30 (s, 3H, N- $\text{CH}_3$ ), 2.35-2.45 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 3.17 (d, 1H,  $J = 5.93$  Hz, N-CH-CO), 4.07 (d, 1H,  $J = 10.67$  Hz, 4a-H), 4.54 (d, 1H,  $J = 10.67$  Hz, 10b-H), 6.89 (d, 1H,  $J = 8.53$  Hz, 7H), 7.46 (dd, 1H,  $J = 8.53$  Hz, 1.51 Hz, 8H); 7.78 (br, 1H, 10H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta = 19.08$  ( $\text{CH}(\text{CH}_3)_2$ ), 20.00 (5- $\text{CH}_3$ ), 25.99 (5- $\text{CH}_3$ ), 31.82 ( $\text{CH}(\text{CH}_3)_2$ ), 37.94 (N- $\text{CH}_3$ ), 53.19 (10b-C), 70.86 (N-CH-CO), 73.64 (4a-C), 78.01 (5-C), 104.63 (9-C), 118.71 (7-C), 118.95 (CN), 120.27 (10a-C), 131.82, 133.11 (8-C, 10-C), 156.74 (6a-C), 170.72 (C=O); MS  $m/z$  314 (5.36%, M<sup>+</sup>), 243 (18.2%), 170 (100%); Anal. Calcd for  $\text{C}_{18}\text{H}_{22}\text{N}_2\text{O}_3$  %: C 68.77, H 7.05, N 8.91; found: C 68.32, H 7.06, N 8.71.

**6.1.4.7. (2R,4aR,10bS)-2-isopropyl-1,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (18b).** From **11b** gave **18b**, 52% yellowish solid; mp: 93-96 °C;  $[\alpha]^{20} = (-) 125.71$  ( $c = 0.105$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 2225 (CN)  $\text{cm}^{-1}$ , 1757 (lactone)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta = 0.97$ , 1.06 (each d, each 3H,  $J = 6.57$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.37 (s, 3H, 5- $\text{CH}_3$ ), 1.54 (s, 3H, 5- $\text{CH}_3$ ), 1.62-1.81 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.71 (s, 3H, N- $\text{CH}_3$ ), 2.91 (d, 1H,  $J = 10.23$  Hz, N-CH-CO), 3.60 (d, 1H,  $J = 11.74$  Hz, 4a-H), 4.25 (d, 1H,  $J = 11.74$  Hz, 10b-H), 6.83 (d, 1H,  $J = 8.56$  Hz, 7H), 7.44 (dd, 1H,  $J = 8.56$  Hz, 1.52 Hz, 8H); 7.71 (br, 1H, 10H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta = 19.28$  (CH- $\text{CH}_3$ ), 20.38 (5- $\text{CH}_3$ ), 21.17 (CH- $\text{CH}_3$ ), 27.12 (5- $\text{CH}_3$ ), 32.52  $\text{CH}(\text{CH}_3)_2$ , 47.12 (N- $\text{CH}_3$ ), 57.01 (10b-C), 73.27 (N-CH-CO), 76.09 (4a-C), 78.55 (5-C), 104.12 (9-C), 117.80 (7-C), 119.21 (CN), 123.87 (10a-C), 130.93, 132.94 (8-C, 10-C), 155.56 (6a-C), 170.30 (C=O); MS  $m/z$  314 (1.14%, M<sup>+</sup>), 243 (16.2%), 185 (45%), 170 (100%); Anal. Calcd for  $\text{C}_{18}\text{H}_{22}\text{N}_2\text{O}_3$  %: C 68.77, H 7.05, N 8.91; found: C 68.78, H 7.12, N 8.61.

**6.1.5. General Procedure for reaction of epoxide with valinol (19a,b).** A solution of enantiomeric pure epoxide (**4a**, **4b**) (4.97mmol) and L-valinol (4.47mmol) in 50mL 96% ethanol was heated to reflux at 65 °C for 5 days. The solvent was removed under reduced pressure. Purification was done by flash column chromatography (petroleum ether: ethylacetate /8:2)

**6.1.5.1 (2S,3S,4R)-3-hydroxy-4-(1-hydroxy-3-methyl-2-butyl-amino)-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (19a).** Yielded 67% pale yellowish crystal; mp: 141-142 °C;  $[\alpha]^{20} = (+) 13.45$  ( $c = 0.119$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr)

3234 (OH)  $\text{cm}^{-1}$ , 2224 (CN)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 1.00, 1.03 (each d, each 3H,  $J$  = 1.63 Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.22 (s, 3H, 2- $\text{CH}_3$ ), 1.47 (s, 3H, 2- $\text{CH}_3$ ), 1.61 (br, 1H, NH), 1.74-1.91 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 3.02-3.10 (m, 1H, N-CH), 3.48 (d, 1H,  $J$  = 10.04 Hz, 3H), 3.61 (m, 2H,  $\text{CH}_2\text{O}$ ), 3.86 (dd, 1H,  $J$  = 10.04 Hz, 3.26 Hz, 4H), 4.83 (br, 1H, OH), 6.78 (d, 1H,  $J$  = 8.53 Hz, 8H); 7.38 (dd, 1H,  $J$  = 8.53 Hz, 1.51 Hz, 7H), 7.92 (d, 1H,  $J$  = 1.51 Hz, 5H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta$  = 18.79, 18.98, 19.34 (2- $\text{CH}_3$ ,  $\text{CH}(\text{CH}_3)_2$ ), 26.85 (2- $\text{CH}_3$ ), 32.00 ( $\text{CH}(\text{CH}_3)_2$ ), 57.32 ( $\text{CH}_2\text{OH}$ ), 63.94 (4-C), 64.31 (N-CH), 77.21 (3-C), 97.80 (2-C), 103.30 (6-C), 117.94 (8-C), 119.61 (CN), 127.08 (4a-C), 123.38, 132.69 (5-C, 7-C), 156.65 (8a-C); MS  $m/z$  305 (0.72%, M+), 273 (39.8%), 160 (60.6%); Anal. Calcd for  $\text{C}_{17}\text{H}_{24}\text{N}_2\text{O}_3$  %: C 67.08, H 7.95, N 9.20; found: C 67.07, H 7.77, N 9.20.

**6.1.5.2. (2S,3R,4S)-3-hydroxy-4-(1-hydroxy-3-methyl-2-butyl-amino)-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (19b).** Yielded 65% pale yellowish oil;  $[\alpha]^{20} = (+)$  8.65 ( $c = 0.104$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 3446 (OH)  $\text{cm}^{-1}$ , 2226 (CN)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 0.90, 0.94 (each d, each 3H,  $J = 3.91$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.21 (s, 3H, 2- $\text{CH}_3$ ), 1.47 (s, 3H, 2- $\text{CH}_3$ ), 1.70-1.86 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.92-3.00 (m, 1H, N-CH), 3.50 (m, 2H,  $J = 8.65$  Hz,  $\text{CH}_2\text{O}$ ), 3.69 (d, 1H,  $J = 10.05$  Hz, 3H), 3.81 (dd, 1H,  $J = 10.05$  Hz, 3.16 Hz, 4H), 6.77 (d, 1H,  $J = 8.56$  Hz, 8H), 7.36 (dd, 1H,  $J = 8.56$  Hz, 1.84 Hz, 7H), 8.02 (s, 1H, 5H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta$  = 16.96 (CH- $\text{CH}_3$ ), 19.30, 19.58 (2- $\text{CH}_3$ , CH- $\text{CH}_3$ ), 26.63 (2- $\text{CH}_3$ ), 30.28 ( $\text{CH}(\text{CH}_3)_2$ ), 54.71 (4-C), 61.75 ( $\text{CH}_2\text{OH}$ ), 62.24 (N-CH), 73.99 (3-C), 79.58 (2-C), 102.88 (6-C), 117.85 (8-C), 119.48 (CN), 126.41 (4a-C), 132.20, 132.83 (5-C, 7C), 156.81 (8a-C); MS  $m/z$  305 (0.47%, M+), 273 (16.6%), 160 (21.9%); Anal. Calcd for  $\text{C}_{17}\text{H}_{24}\text{N}_2\text{O}_3$  %: C 67.08, H 7.95, N 9.20; found: C 67.12, H 7.88, N 9.10.

**6.1.6. General Procedure for N-methylation of 3-hydroxy-4-(1-hydroxy-3-methyl-2-butyl-amino)-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril.**

The suspension of educt **26a,b**, paraformaldehyde (250mg) and 95% sodium cyanoborohydride ( $\text{NaBH}_3\text{CN}$ ) (326mg) in MeOH (30mL) were adjusted pH to 6 with glacial acetic acid. The suspension was stirred at room temperature overnight. Flash chromatography (petroleum ether: ethylacetate /9:1) were used to purify.

**6.1.6.1. (2S,3S,4R)-3-hydroxy-4-(1-hydroxy-3-methyl-2-butyl-N-methyl-amino)-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (20a).** From **19a** yielded 87% colourless crystal of **20a**; mp 143-144  $^\circ\text{C}$ ;  $[\alpha]^{20} = (+)$  39.45 ( $c = 0.109$ , in

CH<sub>2</sub>Cl<sub>2</sub>); IR (KBr) 3216 (OH) cm<sup>-1</sup>, 2227 (CN) cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ = 1.00, 1.03 (each d, each 3H, J = 6.00 Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.20 (s, 3H, 2-CH<sub>3</sub>), 1.48 (s, 3H, 2-CH<sub>3</sub>), 1.73-1.89 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 2.22 (s, 3H, N-CH<sub>3</sub>), 3.07-3.17 (dt, 1H, J = 10.7 Hz, 4.30 Hz, N-CH), 3.74 (d, 1H, J = 9.85 Hz, 4H), 3.78 (d, 1H, J = 9.85 Hz, 3H), 3.85 (dd, 1H, J = 10.7 Hz, 4.30 Hz, OCH<sub>A</sub>), 3.98 (t, 1H, J = 10.7 Hz, OCH<sub>B</sub>), 6.79 (d, 1H, J = 8.46 Hz, 8H), 7.38 (dd, 1H, J = 8.46 Hz, 1.65 Hz, 7H), 7.96 (d, 1H, J = 1.65 Hz, 5H); <sup>13</sup>C NMR(200MHz, CDCl<sub>3</sub>): δ = 19.03 (CH-CH<sub>3</sub>), 19.95, 20.07 (2-CH<sub>3</sub>, CH-CH<sub>3</sub>), 27.17 (2-CH<sub>3</sub>), 28.64 (N-CH<sub>3</sub>), 32.98 (CH(CH<sub>3</sub>)<sub>2</sub>), 62.43 (CH<sub>2</sub>OH), 66.04 (4-C), 69.08 (3-C), 69.23 (N-CH), 79.63 (C (CH<sub>3</sub>)<sub>3</sub>), 103.28 (6-C), 118.17 (8-C), 119.88 (CN), 125.12 (4a-C), 132.23, 133.01 (5-C, 7-C), 158.01 (8a-C); MS m/z 318 (0.38%, M<sup>+</sup>), 160 (43.2%), 86 (100%); Anal. Calcd for C<sub>18</sub>H<sub>26</sub>N<sub>2</sub>O<sub>3</sub> %: C 67.90, H 8.23, N 8.80; found: C 68.04, H 8.01, N 8.97.

**6.1.6.2. (2S,3R,4S)-3-hydroxy-4-(1-hydroxy-3-methyl-2-butyl-N-methyl-amino)-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (20b).** From **19b** yielded 95% pale yellowish solid of **20b**; mp: 137-138 °C; [α]<sup>20</sup> = (-) 20.91 (c = 0.112, in CH<sub>2</sub>Cl<sub>2</sub>); IR (KBr) 3215 (OH) cm<sup>-1</sup>, 2225 (CN) cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ = 0.92, 1.02 (each d, each 3H, J = 6.82 Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.20 (s, 3H, 2-CH<sub>3</sub>), 1.48 (s, 3H, 2-CH<sub>3</sub>), 2.00-2.16 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 2.32 (s, 3H, N-CH<sub>3</sub>), 2.80-2.88 (m, 1H, N-CH), 3.72 (d, 1H, J = 10.1 Hz, 4H), 3.78 (d, 1H, J = 11.56 Hz, OCH<sub>A</sub>), 3.92 (dd, 1H, J = 11.56 Hz, 2.21 Hz, OCH<sub>B</sub>), 4.08 (d, 1H, J = 10.01 Hz, 3H), 6.78 (d, 1H, J = 8.46 Hz, 8H), 7.37 (dd, 1H, J = 8.46 Hz, 1.90 Hz, 7H), 7.90 (d, 1H, J = 1.90 Hz, 5H); <sup>13</sup>C NMR (200 MHz, CDCl<sub>3</sub>): δ = 17.43 (CH-CH<sub>3</sub>), 19.07(2-CH<sub>3</sub>), 21.22 (CH-CH<sub>3</sub>), 27.08 (2-CH<sub>3</sub>), 28.97(N-CH<sub>3</sub>), 34.90 (CH(CH<sub>3</sub>)<sub>2</sub>), 60.64 (CH<sub>2</sub>OH), 69.81 (3-C), 70.65 (N-CH), 80.21 (2-C), 103.07 (6-C), 117.93 (8-C), 119.83 (CN), 125.65 (4a-C), 132.20, 133.12 (5-C, 7-C), 157.99 (8a-C); MS m/z 318 (0.35%, M<sup>+</sup>), 160 (43.5%), 86 (100%); Anal. Calcd for C<sub>18</sub>H<sub>26</sub>N<sub>2</sub>O<sub>3</sub> %: C 67.90, H 8.23, N 8.80; found: C 67.93, H 8.14, N 8.73.

**6.1.7. General Procedure for ring closure via intramolecular ether formation (21a,b).** A suspension of **19a,b** (600mg), trimethylamine (0.44mL) and trimethylamine hydrochloride (0.2mg) in toluene (2mL), cooled at 0 °C in an ice-acetone bath, than added cooled solution of methane sulfonyl chloride (0.5mg) in toluene and stirred at 0°C for 1.5 hr. The solvent was removed under reduced pressure. The residue was dissolved in dry tetrahydrofuran 10mL and 60% sodium hydride dispersion in mineral



oil (6.25 mmol) was added bit by bit, and then stirred overnight. Excess of sodium hydride was destroyed by adding water, organic phase was separated. The water phase was extracted two times with ethyl acetate. Combined organic phase was dried with anhydrous sodium sulphate and brought to dryness. The residue was purified by flash column chromatography (petroleum ether: ethylacetate /9:1). Surprisingly, in case of the tertiary amine (**20a,b**), the cyclisation failed and the respective chloro derivatives **22a,b** were obtained.

**6.1.7.1. (2S,4aS,10bR)-2-isopropyl-5,5-dimethyl-1,4a,5,10b-tetrahydro-3H [1]benzopyrano-[3, 4-b] [1, 4]oxazine-9-carbonitril (21a).** Yielded 50% yellowish oil; IR (KBr) 2224 (CN)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 0.98, 1.02 (each d, each 3H,  $J$  = 2.66 Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.25 (s, 3H, 5- $\text{CH}_3$ ), 1.66 (s, 3H, 5- $\text{CH}_3$ ), 1.76- 1.83 (m, 1H, 2H), 1.99-2.15 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.50 (d, 1H,  $J$  = 6.83 Hz, 4a-H), 2.56 (d, 1H,  $J$  = 6.83 Hz, 10b-H), 3.69 (d, 2H,  $J$  = 4.68 Hz, 3- $\text{CH}_2$ ), 6.86 (d, 1H,  $J$  = 8.37 Hz, 7H), 7.47 (dd, 1H,  $J$  = 8.37 Hz, 2.02 Hz, 8H), 7.57 (d, 1H,  $J$  = 2.02Hz, 10H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta$  = 19.19, 19.27 ( $\text{CH}(\text{CH}_3)_2$ ), 24.47 (5- $\text{CH}_3$ ), 26.12 (5- $\text{CH}_3$ ), 31.32 ( $\text{CH}(\text{CH}_3)_2$ ), 37.45 (10b-C), 45.98 (3-C), 48.90 (2-C), 73.95 (5-C), 74.26 (4a-C), 103.85 (9-C), 119.21(7-C), 123.98 (10a-C), 132.51, 132.70 (8-C, 10-C), 156.12 (6a-C); MS  $m/z$  269 (54.4%,  $M^+$ ), 170 (22.9%), 157 (75.4%); Anal. Calcd for  $\text{C}_{17}\text{H}_{22}\text{N}_2\text{O}_2$  %: C 71.30, H 7.78, N 9.78; found: C 71.12, H 7.65, N 9.85

**6.1.7.2. (2S,4aR,10bS)-2-isopropyl-5,5-dimethyl-1,4a,5,10b-tetrahydro-3H [1]benzopyrano-[3, 4-b] [1, 4]oxazine-9-carbonitril (21b).** Yielded 51% yellowish oil; IR (KBr) 2224 (CN), 1572 (NH)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 0.98, 1.10 (each d, each 3H,  $J$  = 6.26 Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.25 (s, 3H, 5- $\text{CH}_3$ ), 1.44 (s, 3H, 5- $\text{CH}_3$ ), 2.15-2.29 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.36 (dd, 1H,  $J$  = 10.36 Hz, 2.91 Hz, 2H), 3.08 (d, 1H,  $J$  = 10.11 Hz, 4a-H), 3.62 (dd, 1H,  $J$  = 11.75 Hz, 3.03 Hz, 3 $\text{CH}_2\text{A}$ ), 3.80 (d, 1H,  $J$  = 10.11 Hz, 10b-H), 4.16 (d, 1H,  $J$  = 11.87 Hz, 3 $\text{CH}_2\text{B}$ ), 6.79 (d, 1H,  $J$  = 8.59Hz, 7H), 7.39 (dd, 1H,  $J$  = 8.59 Hz, 1.51 Hz, 8H), 7.69 (d, 1H,  $J$  = 1.51 Hz, 10H);  $^{13}\text{C}$  NMR( $\text{CDCl}_3$ ):  $\delta$  = 19.86 ( $\text{CH}-\text{CH}_3$ ), 20.70, 20.81 ( $\text{CH}-\text{CH}_3$ , 5- $\text{CH}_3$ ), 25.80 ( $\text{CH}(\text{CH}_3)_2$ ), 26.69 (5- $\text{CH}_3$ ), 45.36 (10b-C), 58.77 (2-C), 69.40 (3-C), 78.79 (5-C), 81.88 (4a-C), 103.41 (9-C), 117.75 (7-C), 119.35 (CN), 124.10 (10a-C), 130.06, 132.51 (8-C, 10-C), 156.44 (6a-C); MS  $m/z$  286 (11.9%,  $M^+$ ), 243 (100%), 170 (25%); Anal. Calcd for  $\text{C}_{17}\text{H}_{22}\text{N}_2\text{O}_2$  %: C 71.30, H 7.74, N 9.78; found: C 71.45, H 7.64, N 9.61

**6.1.7.3. (2S,3S,4R)-4-(N-(1-chloro-3-methyl-2-butyl-N-methyl-amino)-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (22a).** From **20a** yielded 55% yellowish oil;  $[\alpha]^{20} = (-) 7.77$  ( $c = 0.103$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 3479 (OH)  $\text{cm}^{-1}$ , 2226 (CN)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta = 0.90, 1.02$  (each d, each 3H, each  $J = 6.63$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.18 (s, 3H, 2- $\text{CH}_3$ ), 1.48 (s, 3H, 2- $\text{CH}_3$ ), 1.90-2.05 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.51 (s, 3H, N- $\text{CH}_3$ ), 2.95 (br, 2H,  $\text{CH}_2\text{Cl}$ ) 3.24 (br, 1H, OH), 3.61 (d, 1H,  $J = 10.17$  Hz, 3H), 3.69 (d, 1H,  $J = 10.17$  Hz, 4H), 3.97-4.05 (m, 1H, N-CH), 6.80 (d, 1H,  $J = 8.50$  Hz, 8H), 7.36 (dd, 1H,  $J = 8.50$  Hz, 1.77 Hz, 7H), 7.64 (br, 1H, 5H);  $^{13}\text{C}$  NMR (200MHz,  $\text{CDCl}_3$ ):  $\delta = 16.64$  (2- $\text{CH}_3$ ), 18.73 (CH- $\text{CH}_3$ ), 20.35 (2- $\text{CH}_3$ ), 26.84 (CH- $\text{CH}_3$ ), 31.84 ( $\text{CH}(\text{CH}_3)_2$ ), 38.84 (N- $\text{CH}_3$  at 60 °C 300 MHz), 60.18 ( $\text{CH}_2\text{Cl}$  at 60 °C 300 MHz), 63.50 (4-C), 68.54 (N-CH), 70.01 (3-C), 79.61 ( $\text{C}(\text{CH}_3)_3$ ), 103.28 (6-C), 118.48 (8-C), 119.38 (CN), 123.46 (4a-C), 132.39, 132.44 (5-C, 7-C), 157.57 (8a-C); MS  $m/z$  336 (0.43%, M+), 229 (56.5%), 160 (35.1%); Anal. Calcd for  $\text{C}_{18}\text{H}_{25}\text{ClN}_2\text{O}_2$  %: C 64.18, H 7.48, N 8.32; Cl 10.52 found: C 63.61, H 7.17, N 7.79, Cl 9.93.

**6.1.7.4. (2S,3R,4S)-4-(N-(1-chloro-3-methyl-2-butyl-N-methyl-amino)-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-6-carbonitril (22b).** From **20b** yielded 60% yellowish oil;  $[\alpha]^{20} = (-) 31.98$  ( $c = 0.111$ , in  $\text{CH}_2\text{Cl}_2$ ); IR (KBr) 3310 (OH)  $\text{cm}^{-1}$ , 2227 (CN)  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta = 1.00, 1.08$  (each d, each 3H, each  $J = 6.70$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.19 (s, 3H, 2- $\text{CH}_3$ ), 1.50 (s, 3H, 2- $\text{CH}_3$ ), 1.97-2.12 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.22 (s, 3H, N- $\text{CH}_3$ ), 3.12 (dd, 1H,  $J = 13.64$  Hz, 10.23 Hz,  $\text{CH}_2\text{ACl}$ ), 3.40 (dd, 1H,  $J = 14.08$  Hz, 7.14Hz,  $\text{CH}_2\text{BCl}$ ), 3.59 (d, 1H,  $J = 10.04$  Hz, 3H), 3.71 (d, 1H,  $J = 10.04$  Hz, 4H), 4.02-4.11 (m, 1H, N-CH), 6.84 (d, 1H,  $J = 8.08$  Hz, 8H), 7.37 (s, 1H, 5H), 7.39 (d, 1H,  $J = 8.08$  Hz, 7H);  $^{13}\text{C}$  NMR(200MHz,  $\text{CDCl}_3$ ):  $\delta = 16.80$  (CH- $\text{CH}_3$ ), 18.66 (2- $\text{CH}_3$ ), 20.30 (CH- $\text{CH}_3$ ), 26.88 (2- $\text{CH}_3$ ), 32.07 ( $\text{CH}(\text{CH}_3)_2$ ), 34.96 (N- $\text{CH}_3$  at 60 °C 300 MHz), 64.07 ( $\text{CH}_2\text{Cl}$  at 60 °C 300 MHz), 65.37 (4-C), 69.67 (N-CH), 70.89 (3-C), 79.74 ( $\text{C}(\text{CH}_3)_3$ ), 103.17 (6-C), 118.83 (8-C), 119.41 (CN), 122.48 (4a-C), 132.20, 132.59 (5-C, 7-C), 157.77 (8a-C); MS  $m/z$  336 (0.51%, M+), 229 (37.08%), 160 (71.2%); Anal. Calcd for  $\text{C}_{18}\text{H}_{25}\text{ClN}_2\text{O}_2$  %: C 64.18, H 7.48, N 8.32, Cl 10.52; found: C 64.25, H 7.57, N 7.80, Cl 9.01.

## 6.2. Biological Essay

The human T-lymphoblast cell line CCRF-CEM and the multidrug resistant CEM/vcr1000 cell line were provided by V. Gekeler (Byk Gulden, Konstanz, Germany). The resistant CEM/vcr1000 line was obtained by stepwise selection in vincristine containing medium. Cells were kept under standard culture conditions (RPMI1640 medium supplemented with 10% fetal calf serum). P-gp-expressing resistant cell line was cultured in presence of 1000ng/ml vincristine. One week prior to the experiments cells were transferred into medium without selective agents or antibiotics. Briefly, cells were pelleted, the supernatant was removed by aspiration and cells were resuspended at a density of  $1 \times 10^6$ / mL in PRMI1640 medium containing  $3 \mu\text{mol/l}$  daunomycin. Cell suspensions were incubated at  $37^\circ\text{C}$  for 30 min. After this time a steady state of daunorubicin accumulation was reached. Tubes were chilled on ice and cells were pelleted at  $500 \times g$ . Cells were washed once in RPMI1640 medium to remove extracellular daunorubicin. Subsequently, cells were resuspended in medium prewarmed to  $37^\circ\text{C}$ , containing either no modulator or chemosensitizer at various concentrations ranging from 3nM to  $500 \mu\text{M}$ , depending on solubility and expected potency of the modifier. Generally, 8 serial dilutions were tested for each modulator. After 1, 2, 3 and 4 min aliquots of the incubation mixture were drawn and pipetted into 4 volumes of ice cold stop solution (RPMI1640 medium containing verapamil at a final concentration of  $100 \mu\text{M}$ ). Parental CCRF-CEM cells were used to correct for simple membrane diffusion, which was less than 3% of the efflux rates observed in resistant cells. Samples drawn at the respective time points were kept in an ice water bath and measured within one hour on a Becton Dickinson FACS Calibur (Becton Dickinson, Heidelberg, Germany) flow cytometer as described. Dose response curves were fitted to the data points using non-linear least squares and  $\text{IC}_{50}$  values were calculated as described by Chiba *et al* [20].  $\text{IC}_{50}$  values of individual compounds the average of at least triplicate determinations (Table 1). A CV of below 20% was obtained in all determinations.

## 6.3. Computational Methods

### 6.3.1. Hansch Analysis

Molecular descriptors supplied by the program MOE (atom and bond counts,

connectivity indices, partial charge descriptors, pharmacophore feature descriptors, calculated physical property descriptors, etc.) were computed. QSAR-Contingency [42], a statistical application in MOE, was used for selection of relevant descriptors. PLS analysis was performed to determine the relationship between 2D molecular descriptors and biological activity of the compounds. The predictive value of the model was determined by leave one out cross validation (LOO).

### 6.3.2. GRIND

Molecular Interaction Fields (MIF) were calculated using GRID based fields in Pentacle [29] using four different probes: Dry probe to represent hydrophobic interactions, O sp<sup>2</sup> carbonyl oxygen probe to represent H-bond donor feature of the molecules, N1 probe to represent –NH which is a neutral flat probe as an H-bond acceptor in the molecules and the TIP probe that represents the shape of the molecule and refers to steric hot spots. The regions with highest MIF were extracted by applying the AMANDA algorithm [43] that uses the intensity of the field at a node and the mutual node-node distances between the chosen nodes. At each point, the interaction energy ( $E_{xyz}$ ) was calculated as a sum of Lennard-Jones energy ( $E_{lj}$ ), Hydrogen bond ( $E_{hb}$ ) and Electrostatic ( $E_{el}$ ) interactions.

$$E_{xyz} = \sum E_{lj} + \sum E_{el} + \sum E_{hb}$$

Default values of probe cutoff (DRY= -0.5, O= -2.6, N1= -4.2, TIP= -0.74) was used for discretization of MIF. Nodes with an energy value below this cutoff were discarded. The Consistently Large Auto and Cross Correlation (CLACC) algorithm [29] was used for encoding the prefiltered nodes into GRIND thus producing most consistent variables as compared to MACC [44]. The values obtained from the analysis can be represented directly in correlogram plots, where the product of node-node energies is reported versus the distance separating the nodes. Highest energy product can be define by the same probe (obtaining four auto correlograms: Dry-Dry, O-O, N1-N1 and TIP-TIP) and by pair of probes (obtaining six cross correlograms: Dry-O, Dry-N1, Dry-TIP, O-N1, O-TIP, and N1-TIP). The QSAR model obtained was assessed by means of the  $q^2$  and standard deviation error of prediction (SDEP). Classical leave one out (LOO) method was applied to calculate  $q^2$  values.

**Acknowledgement:** We are grateful to the Austrian Science Fund for financial support (grant SFB 3502 and SFB 3509). Ishrat Jabeen thanks the Higher Education Commission of Pakistan for financial support.

### Reference

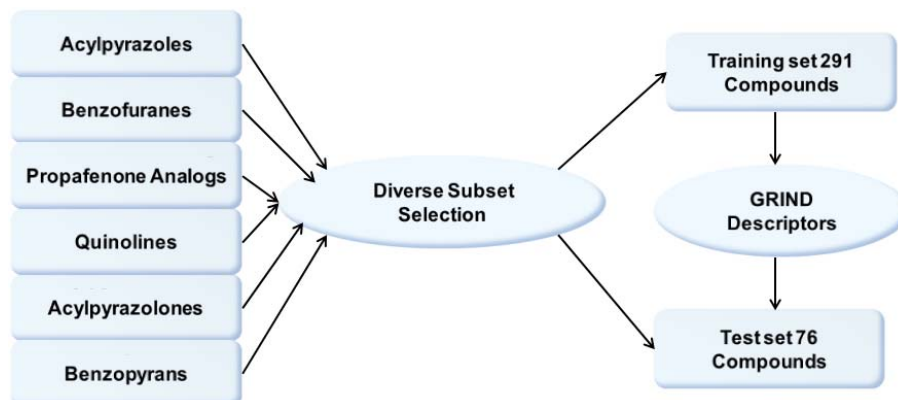
- [1] C. Wandel, R.B. Kim, S. Kajiji, P. Guengerich, G.R. Wilkinson, A.J. Wood, P-glycoprotein and cytochrome P-450 3A inhibition: dissociation of inhibitory potencies, *Cancer Res*, 59 (1999) 3944-3948.
- [2] M.M. Gottesman, V. Ling, The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research, *FEBS Lett*, 580 (2006) 998-1009.
- [3] J.M. Ford, Experimental reversal of P-glycoprotein-mediated multidrug resistance by pharmacological chemosensitisers, *Eur J Cancer*, 32A (1996) 991-1001.
- [4] S. Ekins, Predicting undesirable drug interactions with promiscuous proteins in silico, *Drug Discov Today*, 9 (2004) 276-285.
- [5] G.F. Ecker, Inhibitors of P-glycoprotein-Hit identification and lead optimization, in: N.A. Colabufo (Ed.) *Multidrug Resistance: Biological and Pharmaceutical Advances in Antitumour*, Research Signpost, Kerala, India, 2008, pp. 243-259.
- [6] M.M. Gottesman, T. Fojo, S.E. Bates, Multidrug resistance in cancer: role of ATP-dependent transporters, *Nat Rev Cancer*, 2 (2002) 48-58.
- [7] R. Hiessbock, C. Wolf, E. Richter, M. Hitzler, P. Chiba, M. Kratzel, G. Ecker, Synthesis and in vitro multidrug resistance modulating activity of a series of dihydrobenzopyrans and tetrahydroquinolines, *J Med Chem*, 42 (1999) 1921-1926.
- [8] P. Chiba, W. Holzer, M. Landau, G. Bechmann, K. Lorenz, B. Plagens, M. Hitzler, E. Richter, G. Ecker, Substituted 4-acylpyrazoles and 4-acylpyrazolones: synthesis and multidrug resistance-modulating activity, *J Med Chem*, 41 (1998) 4001-4011.
- [9] D. Kaiser, M. Smiesko, S. Kopp, P. Chiba, G.F. Ecker, Interaction field based and hologram based QSAR analysis of propafenone-type modulators of multidrug resistance, *Med Chem*, 1 (2005) 431-444.
- [10] I. Jabeen, P. Wetwitayaklung, F. Klepsch, Z. Parveen, P. Chiba, G.F. Ecker, Probing the stereoselectivity of P-glycoprotein-synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-b][1,4]oxazines, *Chem Commun (Camb)*, 47 (2011) 2586-2588.
- [11] M. Pastor, G. Cruciani, I. McLay, S. Pickett, S. Clementi, GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors, *J Med Chem*, 43 (2000) 3233-3243.
- [12] F. Fontaine, M. Pastor, I. Zamora, F. Sanz, Anchor-GRIND: filling the gap between standard 3D QSAR and the GRid-INdependent descriptors, *J Med Chem*, 48 (2005) 2687-2694.
- [13] S. Sciabola, R.V. Stanton, J.E. Mills, M.M. Flocco, M. Baroni, G. Cruciani, F. Perruccio, J.S. Mason, High-throughput virtual screening of proteins using GRID molecular interaction fields, *J Chem Inf Model*, 50 155-169.
- [14] G. Cruciani, M. Pastor, W. Guba, VolSurf: a new tool for the pharmacokinetic optimization of lead compounds, *Eur J Pharm Sci*, 11 Suppl 2 (2000) S29-39.
- [15] C. Obiol-Pardo, J. Gomis-Tena, F. Sanz, J. Saiz, M. Pastor, A multiscale simulation system for the prediction of drug-induced cardiotoxicity, *J Chem Inf Model*, 51 483-492.

- [16] G. Cruciani, E. Carosati, B. De Boeck, K. Ethirajulu, C. Mackie, T. Howe, R. Vianello, MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist, *J Med Chem*, 48 (2005) 6970-6979.
- [17] J.D. Godfrey, R.H. Mueller, T.C. Sedergran, N. Soundararajan, V.J. Colandrea, Improved synthesis of aryl 1,1-dimethylpropargyl ethers, *Tetrahedron Letters*, 35 (1994) 6405-6408.
- [18] N.H. Lee, A.R. Muci, E.N. Jacobsen, Enantiomerically Pure Epoxychromans via Asymmetric Catalysis, *Tetrahedron Letters*, 32 (1991) 5055-5058.
- [19] F.M. Callahan, G.W. Anderson, R. Paul, J. Zimmerman, The Tertiary Butyl Group as a Blocking Agent for Hydroxyl, Sulfhydryl and Amido Functions in Peptide Synthesis, *J. Am. Chem. Soc.*, 85 (1963) 201-207.
- [20] P. Chiba, G. Ecker, D. Schmid, J. Drach, B. Tell, S. Goldenberg, V. Gekeler, Structural requirements for activity of propafenone-type modulators in P-glycoprotein-mediated multidrug resistance, *Mol Pharmacol*, 49 (1996) 1122-1130.
- [21] K. Sugano, M. Kansy, P. Artursson, A. Avdeef, S. Bendels, L. Di, G.F. Ecker, B. Faller, H. Fischer, G. Gerebtzoff, H. Lennernaes, F. Senner, Coexistence of passive and carrier-mediated processes in drug transport, *Nat Rev Drug Discov*, 9 597-614.
- [22] T.A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *Journal of Computational Chemistry*, 17 (1996) 490-519.
- [23] K. Pleban, C. Hoffer, S. Kopp, M. Peer, P. Chiba, G.F. Ecker, Intramolecular distribution of hydrophobicity influences pharmacological activity of propafenone-type MDR modulators, *Arch Pharm (Weinheim)*, 337 (2004) 328-334.
- [24] I. Pajeva, M. Wiese, Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: a comparative molecular field analysis study, *J Med Chem*, 41 (1998) 1815-1826.
- [25] I.K. Pajeva, M. Wiese, A Comparative Molecular Field Analysis of Propafenone-type Modulators of Cancer Multidrug Resistance, *Quantitative Structure-Activity Relationships*, 17 (1998) 301-312.
- [26] S.G. Aller, J. Yu, A. Ward, Y. Weng, S. Chittaboina, R. Zhuo, P.M. Harrell, Y.T. Trinh, Q. Zhang, I.L. Urbatsch, G. Chang, Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding, *Science*, 323 (2009) 1718-1722.
- [27] M. Wiese, I.K. Pajeva, Structure-activity relationships of multidrug resistance reversers, *Curr Med Chem*, 8 (2001) 685-713.
- [28] C. Tmej, P. Chiba, M. Huber, E. Richter, M. Hitzler, K.J. Schaper, G. Ecker, A combined Hansch/Free-Wilson approach as predictive tool in QSAR studies on propafenone-type modulators of multidrug resistance, *Arch Pharm (Weinheim)*, 331 (1998) 233-240.
- [29] Á. Durán, M. Pastor, An advanced tool for computing and handling GRid-INdependent. Descriptors. User Manual Version 1.06, (2011).
- [30] G. Ecker, M. Huber, D. Schmid, P. Chiba, The importance of a nitrogen atom in modulators of multidrug resistance, *Mol Pharmacol*, 56 (1999) 791-796.
- [31] M. Baroni, G. Costantino, G. Cruciani, D. Riganelli, R. Valigi, S. Clementi, Generating Optimal Linear PLS Estimations (GOLPE): An Advanced Chemometric Tool for Handling 3D-QSAR Problems, *Quantitative Structure-Activity Relationships*, 12 (1993) 9-20.
- [32] P. Chiba, M. Hitzler, E. Richter, M. Huber, C. Tmej, E. Giovagnoni, G. Ecker, Studies on Propafenone-type Modulators of Multidrug Resistance III: Variations on the Nitrogen, *Quantitative Structure-Activity Relationships*, 16 (1997) 361-366.

- [33] T. Langer, M. Eder, R.D. Hoffmann, P. Chiba, G.F. Ecker, Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model, *Arch Pharm (Weinheim)*, 337 (2004) 317-327.
- [34] A. Seelig, A general pattern for substrate recognition by P-glycoprotein, *Eur J Biochem*, 251 (1998) 252-261.
- [35] A. Seelig, How does P-glycoprotein recognize its substrates?, *Int J Clin Pharmacol Ther*, 36 (1998) 50-54.
- [36] P. Crivori, B. Reinach, D. Pezzetta, I. Poggesi, Computational models for identifying potential P-glycoprotein substrates and inhibitors, *Mol Pharm*, 3 (2006) 33-44.
- [37] G. Cianchetta, R.W. Singleton, M. Zhang, M. Wildgoose, D. Giesing, A. Fravolini, G. Cruciani, R.J. Vaz, A pharmacophore hypothesis for P-glycoprotein substrate recognition using GRIND-based 3D-QSAR, *J Med Chem*, 48 (2005) 2927-2935.
- [38] F. Broccatelli, E. Carosati, A. Neri, M. Frosini, L. Goracci, T.I. Oprea, G. Cruciani, A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields, *J Med Chem*, (2011).
- [39] J. Boccard, F. Bajot, A. Di Pietro, S. Rudaz, A. Boumendjel, E. Nicolle, P.A. Carrupt, A 3D linear solvation energy model to quantify the affinity of flavonoid derivatives toward P-glycoprotein, *Eur J Pharm Sci*, 36 (2009) 254-264.
- [40] A. Boumendjel, C. Beney, N. Deka, A.M. Mariotte, M.A. Lawson, D. Trompier, H. Baubichon-Cortay, A. Di Pietro, 4-Hydroxy-6-methoxyaurones with high-affinity binding to cytosolic domain of P-glycoprotein, *Chem Pharm Bull (Tokyo)*, 50 (2002) 854-856.
- [41] A. Boumendjel, A. Di Pietro, C. Dumontet, D. Barron, Recent advances in the discovery of flavonoids and analogs with high-affinity binding to P-glycoprotein responsible for cancer cell multidrug resistance, *Med Res Rev*, 22 (2002) 512-529.
- [42] R.V. Hogg, E.A. Tanis, in: *Probability and Statistical Inference*, Macmillan Publishing, New York, 1993.
- [43] A. Duran, G.C. Martinez, M. Pastor, Development and validation of AMANDA, a new algorithm for selecting highly relevant regions in Molecular Interaction Fields, *J Chem Inf Model*, 48 (2008) 1813-1823.
- [44] M. Clementi, S. Clementi, G. Cruciani, M. Pastor, Chemometric Detection of Binding Sites of 7TM Receptors, in: K. Gundertofte, Jorgensen, F. S (Ed.) *Molecular Modelling and Prediction of Bioreactivity*, Kluwer Academic/Plenum Publishers, New York, 2000, pp. 207-212.

**Development of a Predictive 3D-QSAR Model for a Structurally Diverse Set of Inhibitors of P-glycoprotein (P-gp)**

Page 104-116



In this chapter the GRIND approach was used to build a predictive 3D-QSAR model for an extended data set of P-glycoprotein inhibitors belonging to different chemical scaffolds.

**Contents****Introduction****Data set****Diverse Subset Selection****Results and Discussion****GRID-Independent Molecular Descriptor Analysis****Conclusions****References**

☞ Appendix available: Page 197-207

Smile codes and biological activities of the data set



## Development of a Predictive 3D-QSAR Model for a Structurally Diverse Set of Inhibitors of P-glycoprotein (P-gp)

### Introduction

About two decades ago propafenone, which is originally an anti-arrhythmic agent of the class Ic, has been identified as a promising P-glycoprotein inhibitor.<sup>1,2</sup> Later on some of its derivatives were among few highly effective agents involved in resensitization of multidrug-resistant tumor cells.<sup>3</sup> In order to probe structural features important for P-gp inhibitory activity and to design promising inhibitors of this efflux pump, extensive SAR and QSAR studies have been performed on propafenone analogs including acylpyrazoles, acylpyrazolones, dihydrobenzopyrans, tetrahydroquinolines, benzophenones and benzofuranes.<sup>3-6</sup> Hansch analyses,<sup>7-11</sup> CoMFA, CoMSIA, HQSAR studies,<sup>12</sup> pharmacophore modeling<sup>13</sup> as well as similarity-based descriptors (SIBAR)<sup>14,15</sup> were computed for building local models across a particular chemical scaffold. Most of these studies point towards the importance of hydrogen bond acceptors and their strength, a certain distance between aromatic moieties and hydrogen bond acceptors, as well as the influence of global physicochemical parameters, such as lipophilicity and molar refractivity.<sup>10,11,16</sup>

Ligand based P-gp inhibition models, reviewed in the 1<sup>st</sup> chapter of this thesis, show that their performance diminished when tested against nonlocal external test sets. This represents one of the major drawbacks of classical QSAR models.<sup>13,17-22</sup> 3D-QSAR methods on the other hand need a proper alignment of the molecules. Moreover, pharmacophore-based as well as CoMFA and CoMSIA models do not consider ADME (e.g. membrane permeability) properties of the compounds. This can be overcome by using descriptors derived from molecular interaction fields (MIF), such as Volsurf or GRIND,<sup>23</sup> which are alignment-independent and thus allow the analysis of structurally diverse data series. Recently, Broccatelli *et al.*, used MIF based descriptors to predict optimal ADME properties and to generate a pharmacophore model to identify more potent and nontoxic inhibitors of P-gp.<sup>24</sup> In their model 3D shape of the molecules along with one hydrogen bond acceptor atom and one large hydrophobic region appears as a basic pharmacophoric pattern for P-gp inhibitors across a data set of diverse chemical scaffolds. Within this study we used MIF based descriptors (GRIND) to identify 3D

pharmacophoric features and pinpoint their mutual distances by using a training set of structurally diverse propafenone-type P-gp inhibitors.

#### Data sets

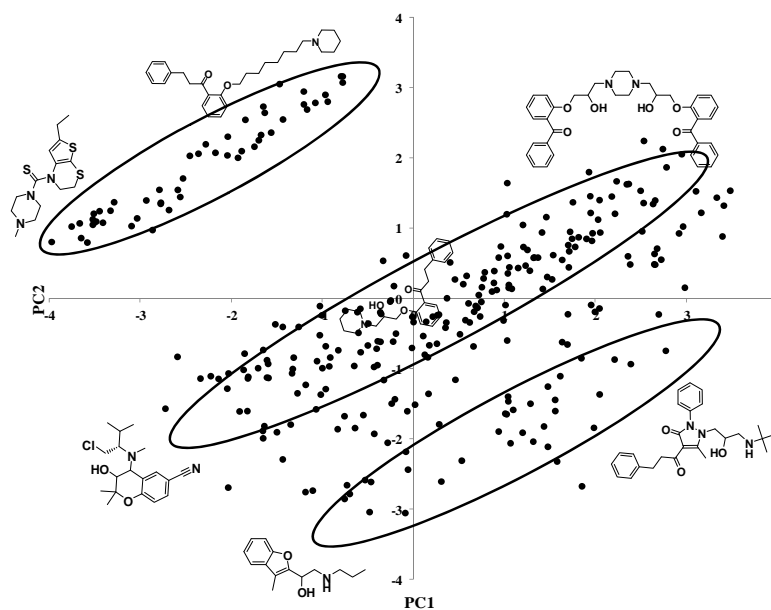
The data set used consists of 375 inhibitors of P-gp, including some previously published propafenones,<sup>3,4,9</sup> acylpyrazoles, acylpyrazolones<sup>10</sup> dihydrobenzopyrans,<sup>25</sup> tetrahydroquinolines,<sup>11</sup> benzofuranes<sup>7</sup> as well as some newly synthesized benzophenone derivatives. Smile codes and biological activity values of the compounds are provided in appendix A3. The diverse subset selection tool in MOE was applied to assign a ranking order to the entries in the database. 185 2D-descriptors together with P-gp inhibitory activity values were used to calculate the pair wise distances between all compounds. A subset (80%) which is diverse with respect to chemical structures (e.g., 2D molecular descriptors) as well as P-gp inhibitory activities was used as training set and the remaining compounds comprised the test set (20%).

#### Results and Discussion

Extended 3D conformations of the molecules were generated by CORINA.<sup>26</sup> All compounds were modeled in their neutral form as the role of the nitrogen atom as a hydrogen bond acceptor has already been demonstrated for P-gp inhibitors,<sup>5</sup> suggesting that molecules should be modeled in their neutral state. The software package Pentacle (v. 1.06)<sup>27</sup> was used to construct 3D-QSAR models using GRIND descriptors. (See methods section of chapter 2 and 3).

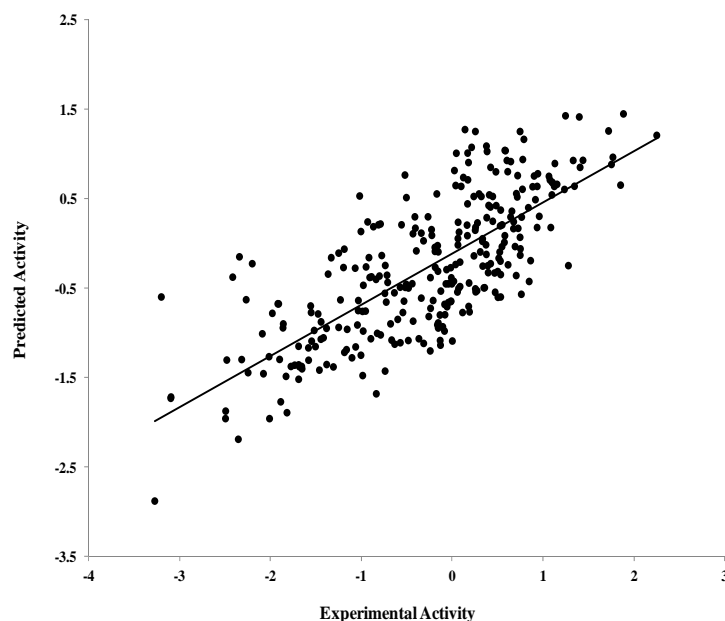
Structural variance of the data was analyzed with principal component analysis (PCA) performed on the complete set of GRIND descriptors. The first two principal components explain about 36% of the chemical variance of the data. The PCA on the data matrix showed that the whole training set consists of structurally diverse compounds divided into three major clusters which differ with respect to their shape/size and number of hydrogen bond donor groups (Figure 1). Compounds in cluster 1, which is located in the upper left corner of the plot do not contain any hydrogen bond donor (-OH, -NH) group. In the central cluster, which consists of simple propafenone derivatives and dihydrobenzopyrans, all compounds contain one hydrogen bond donor group. Compounds in the 3<sup>rd</sup> cluster located just below the central cluster are acylpyrazoles and acylpyrazolones, all possessing two hydrogen bond donors. In

addition, a general trend of decrease in size and flexibility has been observed from the upper right to the lower left section of each cluster (Figure 1). Thus, principal component analyses groups the compounds according to their chemical variability with respect to their shape and size as well as number of hydrogen bond donor groups.



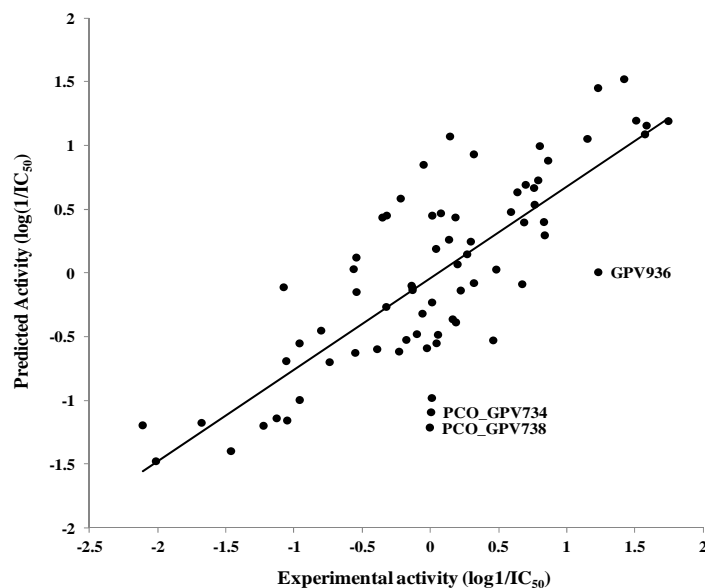
**Figure 1.** PCA score plot showing that the whole data set is divided into three major clusters, which are different with respect to number of hydrogen bond donor groups. Within each cluster the flexibility and size of the ligands decreases from the upper right corner to the lower left part.

In order to explore the pharmacophoric pattern of ligand-protein interaction across structurally different series of P-gp inhibitors, PLS analysis correlating the activity with the complete set of GRIND variables (750) was carried out using the AMANDA algorithm implemented in Pentacle (v 1.06).<sup>27</sup> This resulted in a two-latent variable model with an  $r^2 = 0.54$  and a leave one out cross-validated (LOO)  $q^2$  value of 0.45, which was quite unsatisfactory. Thus, variable selection was applied to reduce the number of active variables by using FFD factorial selection.<sup>28</sup> Finally, the performance of the model increased ( $r^2$  of 0.61,  $q^2 = 0.56$ , standard error of prediction 0.68) when the number of active variables decreases from 552 to 422. Figure 2 shows the plot of the experimental versus calculated biological activities.



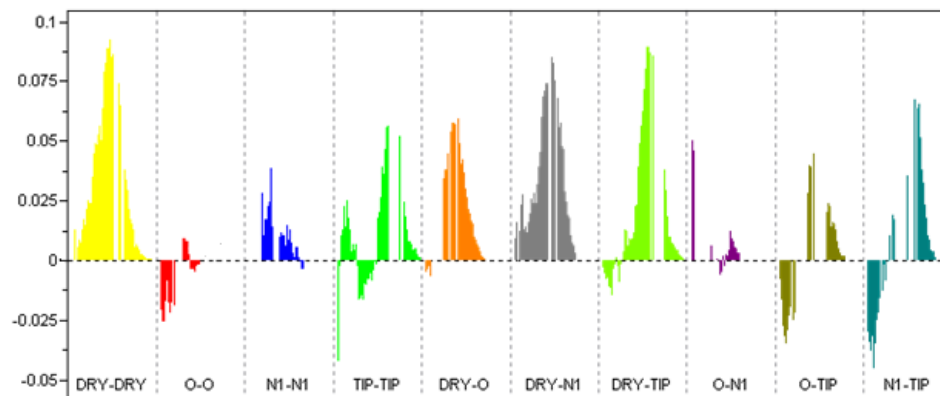
**Figure 2.** Plot of observed *vs.* predicted MDR-modulating activity ( $\log_1/IC_{50}$ ) of compounds in the training set, predicted values were obtained by leave-one-out cross-validation procedure.

The entire training set could be modeled nicely although some of the compounds in the training set reveal residual values above one log unit. This might be due to the quite large diversity in chemical structures in the training set. Figure 3 shows the experimental *vs.* predicted biological activities of the external test set. Biological activities of all compounds could be predicted with a difference of less than one log unit, except for PCO\_GPV738 (obs: -0.004; pred: -1.21) and PCO\_GP734 (obs: 0.005; pred: -1.93), which belong to the series of dihydrobenzopyrans, and for GPV936 (obs 1.23; pred: 0.005), which is a propafenone derivative. All three compounds have been predicted as being about one log unit more active than observed. However, the overall performance of the model is quite satisfactory, which indicates that a similar pharmacoporic pattern across different chemical scaffolds of P-gp inhibitors may exist.



**Figure 3.** Experimental vs. predicted biological activity values of the external test set compounds.

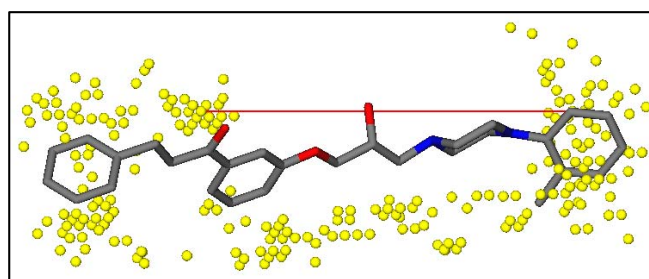
Analysis of the PLS coefficients profile allows to identify those descriptors which exhibit strong contribution to the model. According to the correlogram plot given in figure 4, Dry-Dry, Dry-TIP, Dry-N1 and Dry-O are GRIND variables that contribute most in explaining the variance in activity values.



**Figure 4.** PLS Coefficient correlograms showing the descriptors which are directly (positive values) or inversely (negative values) correlated to the biological activity. The activity predominantly increases with the increase in (DRY-DRY), (DRY-TIP), (DRY-N1) and (DRY-O) descriptor values.

The DRY-DRY correlogram is important for explaining the model, as all coefficients are positive and well spread in distance between 2.00-15.60 Å. The correlogram indicates that two large hydrophobic regions at a distance of 14.40-14.80 Å are present in all compounds exhibiting an activity value (IC<sub>50</sub>) below 1µM, while in least active

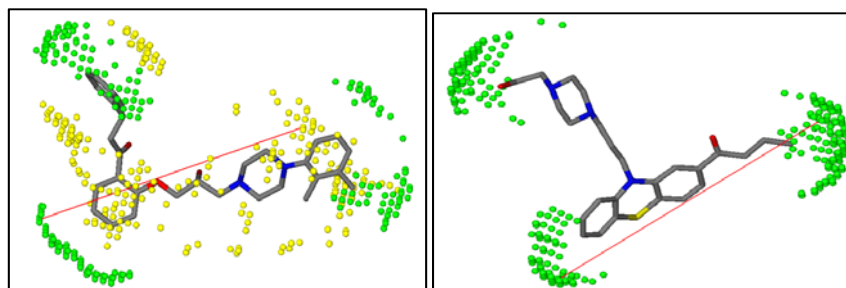
compounds ( $IC_{50} > 1$ ) the distance is shorter (Figure 5). These results demonstrate once more that hydrophobicity is a key property for P-gp inhibitors across a wide range of chemical scaffolds. Most of the studies in the past already highlighted the importance of lipophilicity for high potency of P-gp inhibitors.<sup>10,11,16</sup> Several other studies showed that lipophilicity might influence pharmacological activity in a space-directed manner rather than as a general physicochemical determinant.<sup>29,30</sup> This space-directedness might be indicative of different orientations of molecules within the binding pocket of P-gp as reported by Pleban *et al.*<sup>31</sup> This is perfectly in-line with the findings of our model, as all important correlograms for high biological activity measure the distance of a hydrogen bond acceptor, -donor and shape probes from one particular hydrophobic probe of the molecules (DRY-N1, DRY-O and DRY-TIP). Thus, one of the hydrophobic substitutions seems to be a crucial hallmark for the mapping of the pharmacophoric pattern of P-gp inhibitors.



**Figure 5.** DRY-DRY pair of probes representing important hydrophobic regions separated by a distance of 14.40-14.80 Å present in the highly potent ( $IC_{50} < 1\mu M$ ) P-gp inhibitor GPV0647.

The DRY-TIP correlogram refers to the distance of highly hydrophobic substituents from different edges of the molecules. This pair of probes at a distance of 16.00-16.80 Å is present in all compounds of the training set having  $IC_{50} < 1\mu M$ . In propafenones this descriptor predominantly represents the distance between the central aromatic ring of the scaffold and the N-substituted hydrophobic moiety (Figure 6a). A similar distance range (17.60-18.00 Å) between two pharmacophores has been observed in chalcone derivatives as discussed previously in chapter 2. However, Cianchetta and co-workers identified a distance of 20.5 Å between the same probes.<sup>23</sup> The steric hot spots (TIP-TIP) were identified making three important boundaries of the molecules. This includes N-substituted hydrophobic groups in propafenone derivatives, which is separated by a distance of 18.40-18.80 Å from any of the two other ends of the molecules (Figure 6b).

It demonstrates the importance of hydrophobic molecular boundaries for high biological activity of P-gp inhibitors. Similar molecular shape has been identified for chalcones derivatives (Chapter 2). This is in agreement with Broccatelli and co-workers, who provided first evidence for the importance of an optimal shape for P-gp inhibitors.<sup>24</sup>

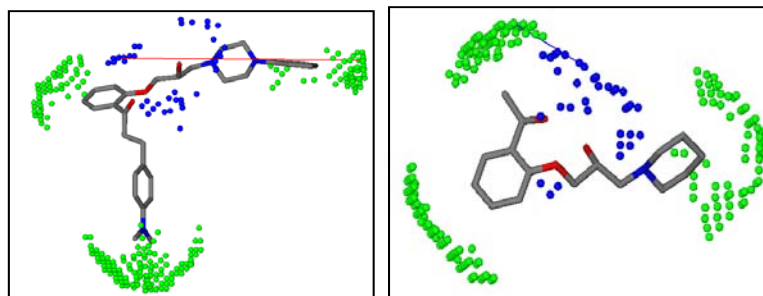


(a) GPV0576

(b) WISE\_B005

**Figure 6.** (a) Represents a distance of 16.00-16.80 Å between a hydrophobic group at the nitrogen atom (Dry: yellow region) and one of the three molecular edges (TIP; green region), (b) represents two edges of the molecule (TIP-TIP) 18.40-18.80 Å apart from each other.

The N1-TIP correlograms provide the important distances of a hydrogen bond acceptor from different edges of the molecules. These coefficients show an interesting behavior, having negative value at shorter distances (3.60-4.00 Å), but become positive for larger distances (17.20-17.60 Å) (Figure 7a,b). This indicates that potent P-gp inhibitors ( $IC_{50} < 1\mu M$ ) show elongated conformations and have a hydrogen bond acceptor far away from molecular boundaries. Weak P-gp inhibitors ( $IC_{50} > 1\mu M$ ) seem to be more compact have their hydrogen bond acceptor group close to one of its edges. This is well pronounced in the least active propafenone analogs ( $IC_{50} > 100\mu M$ ), where the phenylpropionyl-moiety of the propiophenone scaffold was replaced by a  $CH_3$  group.

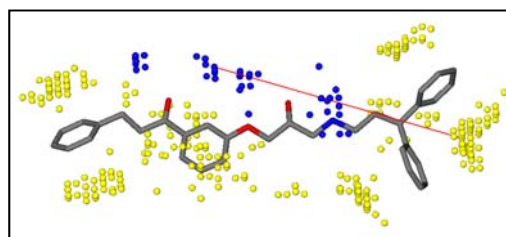


(a) GPV0610

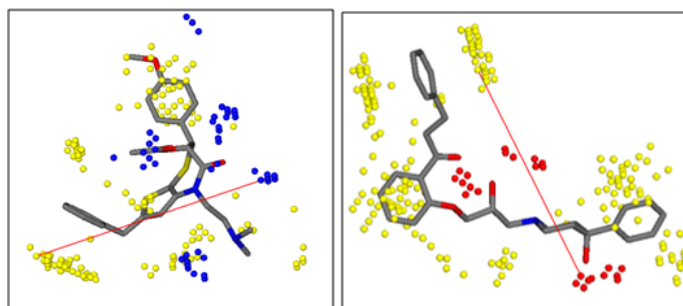
(b) GPV0017

**Figure 7.** (a) Shows a region (blue color) surrounding a hydrogen bond acceptor group (C=O) present in the middle of the molecule, at a distance of 17.20-17.60 Å away from a hydrophobic edge, identified in all potent P-gp inhibitors ( $IC_{50} < 1\mu M$ ). (b) A representative of compounds having  $IC_{50} > 100\mu M$ , where (C=O) is found close (3.60-4.00 Å) to one of the edges.

The DRY-N1 correlogram is identical with the highest positive variable of the N1-TIP correlogram and refers to the distance of a hydrogen bond acceptor from a large hydrophobic moiety. In propafenone derivatives it represents the distance (14.40-14.80 Å) between the carbonyl group and the hydrophobic substituent at the basic nitrogen atom. It is present in most of the compounds having  $IC_{50}$  values below  $1\mu M$  (Figure 8a,b), except for a few smaller compounds such as benzopyrano-[3,4-b][1-4]oxazines PCO770 and PCO726 which are discussed separately in chapter 3. The Dry-O correlogram provides the distance of a hydrogen bond donor from the same hydrophobic N-substituent (12.40-12.80 Å). However, in propafenone analogs containing a 4-hydroxypiperidine moiety, it represents the distance between the 3-phenyl substitution of the propafenone scaffold and the 4-hydroxy group at the piperidine (Figure 8c) and thus confirms our previous findings about the importance of 4-hydroxy-4-phenyl piperidines for high biological activity of propafenones.<sup>8,32</sup> This pair of probes is present in highly active ( $IC_{50} < 1\mu M$ ) compounds of different series and thus further emphasizes the importance of a hydrophobic moiety which also reflects an important shape parameter in the TIP-TIP and N1-TIP correlograms.



(a) GPV0649



(b) ERK\_PA008

(c) GPV 0062



**Figure 8.** (a) (b) Shows a hydrogen bond acceptor at a distance of 14.40-14.80 Å apart from a large hydrophobic group in P-gp inhibitors having different chemical scaffolds. (c) Represent a region (red color) around a hydrogen bond donor group and map its distance (12.40-12.80 Å) from a hydrophobic moiety.

Finally, the N1-N1 correlogram outlines the influence of the distance separating two hydrogen bond acceptors (8.80-9.20 Å). However, this correlogram is not consistent in the present model and could not separate completely the highly active ( $IC_{50} < 1\mu M$ ) compounds from low active ( $IC_{50} > 1\mu M$ ) ones. This might reflect the fact that the highly promiscuous binding site of P-gp possesses multiple spots able to participate in hydrophobic and hydrogen bond interactions and that different chemical series most probably utilize different hydrogen bond interaction patterns.

Overall, the present model, by using structurally diverse compounds, reflects a large hydrophobic moiety present at a specific distance from a hydrogen bond acceptor as global property for P-gp inhibitors. Our GRIND model further points towards the importance of the distance of a hydrophobic group from hydrogen bond acceptor/donor and from different edges of the molecules and thus elucidates the crucial attributes for variations in biological activity of P-gp inhibitors.

### Conclusions

P-gp can accommodate a wide range of structurally diverse compounds in its binding pocket. Our 3D-QSAR model containing different chemical scaffolds identified three important molecular features of P-gp inhibitors, one of which is a large hydrophobic moiety. The GRIND models indicate a favourable distance range of this hydrophobic edge from different hydrogen bond acceptor and donor groups of the molecules. The distances remain consistent for all compounds having  $IC_{50} < 1\mu M$ . This indicates that large hydrophobic groups in the molecules get an optimal fit within the binding pocket and orientate the rest of the molecule in a way that hydrogen bond acceptors, donors as well as the other edges of the molecule get most favorable positions of these groups within the binding pocket.

### References

1. Ecker, G. F.; Chiba, P. Pharmakologisch wirksame o-Acylaryloxypropanolamine mit tertiarem und quartfirem Stickstoff. (Pharmacologically active o-Acylaryloxypropanolamines with tertiary and quaternary nitrogen. 1993.

2. Chiba, P.; Ecker, G.; Tell, B.; Moser, A.; Schmid, D.; Drach, J. Modulation of PGP-mediated multidrug-resistance by propafenone analogs. *Proc. Am. Assoc. Cancer Res* **1994**, 35.
3. Chiba, P.; Burghofer, S.; Richter, E.; Tell, B.; Moser, A.; Ecker, G. Synthesis, pharmacologic activity, and structure-activity relationships of a series of propafenone-related modulators of multidrug resistance. *J Med Chem* **1995**, 38, 2789-93.
4. Chiba, P.; Annibali, D.; Hitzler, M.; Richter, E.; Ecker, G. Studies on propafenone-type modulators of multidrug resistance VI. Synthesis and pharmacological activity of compounds with varied spacer length between the central aromatic ring and the nitrogen atom. *Il Farmaco* **1998**, 53, 357-364.
5. Ecker, G.; Huber, M.; Schmid, D.; Chiba, P. The importance of a nitrogen atom in modulators of multidrug resistance. *Mol Pharmacol* **1999**, 56, 791-6.
6. Schmid, D.; Ecker, G.; Kopp, S.; Hitzler, M.; Chiba, P. Structure-activity relationship studies of propafenone analogs based on P-glycoprotein ATPase activity measurements. *Biochem Pharmacol* **1999**, 58, 1447-56.
7. Ecker, G.; Chiba, P.; Hitzler, M.; Schmid, D.; Visser, K.; Cordes, H. P.; Csollei, J.; Seydel, J. K.; Schaper, K. J. Structure-activity relationship studies on benzofuran analogs of propafenone-type modulators of tumor cell multidrug resistance. *J Med Chem* **1996**, 39, 4767-74.
8. Chiba, P.; Hitzler, M.; Richter, E.; Huber, M.; Tmej, C.; Giovagnoni, E.; Ecker, G. Studies on Propafenone-type Modulators of Multidrug Resistance III: Variations on the Nitrogen. *Quantitative Structure-Activity Relationships* **1997**, 16, 361-366.
9. Tmej, C.; Chiba, P.; Huber, M.; Richter, E.; Hitzler, M.; Schaper, K. J.; Ecker, G. A combined Hansch/Free-Wilson approach as predictive tool in QSAR studies on propafenone-type modulators of multidrug resistance. *Arch Pharm (Weinheim)* **1998**, 331, 233-40.
10. Chiba, P.; Holzer, W.; Landau, M.; Bechmann, G.; Lorenz, K.; Plagens, B.; Hitzler, M.; Richter, E.; Ecker, G. Substituted 4-acylpyrazoles and 4-acylpyrazolones: synthesis and multidrug resistance-modulating activity. *J Med Chem* **1998**, 41, 4001-11.
11. Hiessbock, R.; Wolf, C.; Richter, E.; Hitzler, M.; Chiba, P.; Kratzel, M.; Ecker, G. Synthesis and in vitro multidrug resistance modulating activity of a series of dihydrobenzopyrans and tetrahydroquinolines. *J Med Chem* **1999**, 42, 1921-6.
12. Kaiser, D.; Smiesko, M.; Kopp, S.; Chiba, P.; Ecker, G. F. Interaction field based and hologram based QSAR analysis of propafenone-type modulators of multidrug resistance. *Med Chem* **2005**, 1, 431-44.
13. Langer, T.; Eder, M.; Hoffmann, R. D.; Chiba, P.; Ecker, G. F. Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model. *Arch Pharm (Weinheim)* **2004**, 337, 317-27.
14. Kaiser, D.; Zdrzil, B.; Ecker, G. F. Similarity-based descriptors (SIBAR)--a tool for safe exchange of chemical information? *J Comput Aided Mol Des* **2005**, 19, 687-92.
15. Zdrzil, B.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity-Based Descriptors (SIBAR) as Tool for QSAR Studies on P-Glycoprotein Inhibitors: Influence of the Reference Set. *QSAR & Combinatorial Science* **2007**, 26, 669-678.
16. Wiese, M.; Pajeva, I. K. Structure-activity relationships of multidrug resistance reversers. *Curr Med Chem* **2001**, 8, 685-713.
17. Pearce, H. L.; Safa, A. R.; Bach, N. J.; Winter, M. A.; Cirtain, M. C.; Beck, W. T. Essential features of the P-glycoprotein pharmacophore as defined by a series of

reserpine analogs that modulate multidrug resistance. *Proc Natl Acad Sci U S A* **1989**, 86, 5128-32.

18. Pearce, H. L.; Winter, M. A.; Beck, W. T. Structural characteristics of compounds that modulate P-glycoprotein-associated multidrug resistance. *Adv Enzyme Regul* **1990**, 30, 357-73.

19. Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Three-dimensional quantitative structure-activity relationships of inhibitors of P-glycoprotein. *Mol Pharmacol* **2002**, 61, 964-73.

20. Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of three-dimensional quantitative structure-activity relationships of P-glycoprotein inhibitors and substrates. *Mol Pharmacol* **2002**, 61, 974-81.

21. Pajeva, I. K.; Wiese, M. Pharmacophore model of drugs involved in P-glycoprotein multidrug resistance: explanation of structural variety (hypothesis). *J Med Chem* **2002**, 45, 5671-86.

22. Chang, C.; Swaan, P. W. Computational approaches to modeling drug transporters. *Eur J Pharm Sci* **2006**, 27, 411-24.

23. Cianchetta, G.; Singleton, R. W.; Zhang, M.; Wildgoose, M.; Giesing, D.; Fravolini, A.; Cruciani, G.; Vaz, R. J. A pharmacophore hypothesis for P-glycoprotein substrate recognition using GRIND-based 3D-QSAR. *J Med Chem* **2005**, 48, 2927-35.

24. Broccatelli, F.; Carosati, E.; Neri, A.; Frosini, M.; Goracci, L.; Oprea, T. I.; Cruciani, G. A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields. *J Med Chem* **2011**.

25. Jabeen, I.; Wetwitayaklung, P.; Klepsch, F.; Parveen, Z.; Chiba, P.; Ecker, G. F. Probing the stereoselectivity of P-glycoprotein-synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-b][1,4]oxazines. *Chem Commun (Camb)* **2011**, 47, 2586-8.

26. Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Computer Methodology* **1990**, 3, 537-547.

27. Durán, Á.; Pastor, M. An advanced tool for computing and handling GRIND-Independent Descriptors. User Manual Version 1.06. **2011**.

28. Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating Optimal Linear PLS Estimations (GOLPE): An Advanced Chemometric Tool for Handling 3D-QSAR Problems. *Quantitative Structure-Activity Relationships* **1993**, 12, 9-20.

29. Pajeva, I.; Wiese, M. Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: a comparative molecular field analysis study. *J Med Chem* **1998**, 41, 1815-26.

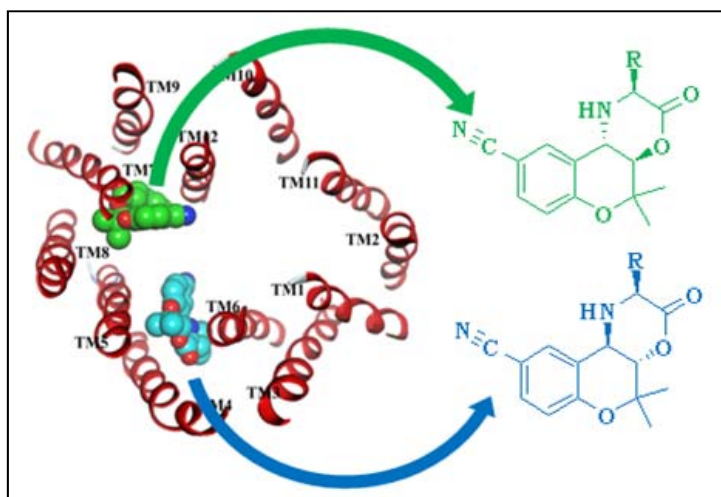
30. Pajeva, I. K.; Wiese, M. A Comparative Molecular Field Analysis of Propafenone-type Modulators of Cancer Multidrug Resistance. *Quantitative Structure-Activity Relationships* **1998**, 17, 301-312.

31. Pleban, K.; Hoffer, C.; Kopp, S.; Peer, M.; Chiba, P.; Ecker, G. F. Intramolecular distribution of hydrophobicity influences pharmacological activity of propafenone-type MDR modulators. *Arch Pharm (Weinheim)* **2004**, 337, 328-34.

32. Klepsch, F.; Chiba, P.; Ecker, G. F. Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Comput Biol* **2011**, *7*, e1002036.

**Probing the Stereoselectivity of P-glycoprotein – Synthesis, Biological Activity and Ligand Docking Studies of a Set of Enantiopure Benzopyrano[3,4b][1,4]oxazines**

Page 95-99



In this chapter a data set of diastereoisomers of benzopyrano[3,4-b][1,4]oxazines were docked into a homology model of P-glycoprotein to probe stereoselective interaction of diastereoisomeric pairs.

**Contents**

1. Introduction
2. Chemistry
3. Structure Activity Relationships
4. Docking

☞ Appendix available: Page 173-193

1. Biological Essay
2. Chemistry

① **Information:** This chapter was published in *Chem Comm*, 2011, Volume 47, 2586-2588 by **Jabeen Ishrat**, Wetwitayaklung Penpun, Klepsch Freya, Parveen Zahida, Chiba Peter, Ecker Gerhard F.

# ChemComm

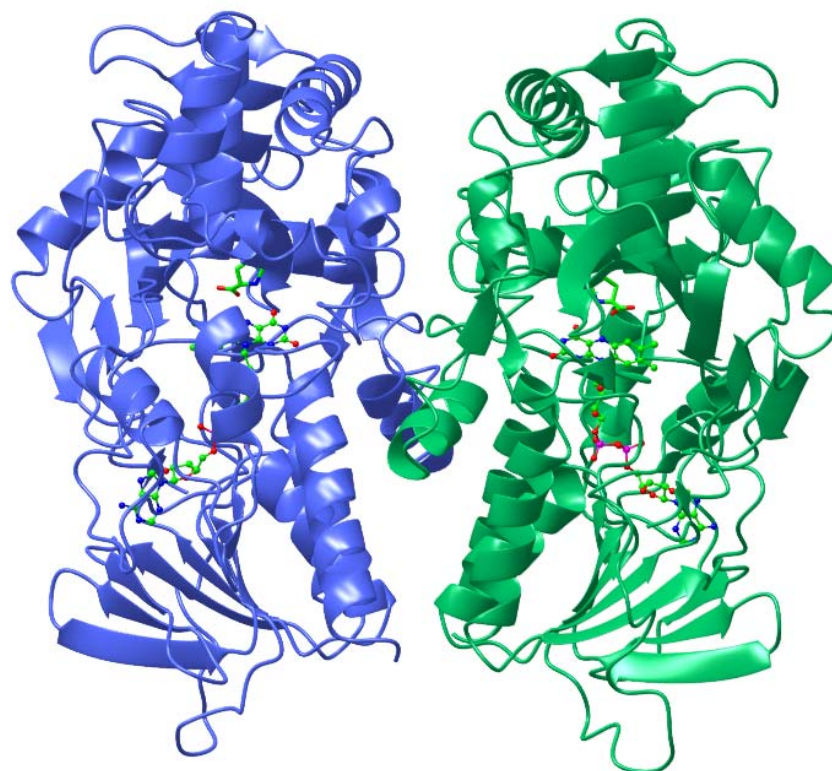
This article is part of the

## Enzymes & Proteins web themed issue

This issue showcases high quality research in the field of enzymes and proteins.

Please visit the website to access the other articles in this issue:-

<http://www.rsc.org/chemcomm/enzymesandproteins>



Cite this: *Chem. Commun.*, 2011, **47**, 2586–2588

www.rsc.org/chemcomm

## Probing the stereoselectivity of P-glycoprotein—synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-*b*][1,4]oxazines†‡

Ishrat Jabeen,<sup>a</sup> Penpun Wetwitayaklung,<sup>ab</sup> Freya Klepsch,<sup>a</sup> Zahida Parveen,<sup>c</sup> Peter Chiba<sup>c</sup> and Gerhard F. Ecker<sup>\*a</sup>

Received 5th August 2010, Accepted 24th November 2010

DOI: 10.1039/c0cc03075a

**A series of enantiomerically pure benzopyrano[3,4-*b*][1,4]-oxazines have been synthesised and tested for their ability to inhibit P-glycoprotein. Reducing the conformational flexibility of the molecules leads to remarkable differences in the activity of diastereoisomers. Docking studies into a homology model of human P-gp provide first insights into potential binding areas for these compounds.**

P-glycoprotein (P-gp) is a transmembrane, ATP-dependent drug efflux pump which transports a wide variety of structurally and functionally diverse compounds out of cells.<sup>1</sup> P-gp is expressed in epithelial cells of the kidney, liver, pancreas, colon, as well as at the blood–brain barrier,<sup>2</sup> underscoring its role in maintaining concentration gradients of (toxic) compounds at physiological barriers.<sup>3</sup> In addition, it is very often overexpressed in tumor cells and thus is one of the major factors responsible for multiple drug resistance in anticancer therapy. Inhibition of P-gp has therefore been advocated as promising concept to overcome the MDR phenotype. However, although several inhibitors of P-gp have been evaluated in clinical studies, none of them has reached the market so far, which questions the druggability of P-gp.<sup>4</sup>

One of the initial candidates for use as P-gp inhibitor was the calcium channel blocker verapamil. However, clinical studies indicated that the serum concentrations required to reverse MDR lead to severe cardiovascular side effects due to the original biological profile of verapamil. As the cardiovascular activity is concentrated in the *S*-enantiomer and both

enantiomers are equipotent at P-gp, *R*-verapamil was used for further clinical studies. Unfortunately, also this compound failed in clinical phase 3 studies.

Lack of significant stereoselectivity in drug/P-gp interaction was also observed for other compounds, such as niguldipine, nitrendipine, felodipine, carvedilol, propranolol, zosuquidar<sup>5</sup> and propafenone.<sup>5</sup>

However, there are also a few reports of remarkable stereospecificity.<sup>6</sup> Furthermore, the recently published crystal structure of mouse P-gp co-crystallised with the two enantiomeric cyclopeptides QZ59-*RRR* and QZ59-*SSS* revealed distinct binding sites for the two enantiomers. QZ59-*RRR* binds in the center of the P-gp binding pocket, whereas QZ59-*SSS* binds at two positions: in one position it interacts with hydrophobic residues between TMs 6 and 12, while in the other position it interacts with TMs 8 and 9 and is surrounded by three polar residues. Amino acid residue Val982 plays an important role having close proximity to all three QZ59 sites.<sup>7</sup> Analogous positions of the QZ-isomers were found in docking experiments of the two isomers into a homology model of human P-gp based on the mouse P-gp structure.<sup>8</sup>

In light of our intense structure–activity relationship studies of inhibitors of P-gp, we also synthesized and tested a series of 3-hydroxy-4-amino-dihydrobenzopyranes.<sup>9</sup> These compounds showed biological activities in the low micromolar range, which is comparable to propafenone and verapamil.

In contrast to our main lead compound propafenone, the dihydrobenzopyranes offer the advantage of remarkably reduced conformational flexibility and thus might be versatile molecular tools for probing stereoselective differences of drug/P-gp interaction. Especially annelation of a third ring leading to benzopyrano[3,4-*b*][1,4]oxazines and introduction of large substituents at position 2 of the tricyclic system should lead to compounds with pronounced configurational differences. The compound design is thus based on synthesis of both enantiomers of epoxide **4**, nucleophilic ring opening with *tert*-butyl esters of selected amino acids followed by ester hydrolysis and cyclisation to yield enantiopure target compounds **11–13** (Scheme 1).

Synthesis of the benzopyrane ring system was achieved according to Godfrey *et al.*<sup>10</sup> *O*-alkylation of 4-hydroxybenzonitril (**1**) with 3-trifluoroacetyl-3-methyl-but-1-yne

<sup>a</sup> Department of Medicinal Chemistry, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.

E-mail: gerhard.f.ecker@univie.ac.at; Fax: +43 1-4277-9551; Tel: +43 1-4277-55110

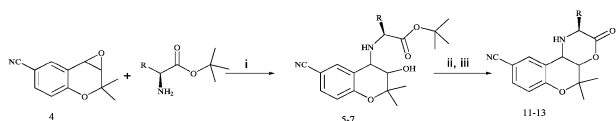
<sup>b</sup> Department of Pharmacognosy, Faculty of Pharmacy, Silpakorn University, Nakhon Pathom, Thailand

<sup>c</sup> Medical University of Vienna, Institute of Medical Chemistry, Weahringerstrasse 10, 1090, Vienna, Austria.

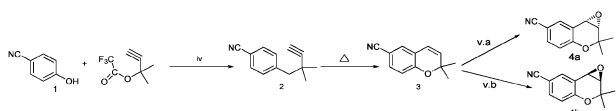
E-mail: peter.chiba@meduniwien.at; Fax: +43 1-4277-60889; Tel: +43 1-4277-60806

† This article is part of the 'Enzymes and Proteins' web-theme issue for ChemComm.

‡ Electronic supplementary information (ESI) available: Experimental procedures and spectroscopic data of all compounds biologically tested as well as NMR-spectra of tricycles **11a–13b**. See DOI: 10.1039/c0cc03075a



**Scheme 1** Synthesis of target compounds **11–13**; (i) 96% ethanol, reflux; (ii) 70% HClO<sub>4</sub>; (iii) 4-dimethylaminopyridine, bis-(2-oxo-3-oxazolidinyl)-phosphinic chloride, triethylamine.



**Scheme 2** Synthesis of the benzopyrane ring system and enantiomeric pure (*S,S*)- and (*R,R*)-epoxide **4a** and **4b**; (iv) CH<sub>3</sub>CN, DBU, CuCl<sub>2</sub>, –4 °C, Ar atmosphere; (v.a) (*S,S*)-Mn(III) Salen NaOCl solution, buffer to pH 11.3, 0 °C; (v.b) (*R,R*)-Mn(III) Salen NaOCl solution, buffer to pH 11.3, 0 °C.

followed by thermal cyclization gave 6-cyano-2,2-dimethyl-2*H*-1-benzopyran **3**. Enantioselective epoxidation using a Jacobsons Mn(III) Salen epoxidation catalyst and commercial household bleach (sodium hypochlorite)<sup>11</sup> as the oxygen source gave (*S,S*)- and (*R,R*)-epoxide **4a** and **4b** (Scheme 2). Enantiomeric purity of both epoxides was confirmed by HPLC analysis, using a LiChroCART (*R,R*)-Whelk-01 column (25 × 0.4 cm) and *n*-hexane/isopropanol (95 : 5) as eluent.

Nucleophilic ring opening of epoxides with *L*-alanine-, *L*-valine- and *L*-phenylalanine-*tert*-butyl-esters yields optically pure *trans*-3,4-disubstituted diastereomeric esters **5a,b–7a,b**. These were hydrolysed with 70% HClO<sub>4</sub><sup>12</sup> to yield the corresponding acids **8a,b–10a,b**, which were subsequently cyclised without further purification using bis-(2-oxo-3-oxazolidinyl)-phosphinic chloride, 4-dimethylaminopyridine and triethylamine to yield the target compounds **11a,b–13a,b** (Table 1).

Biological activity of target compounds **11–13** as well as of intermediates **5–7** was assessed using the daunorubicin efflux protocol as described previously (see ESI†). As negative charges are known to be detrimental for P-gp inhibitory activity the carboxylic acids **8–10** were not measured. In the daunorubicin efflux assay the effect of different modulators on

**Table 1** Chemical structure and biological activity of enantiomerically pure benzopyrano[3,4-*b*][1,4]oxazines

| #          | R  | log <i>P</i> | EC <sub>50</sub> /μM |
|------------|--|--------------|----------------------|
| <b>5a</b>  | CH <sub>3</sub>                                  | 2.84         | 29.85                |
| <b>5b</b>  | CH <sub>3</sub>                                  | 2.84         | 14.55                |
| <b>6a</b>  | CH(CH <sub>3</sub> ) <sub>2</sub>                | 3.82         | 2.40                 |
| <b>6b</b>  | CH(CH <sub>3</sub> ) <sub>2</sub>                | 3.82         | 2.70                 |
| <b>7a</b>  | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 4.38         | 0.55                 |
| <b>7b</b>  | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 4.38         | 0.77                 |
| <b>11a</b> | CH <sub>3</sub>                                  | 1.98         | 1241.65              |
| <b>11b</b> | CH <sub>3</sub>                                  | 1.98         | 76.89                |
| <b>12a</b> | CH(CH <sub>3</sub> ) <sub>2</sub>                | 2.95         | 15.32                |
| <b>12b</b> | CH(CH <sub>3</sub> ) <sub>2</sub>                | 2.95         | 59.33                |
| <b>13a</b> | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 3.51         | 2.69                 |
| <b>13b</b> | CH <sub>2</sub> (C <sub>6</sub> H <sub>5</sub> ) | 3.51         | 259.78               |

the transport rate is measured in a direct functional assay. Furthermore, EC<sub>50</sub> values obtained correlate well with those from cytotoxicity assays and rhodamine 123 efflux studies.<sup>13,14</sup> Values are given in Table 1 and are the mean of at least three independently performed experiments. Generally, interexperimental variation was below 20%.

EC<sub>50</sub> values cover a range of more than three orders of magnitude with the two phenylalanine esters **7a** and **7b** being the most active compounds (**7a**: 0.55 μM; **7b**: 0.77 μM), followed by the valine analogues **6a** (2.40 μM) and **6b** (2.70 μM). Least active compounds in the series of esters were the alanine derivatives with 29.85 μM (**5a**) and 14.55 μM (**5b**), respectively. It has to be noted that for all three diastereoisomeric pairs almost no differences in biological activity was observed. This pattern changes remarkably upon ring closure to the benzopyrano[3,4-*b*][1,4]oxazines. Whereas the valine analogues **12a,b** are still within one order of magnitude, both the alanine and phenylalanine derivatives exhibit remarkable differences in their potential to inhibit P-gp. Most strikingly, in the case of alanine the **4a,S,10b,R**-isomer **11a** is by a factor of 15 less active than the diastereomeric **4a,R,10b,S** analogue **11b**, whereas in the case of the phenylalanine derivatives this behaviour reverses with the **4a,S,10b,R**-isomer **13a** being by two orders of magnitude more active than **13b**.

It is widely accepted that access of substrates/inhibitors to the binding cavity of P-gp occurs directly from the membrane bilayer rather than from the aqueous intracellular medium. Thus, QSAR studies very often show a correlation between lipophilicity of the compounds under investigation and their P-gp inhibitory activity.

In this case, differences in activity more likely reflect the ability of the compounds to enter the membrane bilayer rather than differences in their interaction pattern with P-gp. Calculating the log *P* values of all target compounds with the software package MOE and correlating the values with the log(1/EC<sub>50</sub>) values exhibit remarkable differences between the (a) and (b) series of compounds. Compounds derived from the (*S,S*)-epoxide (a series, showing a (3*S*,4*R*)-configuration at the benzopyrane ring after aminolysis of the epoxide) showed an excellent correlation (*r*<sup>2</sup> = 0.96, *n* = 6). This indicates that within this series of compounds differences in their P-gp inhibitory potency are mainly due to their capability to permeate into the membrane bilayer rather than to protein–ligand interactions at P-gp.

Compounds from the (b) series yielded a significantly lower *r*<sup>2</sup> value (0.42). Considering the remarkable drop of activity for the benzyl-derivative **13b** strongly indicates steric constraints for this series of diastereoisomers and thus maybe leading to different binding modes at P-gp. This is further supported by results of docking studies performed on a homology model of human P-gp based on the X-ray structure of mouse P-gp co-crystallised with the cyclic peptide QZ59 (PDB ID: 3G5U; for details see ESI†).

Both esters **5a,b–7a,b** as well as lactones **11a,b–13a,b** were docked into the homology model of human P-gp. Agglomerative Hierarchical Cluster analysis of the consensus RMSD matrix of esters **5a,b–7a,b** based on the common scaffold identifies 10 clusters. Five clusters contain all compounds of configuration (3*S*,4*R*), while four clusters contain all compounds having configuration (3*R*,4*S*). However, we



also identified one cluster which contains five (**5a,b**; **6a,b** and **7b**) out of six docked esters. Analysis of the ligand protein interaction profile of eight of these clusters showed mainly interactions with amino acid residues of TM5 and TM6. The dominant interacting amino acids for **5a,b–7a,b** include Tyr307, Tyr310, Phe343, Phe336 and Gln347 (Fig. S1, ESI†). However, two clusters (one of series (a), one of series (b)) showed different interaction patterns. Compounds of (3*S*,4*R*)-configuration additionally interact with amino acid residues of TM11, including Phe951, Ser952, Cys956 and Met 69, while compounds of (3*R*,4*S*)-configuration showed interaction with TM1, TM2 and TM11, including Tyr117, Ser952, Phe72 and Met69. Using an identical clustering approach for the tricycles **11a,b–13a,b** identified 15 different clusters. Seven of them contain only compounds with (4*aS*,10*bR*)-configuration (**11a–13a**), and are located close to the potential entry pathway (Fig. 1A). Analysis of the protein–ligand interaction pattern showed mainly interactions with TM 4, 5, and 6, in particular with amino acid residues Tyr307, Phe343, Ala342, and Phe303. Eight clusters contain all compounds with (4*aR*,10*bS*)-configuration (**11b–13b**). These clusters are located in two different positions. One position is identical with those of **11a–13a**, the second position is located close to TM 7, 8, 9 and 12, surrounded by amino acid residues Ala985, Ile765 and Leu724 (Fig. 1B). Similar results are obtained when performing the agglomerative hierarchical clustering on the whole set of poses obtained (**5a–13b**; Table S1, ESI†).

Comparing the main positioning of the benzopyrano-[3,4-*b*][1,4]oxazines with those of QZ59 some overlap could be observed. Especially interaction with Tyr307, Phe343, Phe336, Ala985, Ala342, Met69 and Phe728 was observed for all ligands (Fig. S2, ESI†). Interestingly, almost all clusters observed are located near TM 4, 5, and 6, which are forming one of the two rings. This is consistent with our recent observation of two pseudosymmetric drug translocation pathways.<sup>15</sup>

Furthermore, it is also interesting to note that compounds of series (a), which show excellent correlation between log *P* values and P-gp inhibitory activity, are predominantly positioned at the potential entry gate, whereas compounds

of the series (b), which show a structure–activity pattern independent of log *P* values, are populating both the entry gate and positions deeper inside the protein. This might provide first insights into the entry path for the ligands.

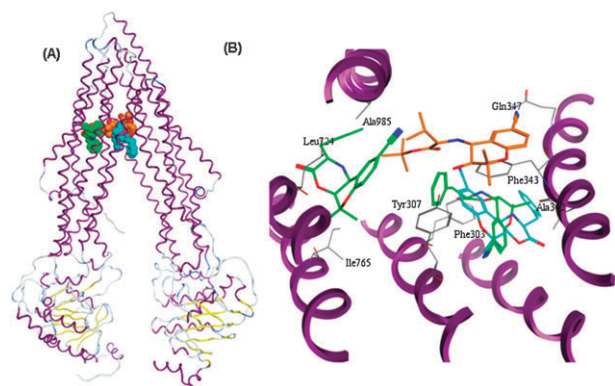
A closer look of ligand–protein interaction profiles of compounds **13a,b** and **7a,b** identified 4 poses of **13b** showing a steric constraint of the benzyl moiety of **13b**, which is about 2 Å apart from Tyr307 and about 2.5 Å apart from Phe343. All these poses are located at the entry gate. No such steric constraint has been observed for **13a** or for **7a,b**. In the case of **7b** this is most probably due to its conformational flexibility, which allows adopting a conformation to minimize the steric interactions. This indicates that the differences observed for the biological activities of phenylalanine derivatives **13a** and **13b** might be due to steric constraints at the entry path rather than differences in drug/transporter binding. Of course, at the current stage this has to be taken very cautiously, as P-gp undergoes major conformational changes during the transport cycle and docking experiments represent only a single snapshot of this complex movement.

Within this manuscript we present a series of stereoisomers, which, upon rigidisation, show significant differences in their inhibitory potency of the drug efflux pump P-glycoprotein. Ligand docking studies into a homology model of P-gp could provide first evidence for different binding areas of the two diastereomeric compound series. Thus, benzopyrano-[3,4-*b*][1,4]oxazines are versatile tools for exploring the stereoselectivity of drug/P-glycoprotein interaction.

We are grateful to the Austrian Science Fund for financial support (grant SFB F35). Ishrat Jabeen and Zahida Parveen thank the Higher Education Commission of Pakistan for financial support.

## Notes and references

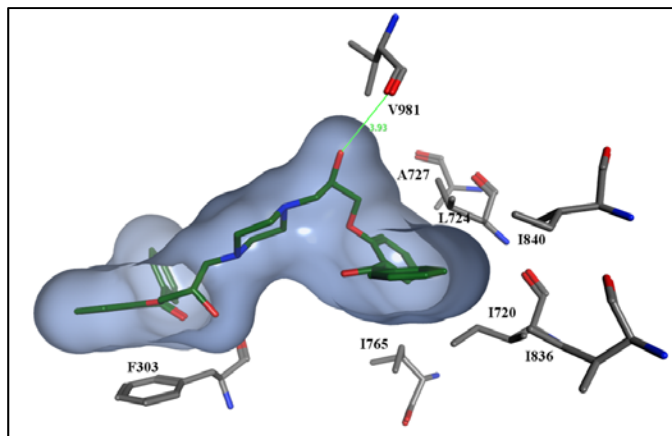
- 1 M. M. Gottesman, T. Fojo and S. E. Bates, *Nat. Rev. Cancer*, 2002, **2**, 48–58.
- 2 M. Kuhnle, M. Egger, C. Muller, A. Mahringer, G. Bernhardt, G. Fricker, B. Konig and A. Buschauer, *J. Med. Chem.*, 2009, **52**, 1190–1197.
- 3 M. F. Fromm, *Trends Pharmacol. Sci.*, 2004, **25**, 423–429.
- 4 *Transporters as Drug Carriers*, ed. G. F. Ecker and P. Chiba, Wiley-VCH, Weinheim, 2009.
- 5 P. Bhatia, M. Kolinski, R. Moaddel, K. Jozwiak and I. W. Wainer, *Xenobiotica*, 2008, **38**, 656–675.
- 6 S. S. Carey, M. Gleason-Guzman, V. Gokhale and L. H. Hurley, *Mol. Cancer Ther.*, 2008, **7**, 3617–3623.
- 7 S. G. Aller, J. Yu, A. Ward, Y. Weng, S. Chittaboina, R. Zhuo, P. M. Harrell, Y. T. Trinh, Q. Zhang, I. L. Urbatsch and G. Chang, *Science*, 2009, **323**, 1718–1722.
- 8 I. K. Pajeva, C. Globisch and M. Wiese, *FEBS J.*, 2009, **276**, 7016–7026.
- 9 R. Hiessbock, C. Wolf, E. Richter, M. Hitzler, P. Chiba, M. Kratzel and G. Ecker, *J. Med. Chem.*, 1999, **42**, 1921–1926.
- 10 J. D. Godfrey, R. H. Mueller, T. C. Sedergran, N. Soundararajan and V. J. Colandrea, *Tetrahedron Lett.*, 1994, **35**, 6405–6408.
- 11 N. H. Lee, A. R. Muci and E. N. Jacobsen, *Tetrahedron Lett.*, 1991, **32**, 5055–5058.
- 12 F. M. Callahan, G. W. Anderson, R. Paul and J. Zimmerman, *J. Am. Chem. Soc.*, 1963, **85**, 201–207.
- 13 G. Ecker, M. Huber, D. Schmid and P. Chiba, *Mol. Pharmacol.*, 1999, **56**, 791–796.
- 14 G. Ecker, P. Chiba and K. J. Schaper, *J. Pharm. Pharmacol.*, 1997, **49**, 305–309.
- 15 Z. Parveen, C. Bentele, T. Stockner, S. Pferschy, M. Kraupp, M. Freissmuth, G. Ecker and P. Chiba, *Mol. Pharmacol.*, 2011, under revision.



**Fig. 1** (A) Shows the three main clusters obtained on the basis of a common scaffold clustering; blue: (4*aS*,10*bR*)-isomers **11a–13a**; green: (4*aR*,10*bS*)-isomers **11b–13b**; brown: **5a–7b**, (B) a docking pose of **13a** (blue) and **13b** (green) near the entry gate showing steric constraints for **13b**, as well as a pose of **13b** deeper inside the membrane (green); brown: a docking pose of the ester **6b** viewed from outside into the TM region.

Structure-Activity Relationships, Ligand Efficiency and Lipophilic Efficiency Profiles of Benzophenone-Type Inhibitors of the Multidrug Transporter P-glycoprotein

Page 122-156



In this chapter a data set of benzophenone analogs along with some compounds in clinical investigations were used for ligand efficiency and lipophilic efficiency profiling, in order to get insights about the importance of these parameters for the design of P-gp inhibitors.

## Contents

### Introduction

### Results and Discussion

#### Chemistry

#### Biological Activity

#### Structure Activity Relationships

#### Ligand Efficiency

#### Lipophilic Efficiency

### Conclusions

### Experimental Section

### Computational Studies

#### Biology

☞ Appendix available: Page 194-226

### Supporting Information

① **Information:** This chapter was submitted in *Journal of Medicinal Chemistry*, 2011 by **Ishrat Jabeen**, Karin Pleban, Peter Chiba and Gerhard F. Ecker.

Structure-activity relationships, ligand  
efficiency and lipophilic efficiency profiles of  
benzophenone-type inhibitors of the multidrug  
transporter P-glycoprotein

*Ishrat Jabeen, <sup>†</sup> Karin Pleban, <sup>†</sup> Peter Chiba, <sup>‡</sup> Gerhard F. Ecker<sup>†\*</sup>*

*<sup>†</sup> University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, 1090,  
Vienna, Austria*

*<sup>‡</sup> Medical University of Vienna, Institute of Medical Chemistry, Waehringerstraße 10,  
1090, Vienna, Austria*

*Abbreviations List<sup>a</sup>*

*PDB ID Codes<sup>b</sup>*

---

\* To whom correspondence should be addressed. University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, 1090 Vienna, Austria. Telephone: +43-1-4277-55110. Fax: +43-1-4277-9551. E-mail Hgerhard.f.ecker@univie.ac.atH.

<sup>a</sup> P-gp, Pglycoprotein; LipE, lipophilic Efficiency; LE, Ligand efficiency; MDR, multidrug resistance; ABC, ATP binding cassette; QSAR, Quantitative structure Activity Relationship.

<sup>b</sup> 3G5U

**Abstract**

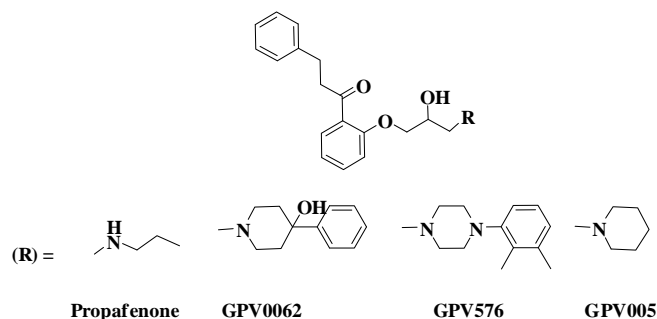
The drug efflux pump P-glycoprotein (P-gp) has been shown to reverse multidrug resistance (MDR) in tumors as well as to influence ADME properties of drug candidates. Here we synthesized and tested a series of benzophenone derivatives structurally analogous to propafenone-type inhibitors of P-gp. Some of the compounds showed ligand efficiency and lipophilic efficiency (LipE) values in the range of compounds which entered clinical trials as MDR modulators. Interestingly, although lipophilicity plays a dominant role for P-gp inhibitors, all compounds investigated showed LipE values below the threshold for promising drug candidates. Docking studies of selected analogs into a homology model of P-glycoprotein suggest that benzophenones show an interaction pattern similar to that previously identified for propafenone-type inhibitors.

**Introduction**

Membrane transporters are increasingly recognised for playing a key role in safety profiles of drug candidates, predominantly by their involvement in drug-drug interactions.<sup>1,2</sup> One of the most intensively studied families in this context is the ATP-binding cassette (ABC) transporter superfamily.<sup>3-5</sup> Several members of these ATP-driven transporters are expressed at tissue barriers and thus influence uptake and elimination of drugs and drug candidates.<sup>6</sup> Originally they have been linked to development of multidrug resistance (MDR) in tumour therapy, as they transport a wide variety of natural product toxins such as anthracyclines, vincristine and taxanes out of tumour cells.<sup>7,8</sup> Thus, P-glycoprotein (P-gp/ABCB1), discovered in 1976 and considered the paradigm ABC transporter,<sup>9,10</sup> shows a remarkably broad substrate pattern, transporting numerous structurally and functionally diverse compounds across cell membranes.<sup>3</sup> P-gp is expressed at the blood brain barrier (BBB), the blood cerebrospinal fluid (B-CSF) barrier and the intestinal barrier, thus modulating the absorption and excretion of xenobiotics across these barriers.<sup>6</sup> P-gp and its ligands (substrates and inhibitors) are therefore extensively studied both with respect to reversing multidrug resistance in tumors and for modifying ADME-Tox properties of drug candidates,<sup>11</sup> such as blood-brain barrier permeation.<sup>12,13</sup> Within the past two decades numerous

modulators of P-gp mediated drug efflux have been identified<sup>14,15</sup> and several entered clinical studies up to phase III. However, up to now no compound achieved approval, which is mainly due to severe side effects and lack of efficacy. This further emphasizes the physiological role of efflux transporters in general and P-gp in particular<sup>16</sup> and stresses the need for a more detailed knowledge on the structure and function of these proteins and the molecular basis of their interaction with small molecules.<sup>17</sup> The latter has been approached by numerous SAR- and QSAR studies, which revealed that high lipophilicity seems to be a general prerequisite for high P-gp inhibitory activity, valid across different chemical scaffolds. This is also in line with recent structure-based studies, which indicate an entry pathway via the membrane bilayer.<sup>18</sup>

In recent years the concepts of “*Binding energy of the ligand per atom*” or ligand efficiency (LE)<sup>19-21</sup> and lipophilic efficiency (LipE),<sup>22,23</sup> which combines both “*potency and lipophilicity*”, have been shown to be useful tools in the lead optimization process.<sup>24,25</sup> In the light of our extensive SAR and QSAR studies on propafenone analogues<sup>26,27</sup> (Figure 1) and related compounds, we also utilized benzophenone based probes, which contain a photoactive arylcarbonyl group as part of the pharmacophore. This led to the identification of key amino acid residues interacting with these ligands.<sup>28,29</sup> Within this study we extended the set of benzophenones in order to identify compounds with higher potency, utilizing also the concepts of LE and LipE. In addition, docking studies of selected compounds into a homology model of P-gp were performed to shed light on the potential binding mode of these compounds and to compare it with the binding hypothesis derived for analogous propafenones.<sup>17</sup>

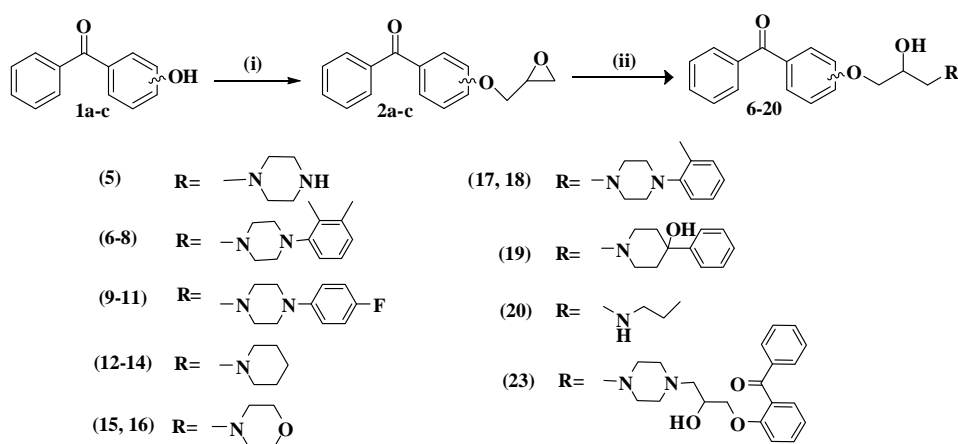


**Figure 1.** Selected propafenone analogs used in this study.

## Results and discussion

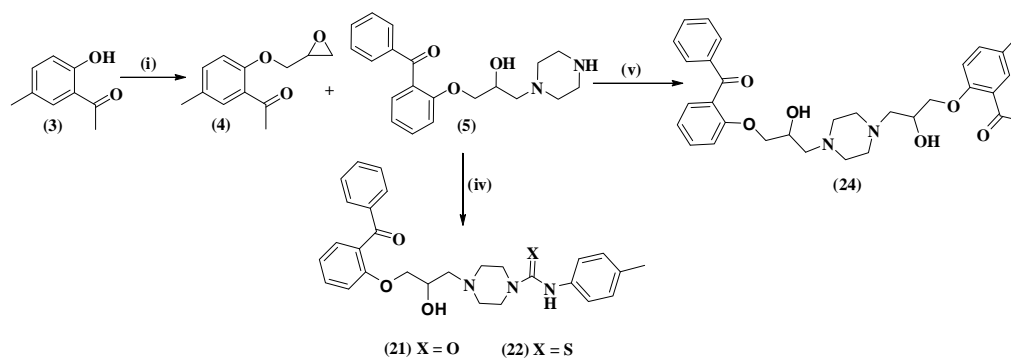
## Chemistry

Synthesis of benzophenone analogs **6-24** was carried out in analogy to the synthesis of propafenone derivatives.<sup>26</sup> Briefly, the respective ortho-, meta- or para-hydroxybenzophenone (**1a-c**) was alkylated with epichlorohydrine yielding ortho-, meta- and para-oxiranes **2a-c**. Subsequent nucleophilic oxirane ring opening with primary or secondary amines (R) gave target compounds **6-20**. Excessive amount of oxirane **2a** upon nucleophilic ring opening with piperazine yielded the homodimer **23**, whereas equimolar amounts of both partners predominantly gave the piperazine analog **5** (Scheme 1).

Scheme 1<sup>a</sup>

<sup>a</sup> Reagents and conditions: (i) NaOH, epichlorohydrine, reflux for 24 hours; (ii) methanol, respective amine (R), reflux for 24 hours (**5-20, 23**)

Further treatment of piperazine analog **5** with phenylisocyanate and its thio-analog as described by Pitha *et al.*<sup>30</sup> yielded **21** and **22**, respectively. O-alkylation of 2-hydroxy 5-methyl acetophenone (**3**) with epichlorohydrine yielded **4**, which, upon subsequent nucleophilic ring opening by piperazine **5** yielded the heterodimer **24** (Scheme 2).

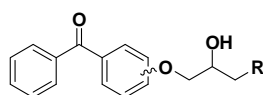
Scheme 2<sup>a</sup>

<sup>a</sup> Reagents and conditions: (i) NaOH, epichlorohydrine, reflux for 24 hours; (iv) p-tolyl isocyanate, CH<sub>2</sub>Cl<sub>2</sub>, stirring 2 hours (X = O); p-tolylisothiocyanate, CH<sub>2</sub>Cl<sub>2</sub>, stirring 2 hours (X = S); (v) methanol, reflux 5 hours

### Biological Activity

Biological activity of target compounds **6-24** was assessed using the daunorubicin efflux protocol as described previously.<sup>31</sup> Briefly, multidrug resistant CCRF-CEM vcr 1000 cells were preloaded with daunorubicin and efflux was monitored by time-dependent decrease in mean cellular fluorescence in the absence and presence of various concentrations of compounds. IC<sub>50</sub> values were calculated from concentration-response curves of efflux  $V_{max}/K_m$  as a function of compound concentration. Thus, the effect of different modulators on the transport rate is measured in a direct functional assay. Values are given in Table 1 and are the mean of at least three independently performed experiments. Generally, interexperimental variation was below 20%.

**Table 1.** Chemical structure, ligand efficiency (LE), lipophilic efficiency (LipE) and pharmacological activity of compounds **6-24**



| Comp | Position <sup>c</sup> | R | IC <sub>50</sub> (μM) | LE   | LE_Scale | clogP | LipE |
|------|-----------------------|---|-----------------------|------|----------|-------|------|
| 6    | Ortho                 |   | 0.08                  | 0.30 | 0.30     | 5.52  | 1.58 |
| 7    | Meta                  |   | 0.17                  | 0.29 | 0.30     | 5.52  | 1.24 |
| 8    | Para                  |   | 0.65                  | 0.26 | 0.30     | 5.52  | 0.66 |
| 9    | Ortho                 |   | 0.15                  | 0.30 | 0.30     | 4.96  | 1.86 |
| 10   | Meta                  |   | 0.58                  | 0.28 | 0.30     | 4.96  | 1.27 |
| 11   | Para                  |   | 0.97                  | 0.27 | 0.30     | 4.96  | 1.04 |
| 12   | Ortho                 |   | 1.20                  | 0.33 | 0.36     | 3.88  | 2.04 |
| 13   | Meta                  |   | 3.55                  | 0.31 | 0.36     | 3.88  | 1.57 |
| 14   | Para                  |   | 2.18                  | 0.32 | 0.36     | 3.88  | 1.78 |
| 15   | Ortho                 |   | 13.37                 | 0.28 | 0.36     | 2.66  | 2.21 |
| 16   | Para                  |   | 5.32                  | 0.30 | 0.36     | 2.66  | 2.61 |
| 17   | Meta                  |   | 0.20                  | 0.30 | 0.30     | 5.07  | 1.62 |
| 18   | Para                  |   | 0.50                  | 0.28 | 0.30     | 5.07  | 1.23 |
| 19   | Ortho                 |   | 0.31                  | 0.29 | 0.30     | 3.65  | 2.86 |
| 20   | Ortho                 |   | 1.21                  | 0.35 | 0.37     | 3.64  | 2.28 |
| 21   | Ortho                 |   | 0.48                  | 0.26 | 0.28     | 4.28  | 2.04 |
| 22   | Ortho                 |   | 0.38                  | 0.26 | 0.28     | 5.07  | 1.34 |
| 23   | Ortho                 |   | 0.05                  | 0.23 | 0.23     | 4.27  | 3.01 |
| 24   | Ortho                 |   | 9.48                  | 0.18 | 0.25     | 3.17  | 1.85 |

<sup>c</sup> Position of the side chain at central aromatic ring.



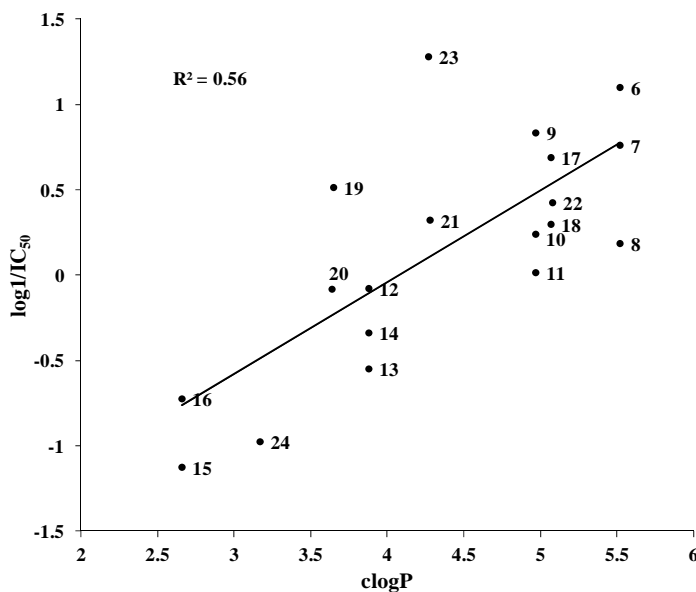
**Structure Activity Relationships**

Table 1 shows the P-gp inhibitory activity of compounds **6-24**. The IC<sub>50</sub> values cover a broad range, spanning from 0.05 μM for the dimer **23** up to 13.37 μM for the morpholine analog **15**. Besides the ortho-benzophenone dimer **23**, also the ortho analogs showing an arylpiperazine moiety (**6**, **9**) are highly active. Interestingly, the heterodimer **24** is one of the least active compounds in the data set, together with the morpholine derivatives **15** and **16**. With respect to substitution pattern at the central aromatic benzene moiety, the rank order for arylpiperazine substituted compounds generally is ortho > meta > para. An analogous trend has also been observed for propafenone analogs.<sup>32</sup> However, for compounds bearing piperidine or morpholine moieties, this trend is partly reversed. In case of piperidine derivatives, the para-derivative is slightly more active than the meta analog (1.20 vs 3.55 vs 2.18). Interestingly, also for the morpholine analogs, the para is by a factor of 2 more active than ortho-derivative (P = 0.01). Thus, the influence of the substitution pattern at the central aromatic ring seems to be more pronounced if the vicinity of the nitrogen comprises large, lipophilic moieties. This is in line with our previous findings using hydrophobic moments as descriptors in QSAR studies.<sup>33</sup>

Our intensive studies on propafenone-type inhibitors of P-gp also revealed the importance of H-bond acceptor and –donor groups in the vicinity of the basic nitrogen atom.<sup>34,35</sup> To further explore this, we synthesized the urea and thiourea analogs **21** and **22**. The two compounds showed activities in the sub-micromolar range. Notably, the urea and thiourea derivatives exhibit almost identical IC<sub>50</sub> values, which might rule out the importance of the urea carbonyl group as H-bond acceptor. Nevertheless, the loss in H-bond capabilities for the thiourea derivative is more than compensated by an increase in its lipophilicity (4.28 vs 5.07). Lipophilicity has been shown in numerous QSAR studies to be a general predictive descriptor for high P-gp inhibitory activity.<sup>36,37,38</sup>

We thus calculated logP values using the software Bio-Loom<sup>39</sup> and correlated them with logIC<sub>50</sub> values (Figure 2). The r<sup>2</sup> value of 0.56 demonstrates that also in the series of benzophenones lipophilicity plays a dominant role. This is in agreement with the notion that compounds most probably enter the binding cavity of P-gp directly from the membrane bilayer. This is additionally supported by the recent X-ray structure of mouse

P-gp, which shows a large inner cavity accessible from the membrane via putative entry ports composed of transmembrane helices 4/6 on one side and 10/12 on the other side.<sup>18</sup>



**Figure 2.** Correlation of P-gp inhibitory activity of compounds **6-24** (expressed as  $\log 1/IC_{50}$  values) vs. calculated  $\log P$  values of the ligands.

The  $\text{clogP}/\log$  (potency) plot further supports our hypothesis about the urea/thiourea compound pair. Urea derivative **21** is located above the correlation line which indicates that it exhibits higher biological activity than would be expected solely from its  $\text{clogP}$  value (0.10 calcd vs 0.32 obs), indicating an additional H-bond mediated by the carbonyl group. The thiourea derivative **22** lies much closer to the line. The 4-hydroxy-4-phenylpiperidine analog **19** is also located above the  $\text{clogP}/\log IC_{50}$  correlation line (-0.24 calcd vs 0.51 obs), which further confirms our previous results on the importance of the 4-hydroxy-4-phenylpiperidines moiety for high biological activity of propafenone derivatives.<sup>34</sup> These results were recently supported by extensive docking studies of propafenone analogs.<sup>17</sup> It is also interesting to note that the homodimer **23** is by a factor of 15 more active than predicted by the  $\text{clogP}/\log IC_{50}$  plot (0.09 calcd vs 1.28 obs). A pair wise comparison of equilipophilic compounds **23** vs **21** ( $\text{clogP}$ : 4.27 vs 4.28;  $IC_{50}$ : 0.05 vs 0.48) and **19** vs **20** ( $\text{clogP}$ : 3.65 vs 3.64;  $IC_{50}$ : 0.31 vs 1.21) indicates that mutual activity differences might also be due to difference in molecular size. The dimer **23** (44 heavy atoms) is about one order of magnitude more active than **21** (35 heavy atoms). Similarly, **19** (32 heavy atoms) is about a factor of four more active than **20** (24 heavy

atoms). This also points towards a commonly observed phenomenon in lead optimisation programmes, i.e. activity increases with the size of the molecules. Therefore, ligand efficiency (LE)<sup>19-21</sup> and lipophilic efficiency (LipE),<sup>22,23</sup> profiles of inhibitors/substrates of P-gp have been used to identify the derivatives with the best activity/size (or logP) ratio, which should provide further insights for the design of new ligands.<sup>24,25</sup>

**Ligand Efficiency (LE)**, most commonly defined as the ratio of free energy of binding over the number of heavy atoms, is a simple metric for assessing whether a ligand derives its potency from optimal fit with the target protein or simply by virtue of making many contacts.<sup>40</sup> In order to get more information on the most promising P-gp inhibitors and to compare them to well established P-gp inhibitors/substrates, we calculated ligand efficiency values of benzophenones **6-24**, selected propafenone analogs, as well as P-gp inhibitors which entered clinical studies. Ligand efficiencies were calculated as described in the methods section. For benzophenones, small ligands such as the N-propyl derivative **20** and the piperidine analog **12** show higher efficiency values (0.35; 0.33) than the large dimers **23** and **24** (0.23; 0.18). For the whole data set it can be observed that ligand efficiencies drop dramatically when the size of the ligands increases above 50 heavy atoms (Figure 3). A similar trend has been observed in the literature, with LE showing generally a dependency on ligand size.<sup>20</sup> As LE in principle is supposed to normalize for the size of the ligand, various proposals have been made to solve this problem.<sup>41,21</sup> As the heavy atom count of the ligands in our data set varies from 24 to 86 (**20**; valsopodar), LE values were subsequently scaled as described by Reynolds *et al.*,<sup>20,21</sup> to retrieve a size-independent ligand efficiency value (LE\_Scale). This was achieved by fitting the top ligand efficiency versus heavy atom count to a simple exponential function, as outlined by Reynolds *et al.*,<sup>20</sup> (Equ 1; Figure 3). Subsequently, the ratio of ligand efficiency over normalized ligand efficiency scale gives a scoring function called “Fit Quality” (FQ) (Equ. 2). According to Reynolds *et al.*, Fit Quality scores close to 1.0 or above indicate near optimal ligand binding, while low fit quality scores are indicative of sub-optimal binding.

$$\text{LE\_Scale} = 0.104 + 0.65 e^{-0.037 \cdot \text{HA}} \quad (\text{Equ. 1})$$

$$FQ = LE / LE\_Scale \quad (\text{Equ. 2})$$

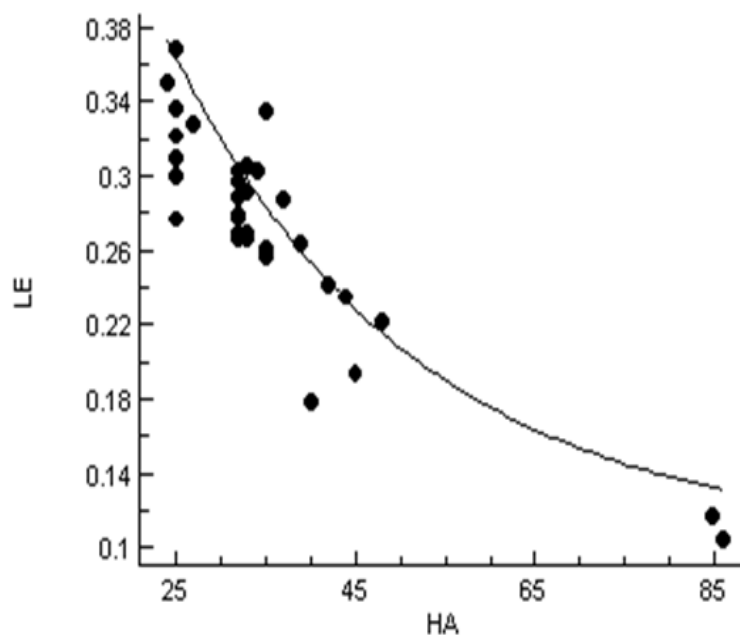
Use of this criterion shows that most of the compounds under clinical investigation show FQ scores above 1, including zosuquidar, ONT093, elacridar and tariquidar, along with benzophenones **6** and **23**, as well as propafenone and its analogs GPV0062 and GPV0576 (Figure 1; Table 2).

**Table 2.** Pharmacological activities, ligand efficiency (LE) and lipophilic efficiency (LipE) profiles of selected propafenones and P-gp inhibitors which entered in clinical studies.

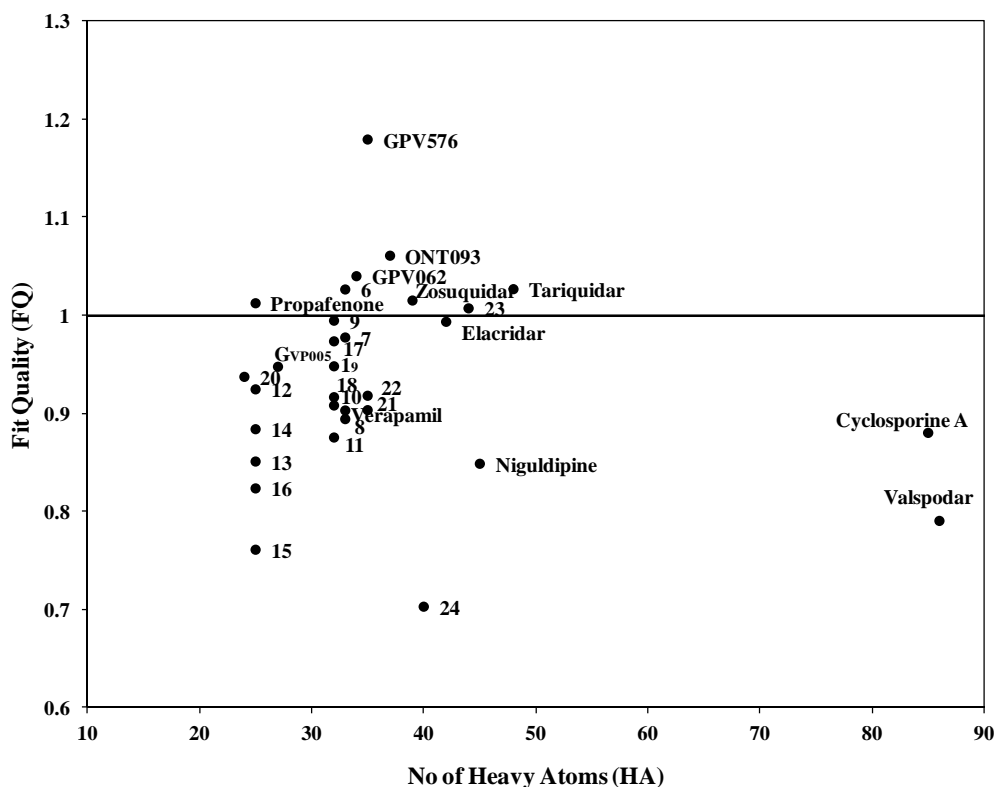
| Comp.          | pIC <sub>50</sub> | HA | LE   | clogP | LipE  |
|----------------|-------------------|----|------|-------|-------|
| Verapamil      | 6.24              | 33 | 0.27 | 4.47  | 1.77  |
| Elacridar      | 7.14              | 42 | 0.24 | 4.21  | 2.93  |
| Tariquidar     | 7.48              | 48 | 0.22 | 5.55  | 1.93  |
| Zosuquidar     | 7.23              | 39 | 0.26 | 4.96  | 2.27  |
| ONT093         | 7.50              | 37 | 0.29 | 7.30  | 0.19  |
| Valspodar      | 6.30              | 86 | 0.10 | 15.09 | -8.79 |
| Cyclosporine A | 6.99              | 85 | 0.12 | 14.36 | -7.37 |
| Niguldipine    | 6.15              | 45 | 0.20 | 7.80  | -1.65 |
| Propafenone    | 6.48              | 25 | 0.37 | 3.64  | 2.84  |
| GPV576         | 8.25              | 35 | 0.33 | 6.02  | 2.23  |
| GPC0062        | 7.24              | 34 | 0.30 | 4.15  | 3.09  |
| GPV005         | 6.22              | 27 | 0.33 | 4.38  | 1.84  |

It is interesting to note that especially those compounds which were specifically designed as P-gp inhibitors (ONT093, zosuquidar, elacridar, tariquidar) show higher FQ values than those originating from drug repurposing attempts (verapamil, cyclosporine and its analog valspodar). With respect to propafenone analogs, GPV0576 is the hitherto most active analog we synthesized showing a highly lipophilic but quite compact substituent at the nitrogen atom (4-tolylpiperazine). Interestingly, the top ranked benzophenone analog **6** also has a 4-tolylpiperazine moiety. This might point towards the tolylpiperazine substituent for being a privileged substructure for P-gp inhibitors. GPV0062 bears a 4-hydroxy-4-phenyl piperidine moiety, which has been shown to influence biological activity independent of lipophilicity, resulting in an

almost tenfold increase of inhibitory activity when compared to compounds having other substituents at the nitrogen atom. This points towards a distinct additional interaction mediated by the 4-hydroxy group, most probably in form of a hydrogen bond. Finally, propafenone itself shows a very good value, thus retrospectively demonstrating its validity as starting point for structural modifications.



**Figure 3.** Plot of ligand efficiency versus heavy atom count for benzophenone analogs, compounds which entered clinical studies and selected propafenones.

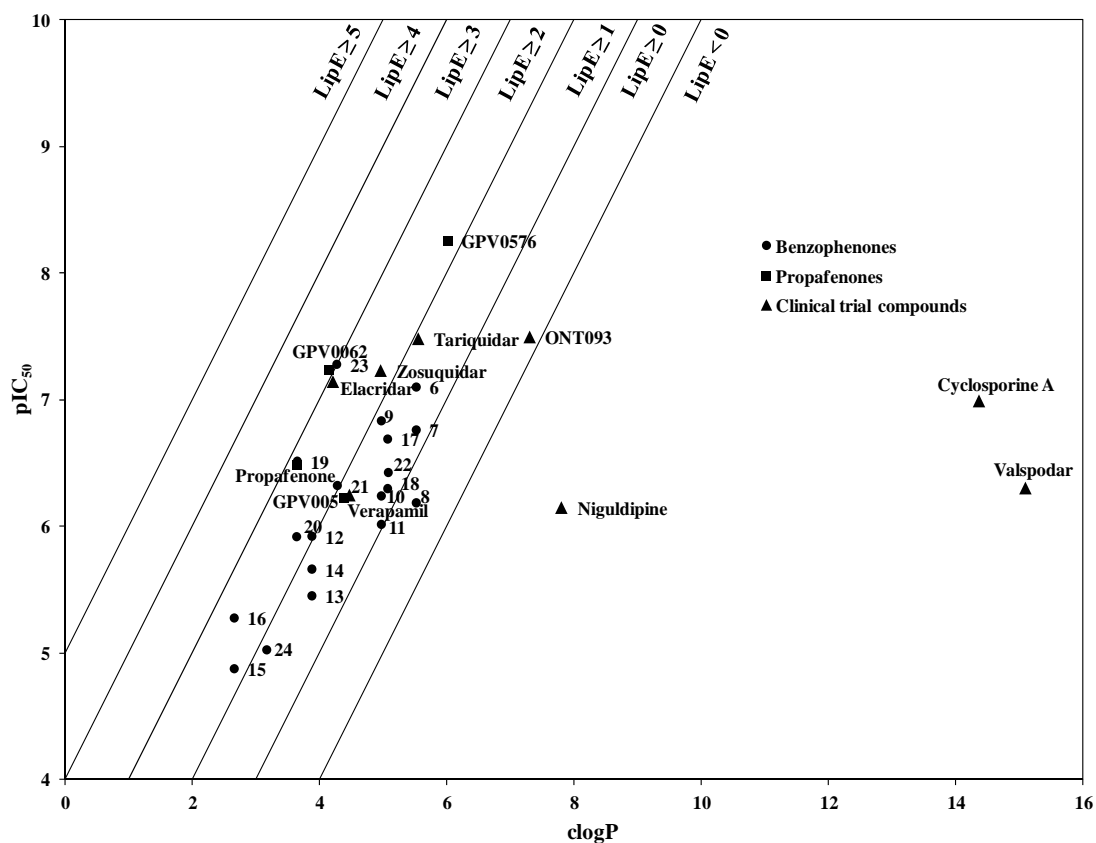


**Figure 4.** Fit quality scores around 1 indicate a near optimal ligand binding affinity for a given number of heavy atoms.

As already outlined, lipophilicity has been shown in numerous studies to be a general predictor for high P-gp inhibitory activity. This most probably is due to the proposed access path of the compounds, which seems to be directly from the membrane bilayer. On the other hand, high lipophilicity is very often associated with poor oral drug-like properties. This led to the assumption that clog P values between 2 and 3 are considered optimal in an oral drug program and prompted Leeson *et al*, to introduce the concept of lipophilic efficiency.<sup>22</sup>

**Lipophilic Efficiency (LipE)** is a parameter that combines both potency and lipophilicity and is defined as a measure of how efficiently a ligand exploits its lipophilicity to bind to a given target. Briefly, in a lead optimization series there is a greater likelihood of achieving good in vivo performance when potency can be increased without increasing logP or logD values. To explore this concept also for P-gp inhibitors, we calculated LipE values for the whole set of benzophenones as well as for the compounds used for the LE study (Table 2). The clogP values vary from 2.66 to

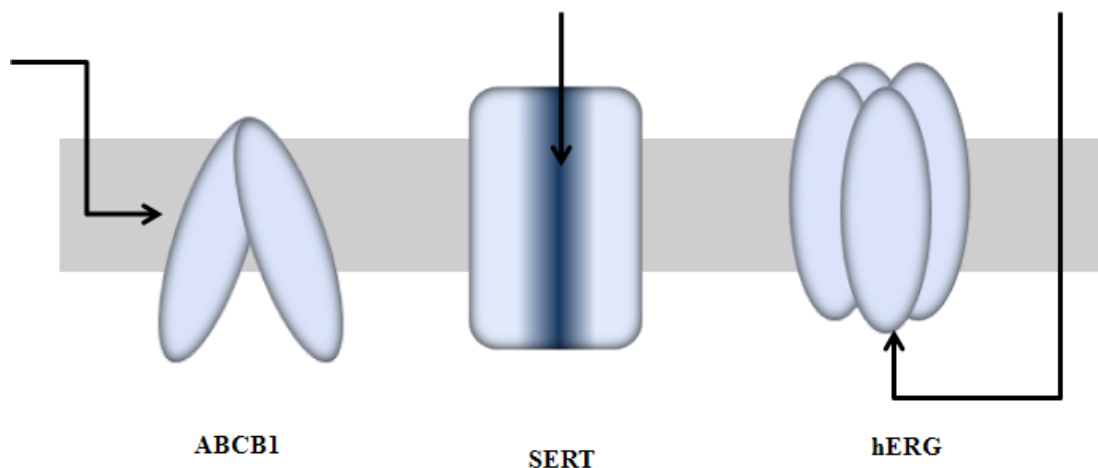
15.09, leading to a lipophilic efficiency range between -8.79 and +3.08. This is somewhat surprising as it has been reported that a lipophilic efficiency greater than 5 combined with clog P values between 2 and 3 is considered optimal for a promising drug candidate.<sup>22,23</sup> None of the clinically tested P-gp inhibitors fulfils these requirements. Only the 4-hydroxy-4phenyl-piperidine analogous propafenone GPV0062 as well as the dimer **23** exhibit values slightly higher than 3. All other compounds show values lower than 3 (Figure 5). It is tempting to speculate whether this is due to the unique entrance pathway directly from the membrane bilayer, which requires a different logP profile than for compounds which access their binding site directly from the extracellular or intracellular aqueous compartment.



**Figure 5.** Plot of clogP versus biological activity of inhibitors of P-gp; LipE values higher than 5 are considered to be the threshold for compounds of clinical interest.

To study in more detail whether the unique access path of P-gp inhibitors directly from the membrane bilayer is linked to this unexpectedly low LipE values, we studied the distribution of LipE profiles for a set of targets showing different access pathways

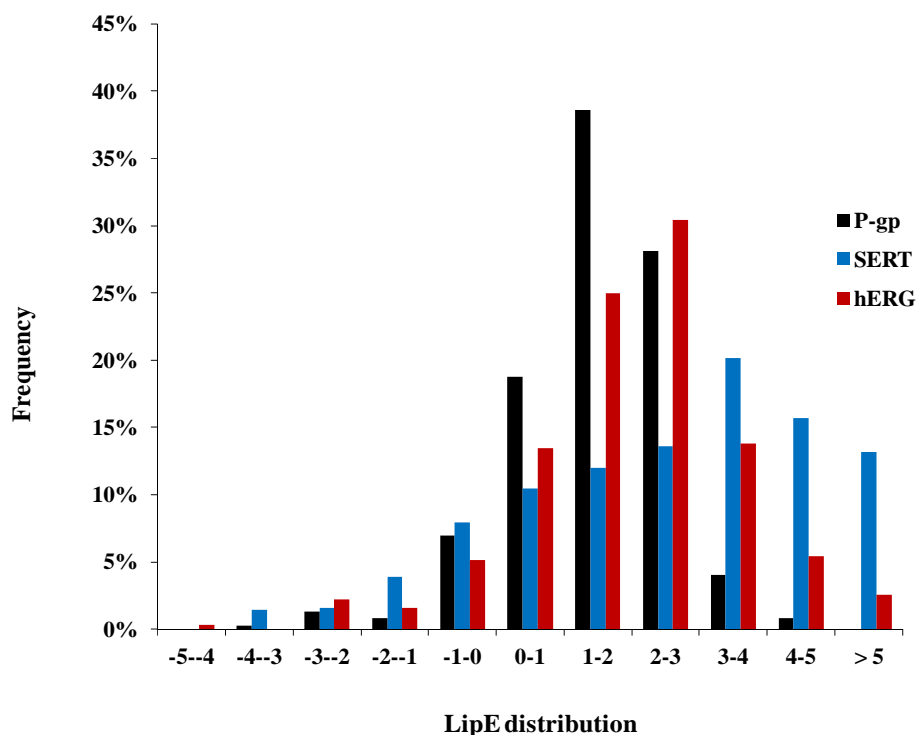
of their ligands: P-glycoprotein (via membrane bilayer), the serotonin transporter SERT (from the extracellular environment), and the hERG potassium channel hERG (from the cytoplasm) (Figure 6). LipE values of inhibitors of SERT (extracted from the ChEMBL data base),<sup>42</sup> hERG blockers,<sup>43</sup> and propafenone-type inhibitors of P-gp (in-house data) were calculated as described in the method section.



**Figure 6.** Schematic representation of access of inhibitors/substrates to the binding sites of P-gp, SERT and hERG along three different pathways. Ligands of P-gp approach the binding cavity via the membrane bilayer, in SERT the ligands get access from the extracellular environment, while in hERG this access occurs via the cytoplasm.

The LipE distribution profile of SERT inhibitors extracted from the ChEMBL data base identified about 13% of the compounds that cross the LipE threshold of 5 (Figure 7). These compounds cover a wide range of activity (0.01 nM- 10 mM) and clogP (-3.42 to 4.66) (SM Figure 1). Moreover, 15 SERT inhibitors have been identified with clogP ~2.5, LiPE > 5 and IC<sub>50</sub> < 10 nM. However, none of them was listed as a marketed drug. In case of hERG only 2.5 % of the compounds cross the LipE threshold of 5. They showed a potency distribution from 5 nM to 18 μM and clogP values between -0.77 and 2.21 (SM Figure 1). Only two compounds, almokalant and dofetilide, complied with the desired profile (clogP~2.5, LipE > 5, potency values < 10 nM). Dofetilide is a registered class III antiarrhythmic agent, while almokalant is in phase II clinical investigations.<sup>44,45</sup>





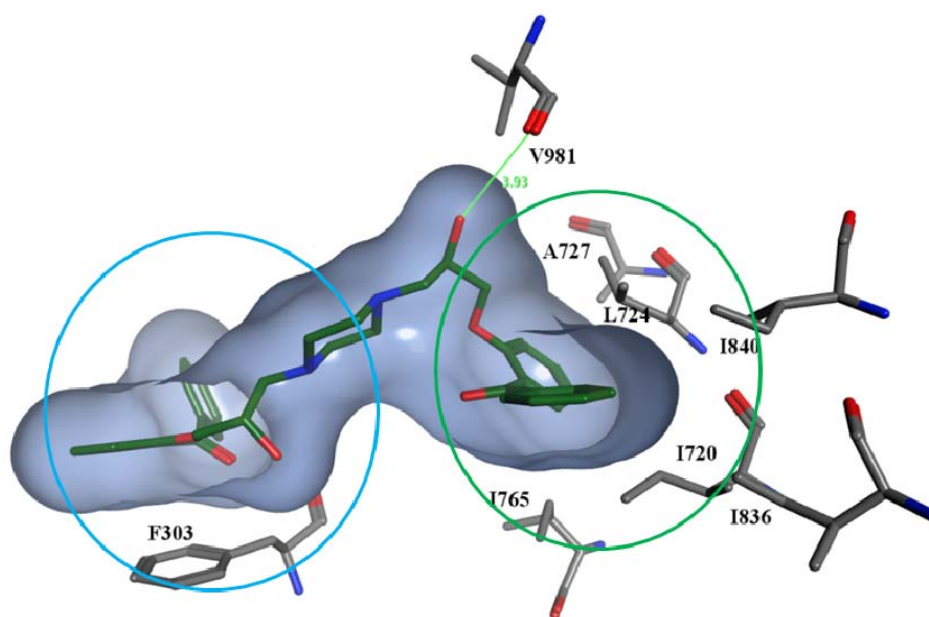
**Figure 7.** LipE distribution profiles of ligands of P-gp, SERT and the hERG potassium channel.

LipE profiles of P-gp inhibitors could not identify any compound that reaches the standard threshold value of 5. Most of the ligands fall in the LipE range of 1-2 (39%) or 2-3 (28%) with wide a range in distribution of their clogP (1.32 to 15.09) as well as IC<sub>50</sub> (5.6 nM to 1.8 mM) values (SM figure 1). Thus, the LipE threshold for ligands of P-gp needs to be reconsidered. Nevertheless, from the benzophenone data set presented here, compounds **15**, **16**, **19**, **20**, and **23**, might be the most promising ones as their LipE values are between 2 and 3, a range where most of the compounds which in the past entered clinical trials are located.

**Docking into a homology model of P-glycoprotein.** To get insights into the potential binding mode of propafenone-type benzophenones we selected compounds **6**, **19**, **20** and the dimer **23** for further *in silico* studies. Compounds **19**, **20**, and **23** were selected as they are ranked high both in LipE and FQ scores, and **6** was additionally included as it is top ranked with respect to FQ. Interestingly, this selection resembles the key features observed for propafenone analogs: compound **6** shows a 4-tolylpiperazine substituent (analogous to GPV0576), compound **19** is analogous to

GPV0062 (4-hydroxy-4-phenyl-piperazine) and derivative **20** is the direct propafenone analog (N-propyl). The docking protocol follows those previously published<sup>46</sup> and is provided in detail in the methods section.

The analysis of the interaction pattern of selected docking poses indicates that the benzophenone scaffold interacts with F343 and F303 near the entry gate, whereas the lipophilic substituents in the vicinity of the basic nitrogen atom are surrounded by hydrophobic amino acid residues L724, I720, V981, I840, I836 and I765 located at TM 7, 9, and 12 (Figure 8). This further supports the importance of high lipophilicity and also is in line with previous studies performed by Pajeva and Wiese, who showed that for a series of inhibitors of P-gp hydrophobicity represents a space directed molecular property rather than a simple overall descriptor.<sup>47</sup> The top ranked cluster of poses are in close vicinity of our previously purposed binding positions for benzopyrano[3,4-*b*][1,4]oxazines, where compounds having 4aS,10bR configuration interact mainly with amino acid residues of TM4, 5 and 6 near the entry gate, while compounds having 4aR,10bS configuration are positioned deeper inside the binding cavity, being mainly surrounded by hydrophobic amino acid residues of TM7, 8, 9 and 12.<sup>46</sup> Interestingly, the top scored dimer **23** is positioned in a way to bridge these two positions (Figure 8). Moreover, this pose might also aid in the explanation for the activity differences of homodimer **23** (0.05  $\mu$ M) and heterodimer **22** (9.48 $\mu$ M): The additional benzene ring in the best scored pose of homodimer **23** is surrounded by several hydrophobic amino acids (I836, L720, I840 and L724). Overall, benzophenones shared a similar interaction profile as propafenones. Amino acids S952, F434, F336, L721 and Y307 have been identified as common interacting amino acid residues of all three classes of propafenone type inhibitors of P-gp (SM Figure 3).



**Figure 8.** Ligand-protein interaction profile of the best cored pose of benzophenone dimer **23**. Blue circle represent the putative position of benzopyrano[3,4-*b*][1,4]oxazines having 4aS,10bR configuration, while the green circle indicates the position of diastereoisomers with 4aR,10bS configuration.

Selected benzophenone analogs have been previously used as photo-affinity ligands to characterize the drug-binding domain of propafenone-type analogs. In these studies, TM 3, 5, 6, 8, 10, 11, and 12 were identified as potential interacting helices.<sup>28,29,48,49</sup> This is well in line with our docking studies, which show main interactions with TM 5 and 6 near the entry gate and TM 7, 8 and 12 deeper inside the cavity (SM figure 4). No significant cluster of poses has been identified on the second wing (2/11 interface), which might be due to the asymmetry in the template used for building the homology model of P-gp, thus narrowing the available space at this side.

### Conclusions

Calculation of ligand efficiency and lipophilic efficiency values for a set of P-gp inhibitors shows that ligands of P-gp exhibit LipE values below the threshold of 5 considered to be optimal for clinical candidates. This might be due to the unique entrance pathway of these classes of compounds, taking a route directly from the membrane bilayer. However, LipE and LE values of benzophenones **6**, **19**, **20**, as well as of the dimer **23**, are close to compounds which entered clinical studies, thus

qualifying them for further studies. Docking studies further strengthen the evidence provided by QSAR studies that the benzophenones bind to the same region as propafenone-type inhibitors. Moreover, the dimer **23** seems to bridge the two distinct binding sites recently proposed for benzopyrano[3,4-*b*][1,4]oxazines. This further supports the general assumption of a binding zone with distinct, but overlapping binding sites for individual scaffolds as a basis for the promiscuity of P-gp.

## Experimental Section

### Chemistry

**Material and Methods.** The data set used consists of a set of previously published benzophenones **9**,<sup>34</sup> **12**, **19** and **20**<sup>28</sup> well as a series of newly synthesised analogs. Melting points were determined on Leica Galen III (ser. no. 1413 WT) and are uncorrected. Elemental analysis was performed at micro analytical laboratory of institute of physical chemistry (Mag. Johannes Theiner); University of Vienna. The used equipment was a “2400 CHN-Elemental Analyzer” Perkin Elmer. Mass spectra were recorded on a Maldi-TOF, Kratos-instruments, matrix assisted laser-desorption-ionization time of flight, reflection mass spectrometer. NMR spectra were recorded on a Bruker spectropin for 200 MHz <sup>1</sup>H-NMR and 50 MHz for <sup>13</sup>C-NMR. CDCl<sub>3</sub> and DMSO at room temperature were used as internal standards. Column chromatographic separations were performed by using silica gel 60 (Particle size 40-63µm, 230-300 mesh) from J.T. Baker or Merck.

**General procedure for the preparation of (2-Oxiranylmethoxy-phenyl)-phenyl-methanone C<sub>16</sub>H<sub>14</sub>O<sub>3</sub> (2a).** 10 g (51mmol) of 2-hydroxy-benzophenone was dissolved in epichlorohydrine (120 mL), treated with 2.04 g (51mmol) sodium hydroxide and refluxed for 24 h. After cooling, the residue was filtered off and washed with diethyl ether. Subsequently the solvent was removed by rotary evaporation. The remaining oil was taken up in diethyl ether and washed with water (175 mL). The organic phase was then dried over anhydrous sodium sulfate. After removal of the solvent by rotary evaporation yellow oil was obtained, yield 12.81g (98.85%); <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.37-2.66 (m, 2H, CH<sub>2</sub>-O-CH), 2.96 – 3.00 (m, 1H, CH), 3.92- 4.15 (m, 2H, Ar-O-CH<sub>2</sub>), 6.97-7.81 (m, 9H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 44.03 (m, 2H, CH-CH<sub>2</sub>-O), 49.59

(CH), 68.62 (Ar-O-CH<sub>2</sub>), 112.72, 121.13, 128.05 (arom C), 129.00 (Ar-CO); 129.38, 129.69, 131.97, 132.66 (arom C); 137.93 (Ph-CO), 156.08 (Ar-O), 196.15 (CO).

**General procedure for the preparation of (3-Oxiranylmethoxy-phenyl)-phenylmethanone C<sub>16</sub>H<sub>14</sub>O<sub>3</sub> (2b).** 5 g (25.25 mmol) of 3-Hydroxy-benzophenone was dissolved in epichlorohydrine (60 mL), treated with 1.01 g (25.25 mmol) sodium hydroxide, refluxed for 6 h and stirred over night. The residue was filtered off and washed with diethyl ether. After removal of the solvents under reduced pressure, the resulting oil was taken up in diethyl ether and washed with water several times. The organic layers were combined, dried over anhydrous sodium sulfate and evaporated to dryness yielding yellow oil. For further purification a column chromatography (Silica gel, ether/ petrol ether, 70+30) was performed. Subsequent removal of the solvents under reduced pressure gave white opalescent oil, yield 6g (93.6%); <sup>1</sup>H-NMR: (CDCl<sub>3</sub>) δ 2.71-2.74 (m, 1H, H<sub>A</sub>), 2.86 (t, 1H, J = 4.42, H<sub>B</sub>), 3.29- 3.37 (m, 1H, CH), 3.93 (dd, 1H, J = 5.94/11.12, H<sub>X</sub>), 4.28 (dd, H, J = 2.90/10.98, H<sub>Y</sub>), 7.10-7.78 (arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 44.34 (CH<sub>2</sub>-O), 49.81 (CH), 68.76 (Ar-O-CH<sub>2</sub>), 114.93, 119.17, 123.10, 128.11, 129.18, 129.81, 132.31, (arom C), 137.28, 138.70 (C), 158.24 (Ar-O), 196.08 (CO).

**General procedure for the preparation of (4-Oxiranylmethoxy-phenyl)-phenylmethanone C<sub>16</sub>H<sub>14</sub>O<sub>3</sub> (2c).** 6 g (30.30 mmol) of 4-Hydroxy-benzophenone was dissolved in epichlorohydrine (50 mL), treated with 2 g (50 mmol) sodium hydroxide, refluxed for 5 h and stirred over night. The residue was filtered off and washed with diethyl ether. Subsequently the solvent was removed by rotary evaporation. The remaining opalescent oil was taken up in diethyl ether and washed with water several times. The organic phase was then dried over anhydrous sodium sulfate. After removal of the solvents by rotary evaporation a white solid was obtained, yield 6.7g (87.54%); <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.76 (dd, 1H, J=2.66/4.8, H<sub>A</sub>), 2.92(t, 1H, J=4.64, H<sub>B</sub>), 3.35-3.40 (m, 1H, CH ), 3.96-4.0 (dd, 1H, J=5.81-11.12, H<sub>X</sub>), 4.29-4.36 (dd, 1H, J = 2.91/11.11, H<sub>Y</sub>), 6.95-7:00 (m, 2H, H-3, H-5), 7.42-7.56 (m, 5H, H-2, H-6, H-3', H-4', H-5'), 7.72-7.83 (m, 2H, H-2', H-6'); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 44.51 (CH<sub>2</sub>-O), 49.84 (CH), 68.82 (Ar-O-CH<sub>2</sub>), 114.07, 128.14, 129.65, 131.90, 132.46, (arom C), 130.54 (Ar-CO), 138.06 (Ph-CO), 161.93 (Ar-O), 195.41 (CO).

**General procedure for the preparation of [2-(2-Hydroxy-3-piperazine-1-yl-propoxy)-phenyl]-phenyl-methanone (5).** 1.82 g (7.20 mmol) of (2-Oxiranylmethoxy-phenyl)-phenyl-methanone (2a) was dissolved in 20-30 mL of methanol, added 1.6 g (18.6 mmol) piperazine and then the reaction mixture was refluxed for 5 h. After removal of the solvent by rotary evaporation a column chromatography was performed (silica gel, CH<sub>2</sub>Cl<sub>2</sub>/methanol/concentrated NH<sub>3</sub>, 100/10/1) subsequent evaporation to dryness yielded 1.88 g (77.33%) yellow oil which was solidified on cooling. ; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.01- 2.42 (m, 6H, CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 2.77-2.81 (m, 4H, (CH<sub>2</sub>)<sub>2</sub>-N), 3.68- 3.75 (m, 1H, CH), 3.90-3.94 (m, 2H, O-CH<sub>2</sub>), 6.96-7.79 (m, 9H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 45.86 (N-(CH<sub>2</sub>)<sub>2</sub>), 54.30 ((CH<sub>2</sub>)<sub>2</sub>-NH), 60.79 (CH<sub>2</sub>-N), 65.27 (CH), 70.76(O-CH<sub>2</sub>), 112.56, 121.05, 128.28, 129.51, 130.07, 132.33, 132.83 (aromC), 128.79 (Ar-CO), 138.36(Ph-CO), 156.55 (arom C-O), 196.56 (CO).

**(2-{3-[4-(2,3-xylyl)-piperazin-1-yl]-2-hydroxy-propoxy}-phenyl)-phenyl-methanone (6).** 700 mg (2.75 mmol) of (2-Oxiranylmethoxy-phenyl)-phenyl-methanone **2a** was dissolved in 15 mL of methanol and treated with 526 mg (2.75 mmol) 1-(2-3-xylyl)-piperazine. The mixture was refluxed for 24 h. Subsequent removal of the solvent yielded yellow oil, crystallization from ethyl acetate / diethylether gave 904 mg (73.8%) white crystals; mp 116-118°C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.19-2.27(m, 2H, CH<sub>2</sub>), 2.21 (s, 3H, CH<sub>3</sub>), 2.27 (s, 3H, CH<sub>3</sub>), 2.42- 2.47 (m, 2H, CH<sub>2</sub>), 2.58-2.66 (m, 2H, CH<sub>2</sub>), 2.83-2.88 (m, 4H, -(CH<sub>2</sub>)<sub>2</sub>-N-Xyl), 3.77-3.83 (m, 1H, CH), 3.91-4.02 (m, 2H, O-CH<sub>2</sub>), 6.90 (dd, 2H, J=2.9/7.83 Hz; arom H), 6.99-7.12 (m, 3H, arom H), 7.44-7.58 (m, 5H, arom H), 7.79-7.83( m, 2H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 13.93 (CH<sub>3</sub>-ortho), 20.61 (CH<sub>3</sub>-meta), 51.97 (N-(CH<sub>2</sub>)<sub>2</sub>), 53.79 (CH<sub>2</sub>-N), 60.36 ((CH<sub>2</sub>)<sub>2</sub>-N-Ph), 65.42(CH), 70.75(O-CH<sub>2</sub>), 112.64, 116.56, 121.12, 125.01, 125.81, 128.33, 128.86, 129.58, 130.12, 132.37, 132.87 (arom C), 137.99 (Ar-CO), 138.43 (Ph-CO), 151.29 (Ar-N), 156.58 (Ar-O). Anal. Calcd for C<sub>28</sub>H<sub>32</sub>N<sub>2</sub>O<sub>3</sub>: C, 74.89; H, 7.29; N, 6.24. Found: C, 74.73; H, 7.59; N, 6.32.

**(3-{3-[4-(2,3-xylyl)-piperazin-1-yl]-2hydroxy-propoxy}-phenyl)-phenyl-methanone (7).** 560 mg (2.20 mmol) of 3-Oxiranylmethoxy-phenyl)-phenyl-methanone **2b** was dissolved in 20 mL of methanol and treated with 418 mg (2.20 mmol) 1-(2, 3-Xylyl)-piperazine. The mixture was refluxed for 6 h. After removal of the solvent under

reduced pressure, a yellow oil was obtained which crystallized from ethyl acetate/diethyl ether yielding 960 mg (98%) white crystals; mp 92-99°C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.23, 2.27 (2s, 6H, 2CH<sub>3</sub>), 2.62-2.66 (m, 4H, N-(CH<sub>2</sub>)<sub>2</sub>), 2.83-2.93 (m, 6H, CH<sub>2</sub>-N, (CH<sub>2</sub>)<sub>2</sub>-N-Ph), 4.08-4.18 (m, 3H, O-CH<sub>2</sub>-CH), 6.92 (d, 2H, J = 7.07, H-4, H-6), 7.05-7.20 (m, 2H, arom H), 7.37-7.79 (m, 6H, arom H), 7.80-7.83 (m, 2H, arom H); <sup>13</sup>C-NMR: (CDCl<sub>3</sub>) δ 13.92 (CH<sub>3</sub>-ortho), 20.60 (CH<sub>3</sub>-meta), 52.15 (N-(CH<sub>2</sub>)<sub>2</sub>), 53.77 ((CH<sub>2</sub>)<sub>2</sub>-N), 60.38 (CH<sub>2</sub>-N), 65.38(CH), 70.49(O-CH<sub>2</sub>), 115.07, 116.58, 119.30, 123.09, 125.03, 125.81, 128.24, 129.26, 130.00 (arom C); 131.19 (C), 132.42 (arom C), 137.52, 137.97, 138.86, 151.31, (C), 158.70 (Ar-O), 196.41(CO); MS m/e 444.68(M, 90%). Anal. Calcd for C<sub>28</sub>H<sub>32</sub>N<sub>2</sub>O<sub>3</sub>·0.3H<sub>2</sub>O: C, 74.64; H, 7.31; N, 6.22. Found: C, 74.76; H, 7.56; N, 6.08.

**(4-{3-[4-(2,3-Xylyl)-piperazin-1-yl]-2-hydroxy-propoxy}-phenyl)-phenyl-methanone (8).** 700 mg (2.76 mmol) of (4-Oxiranylmethoxy-phenyl)-phenyl-methanone **2c** was dissolved in 15 mL of methanol, treated with 523.6 mg (2.756 mmol) 1-(2, 3 Xylyl)-piperazin and the mixture was refluxed for 5 h. Removal of the solvent under reduced pressure yielded white crystals which were recrystallized from methanol giving 1.06 g (94.62%) white crystals; mp 122-126°C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.23-2.27 (2s, 6H, 2CH<sub>3</sub>), 2.55-2.70 (m, 4H, (CH<sub>2</sub>)<sub>2</sub>-N-Ph), 2.70- 3.00 (m, 6H, CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 3.40-3.70 (bs, 1H, OH), 4.10-4.17 (m, 3H, O-CH<sub>2</sub>-CH), 6.90-7.03 (m, 5H, arom H), 7.44 -7.51 (m, 3H, arom H), 7.74-7.81( m, 4H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 13.90 (CH<sub>3</sub>- ortho), 20.58 (CH<sub>3</sub>- meta), 52.15, 53.76 (CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 60.38((CH<sub>2</sub>)<sub>2</sub>-N), 65.31(CH), 70.44 (O-CH<sub>2</sub>), 114.08, 116.57, 125.05, 125.81, 128.15, 129.69 (arom C), 130.37, 131.88 (C), 131.88, 132.50 (arom C), 137.97, 138.18 (C), 151.29 (Ar-N), 162.34 (Ar-O), 195.50(CO); MS m/e 445.4(M<sup>+</sup>, 100%). Anal. Calcd for C<sub>28</sub>H<sub>32</sub> N<sub>2</sub>O<sub>3</sub>: C, 75.65; H, 7.26; N, 6.30. Found: C, 75.61; H, 7.48; N, 6.33.

**(3-{3-[4-(4-Fluoro-phenyl)-piperazin-1-yl]-2hydroxy-propoxy}-phenyl)-phenyl-methanone (10).** 500 mg (1.97 mmol) of (3-Oxiranylmethoxy-phenyl)-phenyl-methanone **2b** was dissolved in 15 mL of methanol, treated 360 mg (2 mmol) p-F-phenyl-piperazine and the mixture was refluxed for 6 h. The solution was allowed to cool down and stirred at room temperature overnight. The obtained solid was filtered off and washed with diethyl ether giving 690 mg (80.76%) white crystals; mp 131-133

°C;  $^1\text{H-NMR}$  ( $\text{CDCl}_3$ )  $\delta$  2.60-2.85 (m, 6H,  $\text{CH}_2\text{-N-(CH}_2)_2$ ), 3.12-3.17 (m, 4H,  $(\text{CH}_2)_2\text{-N-Ph}$ ), 4.05-4.18 (m, 3H,  $\text{O-CH}_2\text{-CH-OH}$ ), 6.87-6.97(m, 4H, arom H), 7.19-7.25 (m, 1H, arom H), 7.37-7.60 (m, 6H, arom H), 7.78-7.82(m, 2H, arom H);  $^{13}\text{C-NMR}$ : ( $\text{CDCl}_3$ )  $\delta$  50.18, 53.28, 60.30 ( $\text{CH}_2\text{-N-(CH}_2)_4$ ), 65.46 (CH), 70.36 ( $\text{O-CH}_2$ ), 115.04, 115.29, 115.73, 117.81, 117.97, 119.28, 123.15, 128.24, 129.99, 132.43 (arom C), 137.48 (Ar-CO), 138.87 (Ph- CO), 147.71 (Ar-N), 157.22 (J = 239 Hz, C-F), 158.64(Ar-O), 196.39(CO); MS m/e 434.77 (M, 100%). Anal. Calcd for  $\text{C}_{26}\text{H}_{27}\text{FN}_2\text{O}_3$ : C, 71.87; H, 6.26; N, 6.45. Found: C, 71.65; H, 6.45; N, 6.46.

**(4-{3-[4-(4-Fluoro-phenyl)-piperazin-1-yl]-2-hydroxy-propoxy}-phenyl)-phenyl-methanone (11).** 700 mg (2.75 mmol) of (4-Oxiranylmethoxy-phenyl)-phenyl-methanone **2c** was dissolved in 15mL of methanol, treated with 496 mg (2.75 mmol) p-F-phenyl-piperazine and refluxed for 5 h. Removal of solvent under reduced pressure left a yellow oil which crystallized from iso-propanol yielding 1 g (83.6%) white crystals; mp 94-97 °C  $^1\text{H-NMR}$  ( $\text{CDCl}_3$ )  $\delta$  2.61-2.69 (m, 4H,  $\text{N-(CH}_2)_2$ ), 2.80-2.91 (m, 2H,  $\text{CH}_2\text{-N}$ ), 3.12- 3.17 (m, 4H,  $(\text{CH}_2)_2\text{-N-Ph}$ ), 3.56 (br, 1H, OH), 4.11- 4.19 (m, 3H,  $\text{CH}_2\text{-CH}$ ), 6.89-7.85 (m, 13H, arom H);  $^{13}\text{C-NMR}$ : ( $\text{CDCl}_3$ )  $\delta$  50.24, 53.28, 60.28 ( $\text{CH}_2\text{-N-(CH}_2)_4$ ), 65.40 (CH), 70.33 ( $\text{O-CH}_2$ ), 114.07, 115.31, 115.75, 117.81, 117.96, 128.17, 129.71, (arom C), 130.42 (Ar-CO), 131.92, 132.52 (arom C), 138.15 (Ph-CO), 147.77 (Ar-N), 162.28 (Ar-O), 195.50 (CO); MS m/e 435.3 ( $\text{M}^+$ , 100%). Anal. Calcd for  $\text{C}_{26}\text{H}_{27}\text{FN}_2\text{O}_3$ : C, 71.87; H, 6.26; N, 6.45. Found: C, 71.61; H, 6.43; N, 6.41.

**[3-(2-Hydroxy-3-piperidin-1-yl-propoxy)-phenyl]-phenyl-methanone (13).** 700 mg (2.75 mmol) of (3- Oxiranylmethoxy-phenyl)-phenyl-methanone was dissolved in 15 mL of methanol and treated with 234 mg (2.75 mmol) piperdine. The mixture was refluxed for 4 h and stirred at room temperature overnight. After removal of the solvent under reduced pressure 336 mg (36%) yellow oil was obtained; mp 64 °C;  $^1\text{H-NMR}$  ( $\text{CDCl}_3$ )  $\delta$  1.45-1.59 (m, 6H,  $\text{CH}_2$ ), 2.37-2.50 (m, 4H,  $\text{N-(CH}_2)_2$ ), 2.56- 2.62(m, 2H,  $\text{CH}_2\text{-N}$ ), 4.02- 4.13(m, 3H,  $\text{O-CH}_2\text{-CH}$ ), 7.13-7.81 (m, 9H, arom H);  $^{13}\text{C-NMR}$  ( $\text{CDCl}_3$ )  $\delta$  24.17, 26.05 ( $\text{CH}_2$ ), 54.67 ( $\text{N-(CH}_2)_2$ ), 60.89 ( $\text{CH}_2\text{-N}$ ), 65.18 (CH), 70.64 ( $\text{O-CH}_2$ ), 115.11, 119.28, 122.99, 128.23, 129.22 130.00, 132.39 (arom C), 137.54, 138.82 (C), 158.75 (Ar-O), 196.44 (CO); MS m/e 340.43 ( $\text{M}^+$ , 100%). Anal. Calcd for  $\text{C}_{21}\text{H}_{25}\text{NO}_3$ : C, 73.53; H, 7.46; N, 4.08. Found: C, 73.51; H, 7.71; N, 4.11.



**[4-(2-Hydroxy-3-piperidin-1-yl-propoxy)-phenyl]-phenyl-methanone (14).** 700 mg (2.75 mmol) of (4-Oxiranylmethoxy-phenyl)-phenyl-methanone **2c** was dissolved in 10 mL of piperidine and refluxed for 2.5 h. After removal of piperidine under reduced pressure a yellow oil was obtained, for further purification a column chromatography was performed (silica gel, CH<sub>2</sub>Cl<sub>2</sub>/ methanol/ concentrated NH<sub>3</sub>, 95/5/1 after removal of piperidine the percentage of methanol was increased to 80/20/1). Solvents were removed by rotary evaporation to yield light yellow oil, which crystallized from ethyl acetate. Further recrystallization from methanol /diethyl ether yielded 400 mg (42.8%) white crystals; mp 106 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 1.49-1.72 (m, 6H, CH<sub>2</sub>), 2.58-2.73 (m, 6H, CH<sub>2</sub>- N-(CH<sub>2</sub>)<sub>2</sub>), 4.05- 4.08 (m, 2H, O-CH<sub>2</sub>), 4.20- 4.29 (m, 1H, CH), 4.32 (s, 1H, OH), 6.97 (d, 2H, J = 8.85, H-3, H-5), 7.45-7.82 (m, 7H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 23.65, 25.33 (CH<sub>2</sub>), 54.78 (N-(CH<sub>2</sub>)<sub>2</sub>), 61.21 (CH<sub>2</sub>-N), 64.89 (CH), 70.64 (O-CH<sub>2</sub>), 114.06, 128.15, 129.67 (arom C<sub>r</sub>), 130.38 (Ar-O), 131.89, 132.48 ( arom C), 138.13 (Ph-CO), 162.21 (Ar-O), 195.49 (CO); MS m/e 340.3(M<sup>+</sup>, 100%). Anal. Calcd for C<sub>21</sub>H<sub>25</sub>NO<sub>3</sub>: C, 71.77; H, 7.55; N, 3.99. Found: C, 71.68; H, 7.51; N, 4.02.

**[3-(2-Hydroxy-3-morpholine-4-yl-propoxy)-phenyl]-phenyl-methanone (15).** 1 g (4.33 mmol) of (2-Oxiranylmethoxy-phenyl)-phenyl-methanone **2a** was dissolved in 20 mL of methanol and treated with 500mg (5.75 mmol) of morpholine. The mixture was refluxed for 4 h. Removal of the solvents gives yellow oil which was purified *via* Column chromatography (silica gel, CH<sub>2</sub>Cl<sub>2</sub> / methanol/ concentrated NH<sub>3</sub>, 200/10/1). Subsequent removal of the solvents under reduced pressure yielded 400 mg (27%) of light yellow oil; mp 185-193°C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 1.97-2.46 (m, 6H, CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 3.73 (t, 2H, J = 4.67, O-CH<sub>2</sub>), 3.71- 3.80 (m, 4H, O-(CH<sub>2</sub>)<sub>2</sub>), 6.95-7.10 (m, 2H, arom H), 7.41-7.58 (m, 5H, arom H), 7.75-7.79 (m, 5H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 53.57, 60.68 (CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 62.24 (CH), 66.68 (O-(CH<sub>2</sub>)<sub>2</sub>), 70.60 (O-CH<sub>3</sub>), 112.54, 121.07, 128.25 (arom C), 128.72 (Ar-CO), 129.48, 130.04, 132.31, 132.80 (arom C), 138.31 (Ph-CO), 156.45 (Ar-O), 196.41 (CO); MS m/e 341.50 (M, 20%). Anal. Calcd for C<sub>20</sub>H<sub>23</sub>NO<sub>4</sub>: C, 68.55; H, 6.90; N, 4.00. Found: C, 68.30; H, 6.62; N, 4.29.

**[4-(2-Hydroxy-3-morpholine-4-yl-propoxy)-phenyl]-phenyl-methanone C<sub>20</sub>H<sub>23</sub>NO<sub>4</sub> (16).** 700 mg (2.75 mmol) of (4-Oxiranylmethoxy-phenyl)-phenyl-methanone **2c** was dissolved in 15 mL of methanol and treated with 240 mg of

morpholine and refluxed for 7 h. Subsequent removal of the solvent under reduced pressure left 910 mg (96.8%) of yellow oil; mp 135-141 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.44-2.71 (m, 6H, CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 3.73 (t, 4H, J = 4.56, (CH<sub>2</sub>)<sub>2</sub>-O), 4.06- 4.17 (m, 3H, O-CH<sub>2</sub>), 6.98 (d, 2H, J = 8.82, H-3, H-5), 7.42-7.56 (m, 3H, arom H), 7.72-7.84 (m, 4H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 53.69, 60.83 (CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>), 65.18 (CH), 66.94 ((CH<sub>2</sub>)<sub>2</sub>-O), 70.28 (O-CH<sub>2</sub>), 114.04, 128.15, 129.68 (arom C), 130.40 (Ar-CO), 131.90, 132.49 (arom C), 138.13 (Ph-CO), 162.24 (Ar-O), 195.48 (CO); MS m/e 342.30 (M<sup>+</sup>, 95%). Anal. Calcd for C<sub>20</sub>H<sub>23</sub>NO<sub>4</sub>: C, 69.15; H, 6.87; N, 4.03. Found: C, 69.16; H, 7.07; N, 4.18.

**{3-[2-Hydroxy-3-(4-o-tolyl-piperazin-1-yl)-propoxy]-phenyl}-phenyl-methanone (17).** 700 mg (2.75 mmol) of (3-Oxiranylmethoxy-phenyl)-phenyl-methanone **2b** was dissolved in 15 mL of methanol and treated with 500 mg (2.84 mmol) of o-tolyl-piperazine and the mixture was refluxed for 6 h. Subsequent removal of the solvent yielded 1.07 g (90.29%) of yellow oil; mp 173-179 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.31(s, 3H, CH<sub>3</sub>), 2.61-2.66 (m, 4H, N(CH<sub>2</sub>)<sub>2</sub>), 2.28-2.96 (m, 6H, CH<sub>2</sub>-N, (CH<sub>2</sub>)<sub>2</sub>-N-Phe), 4.05-4.18 (m, 3H, O-CH<sub>2</sub>-CH), 6.99-7.05 (m, 2H, arom H), 7.37-7.60 (m, 6H, arom H), 7.79-7.83 (m, 2H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 17.82 (CH<sub>3</sub>), 51.71 (N-(CH<sub>2</sub>)<sub>2</sub>), 53.74 (CH<sub>2</sub>-N), 60.34 ((CH<sub>2</sub>)<sub>2</sub>-N-Ph), 65.37(CH), 70.46(O-CH<sub>2</sub>), 115.05, 118.92, 119.28, 123.07, 123.18, 126.53, 128.23, 129.23, 131.02, 132.40 (arom C), 132.40 (Ar-CO), 138.83 (Ph-CO), 151.23(Ar-N), 158.68 (Ar-O); MS m/e 430.52 (M, 20%). Anal. Calcd for C<sub>27</sub>H<sub>30</sub>N<sub>2</sub>O<sub>3</sub>: C, 73.78; H, 7.11; N, 6.37. Found. C, 73.71; H, 7.31; N, 6.24.

**{4-[2-Hydroxy-3-(4-o-tolyl-piperazin-1-yl)-propoxy]-phenyl}-phenyl-methanone (18).** 700 mg (2.75 mmol) (4-Oxiranylmethoxy-phenyl)-phenyl-methanone **2c** was dissolved in 15 mL of methanol and treated with 500 mg (2.84 mmol) of o-tolyl-piperazine, the mixture was refluxed for 5 h. Subsequent removal of the solvent under reduced pressure yield 1.1 g (92.83%) of white crystals which were recrystallized from methanol; mp 107-114 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.31(s, 3H, CH<sub>3</sub>), 2.63-2.67 (m, 4H, N(CH<sub>2</sub>)<sub>2</sub>), 2.84-2.97 (m, 6H, CH<sub>2</sub>-N, (CH<sub>2</sub>)<sub>2</sub>-N-Phe), 4.10- 4.22 (m, 3H, O-CH<sub>2</sub>-CH), 6.99-7.03 (m, 4H, arom H), 7.15-7.21 (m, 2H, arom H), 7.48-7.58 (m, 3H, arom H), 7.74- 7.86 (m, 4H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 17.82 (CH<sub>3</sub>), 51.71 (N-(CH<sub>2</sub>)<sub>2</sub>), 53.74 (CH<sub>2</sub>-N), 60.36 ((CH<sub>2</sub>)<sub>2</sub>-N-Ph), 65.29 (CH), 70.42 (O-CH<sub>2</sub>), 114.09, 118.94, 123.24,

126.57, 128.17, 129.71, (arom C), 130.38 (Ar-CO), 131.06, 131.90, 132.52 (arom C), 138.18 (Ph-CO), 151.22 (Ar-N), 162.34 (Ar-O), 195.51 (CO); MS m/e 431.4 ( $M^+$ , 100%). Anal. Calcd for  $C_{27}H_{30}N_2O_3$ : C, 75.32; H, 7.02; N, 6.51. Found: C, 75.11; H, 7.12; N, 6.39.

**4-[3-(2-Benzoyl-phenoxy)-2-hydroxy-propyl]-piperazine-1-carboxylic acid p-tolylamide (21).** 186 mg (0.54 mmol) of [2-(2-hydroxy-3-piperazine-1-yl-propoxy)-phenyl]-phenyl-methanone (**5**) was dissolved in 15 mL of dichloromethane and a solution of 75 mg (0.57 mmol) of 4-methyl phenyl isocyanate in dichloromethane was added drop wise. The mixture was stirred at room temperature and monitored by TLC. After evaporation to dryness a yellow oil was obtained which was purified by column chromatography (silica gel,  $CH_2Cl_2$  /methanol/ concentrated  $NH_3$ , 100/10/1). Subsequent removal of the solvents under reduced pressure yielded 136 mg (52.56%) of a colorless oil;  $^1H$ -NMR ( $CDCl_3$ )  $\delta$  2.41- 2.53 (m, 6H,  $CH_2$ -N-( $CH_2$ )<sub>2</sub>), 2.43 (s, 3H,  $CH_3$ ), 3.13 (s, 1H, OH), 3.51-3.53 (m, 4H, ( $CH_2$ )<sub>2</sub>-N), 3.86-3.95 (m, 1H, CH), 4.09 (d, 2H,  $J = 4.8$ , O- $CH_2$ ), 7.08-7.96 (m, 14H, arom H, NH);  $^{13}C$ -NMR ( $CDCl_3$ )  $\delta$  20.57 ( $CH_3$ ), 43.70 (( $CH_2$ )<sub>2</sub>-N-CO), 52.76 (N-( $CH_2$ )<sub>2</sub>), 59.97 ( $CH_2$ -N), 65.60 (CH), 70.60 (O- $CH_2$ ), 112.56, 120.32, 120.96, 128.20, 29.07, 129.42, 129.88, 132.27, 132.85 (arom C), 128.54 (Ar-CO), 132.30 (Ar-  $CH_3$ ), 136.38 (NH-Ar), 138.08 (Ph-CO), 155.24 (N-CO-NH), 156.39 (arom C-O), 196.54 (CO); MS m/e 473.8 (M, 18%). Anal. Calcd for  $C_{20}H_{24}N_2O_3$ : C, 71.02; H, 6.60; N, 8.87. Found: C, 70.77; H, 6.89; N, 8.64.

**4-[3-(2-Benzoyl-phenoxy)-2-hydroxy-propyl]-piperazine-1-carbothioic acid p-tolylamide (22).** 200 mg (0.58 mmol) of [2-(2-hydroxy-3-piperazine-1-yl-propoxy)-phenyl]-phenyl-methanone (**5**) was dissolved in 10 mL of dichloromethane and a solution of 91 mg (0.61 mmol) of 4-methyl phenyl isothiocyanate in dichloromethane was added drop wise. The mixture was stirred at room temperature and monitored by TLC. After evaporation to dryness, yellow oil was obtained which was purified by column chromatography (silica gel,  $CH_2Cl_2$ /methanol/ concentrated  $NH_3$ , 100/10/1). Subsequent removal of the solvents under reduced pressure yielded 181 mg (62.92%) of a light yellow oil; mp 99-101°C;  $^1H$ -NMR ( $CDCl_3$ )  $\delta$  2.12- 2.20 (m, 2H,  $CH_2$ -N), 2.33 (s, 3H,  $CH_3$ ), 3.39- 2.47 (m, 4H, N-( $CH_2$ )<sub>2</sub>), 3.74-3.79 (m, 5H, ( $CH_2$ )<sub>2</sub>-N, CH), 3.96 (d, 2H,  $J = 4.9$ , O- $CH_2$ ), 6.96-7.80 (m, 14H, arom H, NH);  $^{13}C$ -NMR ( $CDCl_3$ )  $\delta$  20.89

(CH<sub>3</sub>), 49.02 (N-(CH<sub>2</sub>)<sub>2</sub>), 52.55 ( (CH<sub>2</sub>)<sub>2</sub>-N), 59.78 (CH<sub>2</sub>-N), 65.88 (CH), 70.74 (O-CH<sub>2</sub>), 112.73, 121.15, 123.47, 125.45, 128.31, (arom C), 128.71 (Ar-CO), 129.53, 129.63, 130.12, 132.38, 132.93 (arom H), 135.11 (Ar-CH<sub>3</sub>), 137.34 (NH-Ar), 138.25 (Ph-CO), 156.51 (arom C-O), 183.26 (CS), 196.47(CO); MS m/e 489.5 (M, 15%). Anal. Calcd for C<sub>28</sub>H<sub>31</sub>N<sub>3</sub>O<sub>3</sub>S: C, 68.69; H, 6.38; N, 8.58. Found: C, 68.60; H, 6.53; N, 8.47.

**[2-(3-{Benzoyl-phenoxy}-2-hydroxyl-propyl)-piperazin-1-yl]-2-hydroxy-propoxy)-phenyl]-phenyl-methanone (23).** 356 mg (1.40 mmol) of (2-Oxiranylmethoxy-phenyl)-phenyl-methanone **2a** was dissolved in 20-30 mL of methanol, added 50.8 mg (0.59 mmol) of piperazine and the reaction mixture was refluxed for 5h. After removal of the solvent by rotary evaporation a column chromatography was performed (silica gel, CH<sub>2</sub>Cl<sub>2</sub> /methanol/concentrated NH<sub>3</sub>, 100/10/1). Subsequent evaporation to dryness gave yellow oil which crystallized from isopropanol to leave 424 mg (51%) of white solid; mp 125-135 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 1.98-2.42 (m, 12H, CH<sub>2</sub>-N-(CH<sub>2</sub>)<sub>2</sub>-N-CH<sub>2</sub>), 3.10 (s, 1H, OH), 3.70-3.76 (m, 2H, 2CH), 3.94-4.07 (m, 5H, O-CH<sub>2</sub>, OH), 6.99-7.83 (m, 18H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 53.13 (N-(CH<sub>2</sub>)<sub>4</sub>-N), 60.06 (CH<sub>2</sub>-NH), 65.47 (CH), 70.78 (O-CH<sub>2</sub>), 112.62, 121.10, 128.28, 129.53, 130.09, 132.33, 132.82 (arom C), 128.81 (Ar-CO), 138.34 (Ph-CO), 156.55 (Ar-O), 196.46 (CO); MS m/e 594.7 (M, 100%). Anal. Calcd for C<sub>36</sub>H<sub>38</sub>N<sub>2</sub>O<sub>6</sub>: C, 70.57; H, 7.14; N, 4.22. Found: C, 70.91; H, 7.02; N, 4.26.

**1-[2-(3-{4-[3-(2-Benzoyl-phenoxy)-2-hydroxyl-propyl]-piperazin-1-yl]-2-hydroxy-propoxy)-5-methyl-phenyl]-ethanone (24).** 500 mg (2.43 mmol) of 1-(5-methyl-2-oxiranylmethoxy-phenyl)-ethanone (**4**) was dissolved in 10 mL of methanol, treated with 857.1 mg (2.52 mmol) of [2-(2-hydroxy-3-piperazin-1-yl-propoxy)-phenyl]-phenyl-methanone and refluxed for 5 h. After removal of solvent under reduced pressure a column chromatography was performed (silica gel, CH<sub>2</sub>Cl<sub>2</sub>/methanol/concentrated NH<sub>3</sub>, 120/10/1). Subsequent removal of the solvents under reduced pressure yielded 1.20 g (90.67 %) of an orange oil: mp 89-93 °C; <sup>1</sup>H-NMR (CDCl<sub>3</sub>) δ 2.05-2.13 (m, 2H, CH<sub>2</sub>-N), 2.29 (s, 3H, CH<sub>3</sub>), 2.33-2.55 (m, 10H, N-(CH<sub>2</sub>)<sub>4</sub>-N-CH<sub>2</sub>), 2.62 (s, 3H, CO-CH<sub>3</sub>), 2.98-3.01 (m, 1H, OH), 3.69-4.10 (m, 6H, 2O-CH<sub>2</sub>-CH), 6.83-7.80 (m, 12H, arom H); <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 20.22 (CH<sub>3</sub>-Ar), 31.81 (CH<sub>3</sub>-CO), 53.20 (N-(CH<sub>2</sub>)<sub>2</sub>), 60.03, 60.49 (N-CH<sub>2</sub>), 65.43, 65.51 (CH), 70.75, 71.12

(O-CH<sub>2</sub>), 128.03, 128.82 (Ar-CO), 130.25 (CH<sub>3</sub>-Ar), 112.61, 112.79, 121.10, 128.27, 129.51, 130.08, 130.59, 132.32, 132.81, 134.12 (arom C), 138 (Ph-CO), 156.00, 156.55 (Ar-O), 196.46(CO), 199.88 (Ar-CO-CH<sub>3</sub>); MS m/e 594.7 (M<sup>+</sup>, 100%). Anal. Calcd for C<sub>32</sub>H<sub>38</sub>N<sub>2</sub>O<sub>6</sub>: C, 61.82; H, 6.65; N, 4.51. Found: C, 61.60; H, 6.91; N, 4.37.

### Computational Studies

**Ligand Efficiency (LE)**, ligand efficiency ( $\Delta g$ ) values of the data were calculated by normalizing binding free energy of a ligand for number of heavy atoms. Free energy calculation was carried out as described by Hopkins *et al*, (Equ. i). According to Hopkins *et al*, IC<sub>50</sub> from percentage inhibition can be substituted for *K*<sub>d</sub> (dissociation constant potency)<sup>50</sup> which was further confirmed by experimental results of Kuntz and co-workers.<sup>40</sup> Ligand efficiency calculations was done for a temperature of 310 K and given in kcal per heavy atom (Equ. ii).

$$\Delta G = -RT \ln K_d \quad (\text{Equ. i})$$

$$\Delta g = -\Delta G / \text{HA}_{(\text{non-hydrogen atom})} \quad (\text{Equ. ii})$$

A size independent fit quality score was obtained as described by Reynolds *et al*,<sup>20</sup> by fitting the maximum LE over a large range of molecular size. All calculations regarding ligand efficiency were done by using Excel worksheet. Activity values of the propafenone type inhibitors (GPV576, GPV005, GPV0062 and propafenone) were determined experimentally by a daunorubicin efflux assay.<sup>51,34</sup> Inhibition of rhodamine 123 efflux in the transfectant mouse lymphoma line L5178 VMDR1 C.06 were used to characterize the MDR- modulating activity values of verapamil, nifedipine, and cyclosporine A. IC<sub>50</sub> values of tariquidar,<sup>52</sup> elacridar,<sup>53,54</sup> valspodar,<sup>55</sup> zosuquidar<sup>56,57</sup> and ONT-093<sup>58</sup> were taken from literature (Table 2). IC<sub>50</sub> values for most of the compounds in clinical studies were reported by using rhodamine efflux assays. We use these values as there is a direct correlation between the IC<sub>50</sub> values from daunorubicin and rhodamine efflux assays<sup>37</sup>

**Lipophilic Efficiency (LipE)**, of benzophenones were calculated (Equ. iii), and compared with the compounds which reached clinical studies (verapamil, tariquidar, valspodar, elacridar, zosuquidar, ONT-093, nifedipine and cyclosporine A), as well as with selected propafenone analogs.

$$\text{LipE} = \text{LLE} = \text{pEC}_{50} - \text{clogP} \quad (\text{Equ. iii})$$

clogP values of the data set were computed by using the Bio-loom software package<sup>39</sup> and the LipE calculations were performed by using excel worksheet. In order to compare the standard threshold of LipE along three different entry pathways of ligands into respective binding pockets of P-gp, hERG and SERT, a data set from literature was used. It includes 744 SERT inhibitors extracted from the ChEMBL data base,<sup>59</sup> 313 hERG blockers<sup>43</sup> from literature, and 372 inhibitors of P-gp mediated daunorubicin efflux (in-house data). The data sets are available at our homepage ([pharminfo.univie.ac.at](http://pharminfo.univie.ac.at)) and from Chemspider ([www.chemspider.com](http://www.chemspider.com)).

**Docking,** Compounds **6**, **19**, **20** and **23** were docked in their neutral form into an open state homology model of human P-gp<sup>17</sup> based on the X-ray structure of mouse P-gp (PDB ID: 3G5U)<sup>18</sup> by using the software package GOLD. In order to avoid any bias we considered the whole transmembrane domain region as binding pocket. 100 poses per ligand were obtained and finally ligand protein complexes were minimized by LigX, a minimization tool implemented in MOE, by using the MMFF94 force field.

Agglomerative Hierarchical Cluster analysis of the consensus RMSD matrix based on the common scaffold of the ligands identified 2 interesting clusters of poses containing all four ligands. However, additional 5 clusters have been identified containing three out of four ligands. All seven clusters were occupying the center of the binding cavity mainly interacting with amino acid residues of TM1, 5, 6, 7, 8, 10 and 11 (SM Figure 2A). For a more detailed analysis of the ligand–protein interaction profiles of selected ligands, we used the two clusters containing all four ligands (SM Figure 2B).

In order to prioritize among the two clusters, a rescoring of all docking poses by using four different scoring functions in MOE, (ASE, affinity dG, Alpha HB, London dG) was performed. Subsequently, for each ligand, the top 10 ranked poses according to consensus scoring were taken and analyzed. Out of these 40 poses, 7 poses were present in cluster 1 while only one showed up in cluster 2. In addition, taking only the top ranked pose per ligand, two (**6**, **23**) out of four ligands were located in cluster 1 (SM Figure 2B). Therefore interaction position of cluster 1 was supposed to be the most likely one for benzophenones.

**Biological Assay****Cell lines**

The resistant CCRF vcr1000 cell line was maintained in RPMI 1640 medium containing 10% fetal calf serum (FCS) and 1000ng/ml vincristine. The selecting agent was washed out 1 week before the experiments. This cell line was selected due to its distinct P-gp expression.

**Inhibition of daunorubicin efflux**

IC<sub>50</sub> values for daunorubicin efflux inhibition were determined as reported<sup>31</sup>. Briefly, cells were sedimented, the supernatant was removed by aspiration, and the cells were resuspended at a density of  $1 \times 10^6$ /mL in RPMI 1640 medium containing daunorubicin (Sigma Chemical Co., St. Louis, MO) at a final concentration of 3  $\mu$ mol/l. Cell suspensions were incubated at 37°C for 30 min. Tubes were chilled on ice and centrifuged at 500 g in an Eppendorf 5403 centrifuge (Eppendorf, Hamburg, Germany). Supernatants were removed, and the cell pellet was resuspended in medium pre-warmed to 37°C containing either no inhibitor or compounds at various concentrations ranging from 20  $\mu$ M to 200  $\mu$ M, depending on the solubility and expected potency of the inhibitor. Eight concentrations (serial 1:3 dilution) were tested for each inhibitor. After 60, 120, 180 and 240 seconds, aliquots of the incubation mixture were transferred to tubes containing an equal volume of ice-cold stop solution (RPMI medium containing GPV31 at a final concentration of 5 $\mu$ mol/L). Zero time points were determined by immediately pipetting daunorubicin-preloaded cells into ice cold stop solution. Samples drawn at the respective time points were kept in an ice water bath and measured within 1h on a Becton Dickinson FACS Calibur flow cytometer (Becton Dickinson, Vienna, Austria). Viable cells were selected by setting appropriate gates for forward and side scatter. The excitation and emission wavelengths were 482 nm and 558 nm, respectively. Five thousand gated events were accumulated for the determination of mean fluorescence values.

**Acknowledgement:** We are grateful to the Austrian Science Fund for financial support (grant SFB F35). Ishrat Jabeen thanks the Higher Education Commission of Pakistan for financial support.

### References

1. Giacomini, K. M.; Huang, S. M.; Tweedie, D. J.; Benet, L. Z.; Brouwer, K. L.; Chu, X.; Dahlin, A.; Evers, R.; Fischer, V.; Hillgren, K. M.; Hoffmaster, K. A.; Ishikawa, T.; Keppler, D.; Kim, R. B.; Lee, C. A.; Niemi, M.; Polli, J. W.; Sugiyama, Y.; Swaan, P. W.; Ware, J. A.; Wright, S. H.; Yee, S. W.; Zamek-Gliszczynski, M. J.; Zhang, L. Membrane transporters in drug development. *Nat Rev Drug Discov* **2010**, *9*, 215-36.
2. Lee, E. J.; Lean, C. B.; Limenta, L. M. Role of membrane transporters in the safety profile of drugs. *Expert Opin Drug Metab Toxicol* **2009**, *5*, 1369-83.
3. Szakacs, G.; Paterson, J. K.; Ludwig, J. A.; Booth-Genthe, C.; Gottesman, M. M. Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* **2006**, *5*, 219-34.
4. Couture, L.; Nash, J. A.; Turgeon, J. The ATP-binding cassette transporters and their implication in drug disposition: a special look at the heart. *Pharmacol Rev* **2006**, *58*, 244-58.
5. Szakacs, G.; Varadi, A.; Ozvegy-Laczka, C.; Sarkadi, B. The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). *Drug Discov Today* **2008**, *13*, 379-93.
6. Colabufo, N. A.; Berardi, F.; Contino, M.; Niso, M.; Perrone, R. ABC pumps and their role in active drug transport. *Curr Top Med Chem* **2009**, *9*, 119-29.
7. Gottesman, M. M.; Ling, V. The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research. *FEBS Lett* **2006**, *580*, 998-1009.
8. Glavinas, H.; Krajcsi, P.; Cserepes, J.; Sarkadi, B. The role of ABC transporters in drug resistance, metabolism and toxicity. *Curr Drug Deliv* **2004**, *1*, 27-42.
9. Juliano, R.; Ling, V.; Graves, J. Drug-resistant mutants of chinese hamster ovary cells possess an altered cell surface carbohydrate component. *Journal of Supramolecular Structure* **1976**, *4*, 521-526.
10. Ford, R. C.; Kamis, A. B.; Kerr, I. D.; Callaghan, R. The ABC Transporters: Structural Insights into Drug Transport. In *Transporters as Drug Carriers*, Wiley-VCH Verlag GmbH & Co. KGaA: 2010; pp 1-48.
11. Broccatelli, F.; Carosati, E.; Neri, A.; Frosini, M.; Goracci, L.; Oprea, T. I.; Cruciani, G. A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields. *J Med Chem* **2011**.



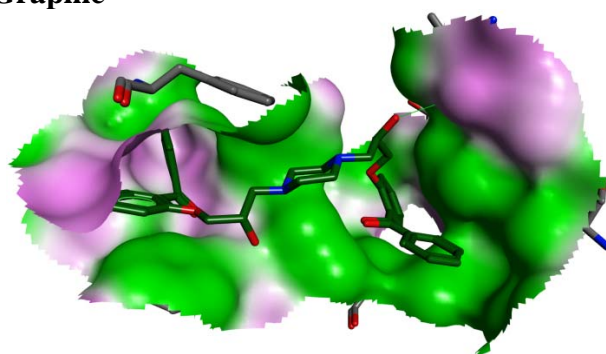
12. Kemper, E. M.; van Zandbergen, A. E.; Cleypool, C.; Mos, H. A.; Boogerd, W.; Beijnen, J. H.; van Tellingen, O. Increased penetration of paclitaxel into the brain by inhibition of P-Glycoprotein. *Clin Cancer Res* **2003**, *9*, 2849-55.
13. Kuhnle, M.; Egger, M.; Muller, C.; Mahringer, A.; Bernhardt, G.; Fricker, G.; Konig, B.; Buschauer, A. Potent and selective inhibitors of breast cancer resistance protein (ABCG2) derived from the p-glycoprotein (ABCB1) modulator tariquidar. *J Med Chem* **2009**, *52*, 1190-7.
14. Ford, J. M. Experimental reversal of P-glycoprotein-mediated multidrug resistance by pharmacological chemosensitisers. *Eur J Cancer* **1996**, *32A*, 991-1001.
15. Thomas, H.; Coley, H. M. Overcoming multidrug resistance in cancer: an update on the clinical strategy of inhibiting p-glycoprotein. *Cancer Control* **2003**, *10*, 159-65.
16. Kannan, P.; John, C.; Zoghbi, S. S.; Halldin, C.; Gottesman, M. M.; Innis, R. B.; Hall, M. D. Imaging the function of P-glycoprotein with radiotracers: pharmacokinetics and in vivo applications. *Clin Pharmacol Ther* **2009**, *86*, 368-77.
17. Klepsch, F.; Chiba, P.; Ecker, G. F. Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Comput Biol* **2011**, *7*, e1002036.
18. Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L.; Chang, G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **2009**, *323*, 1718-22.
19. Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional group contributions to drug-receptor interactions. *J Med Chem* **1984**, *27*, 1648-57.
20. Reynolds, C. H.; Bembenek, S. D.; Tounge, B. A. The role of molecular size in ligand efficiency. *Bioorg Med Chem Lett* **2007**, *17*, 4258-61.
21. Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand binding efficiency: trends, physical basis, and implications. *J Med Chem* **2008**, *51*, 2432-8.
22. Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **2007**, *6*, 881-90.
23. Ryckmans, T.; Edwards, M. P.; Horne, V. A.; Correia, A. M.; Owen, D. R.; Thompson, L. R.; Tran, I.; Tutt, M. F.; Young, T. Rapid assessment of a novel series of selective CB(2) agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis. *Bioorg Med Chem Lett* **2009**, *19*, 4406-9.
24. Keseru, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat Rev Drug Discov* **2009**, *8*, 203-12.
25. Mortenson, P. N.; Murray, C. W. Assessing the lipophilicity of fragments and early hits. *J Comput Aided Mol Des* **2011**.

26. Chiba, P.; Burghofer, S.; Richter, E.; Tell, B.; Moser, A.; Ecker, G. Synthesis, pharmacologic activity, and structure-activity relationships of a series of propafenone-related modulators of multidrug resistance. *J Med Chem* **1995**, *38*, 2789-93.
27. Kaiser, D.; Smiesko, M.; Kopp, S.; Chiba, P.; Ecker, G. F. Interaction field based and hologram based QSAR analysis of propafenone-type modulators of multidrug resistance. *Med Chem* **2005**, *1*, 431-44.
28. Ecker, G. F.; Csaszar, E.; Kopp, S.; Plagens, B.; Holzer, W.; Ernst, W.; Chiba, P. Identification of ligand-binding regions of P-glycoprotein by activated-pharmacophore photoaffinity labeling and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Mol Pharmacol* **2002**, *61*, 637-48.
29. Parveen, Z.; Stockner, T.; Bentele, C.; Pferschy, S.; Kraupp, M.; Freissmuth, M.; Ecker, G. F.; Chiba, P. Molecular Dissection of Dual Pseudosymmetric Solute Translocation Pathways in Human P-Glycoprotein. *Mol Pharmacol* **2011**.
30. Pitha, J.; Szabo, L.; Szurmai, Z.; Buchowiecki, W.; Kusiak, J. W. Alkylating prazosin analogue: irreversible label for alpha 1-adrenoceptors. *J Med Chem* **1989**, *32*, 96-100.
31. Chiba, P.; Ecker, G.; Schmid, D.; Drach, J.; Tell, B.; Goldenberg, S.; Gekeler, V. Structural requirements for activity of propafenone-type modulators in P-glycoprotein-mediated multidrug resistance. *Mol Pharmacol* **1996**, *49*, 1122-30.
32. Tmej, C.; Chiba, P.; Huber, M.; Richter, E.; Hitzler, M.; Schaper, K. J.; Ecker, G. A combined Hansch/Free-Wilson approach as predictive tool in QSAR studies on propafenone-type modulators of multidrug resistance. *Arch Pharm (Weinheim)* **1998**, *331*, 233-40.
33. König, G.; Chiba, P.; Ecker, G. F. Hydrophobic moments as physicochemical descriptors in structure-activity relationship studies of P-glycoprotein inhibitors. *Monatshefte für Chemie / Chemical Monthly* **2008**, *139*, 401-405.
34. Chiba, P.; Hitzler, M.; Richter, E.; Huber, M.; Tmej, C.; Giovagnoni, E.; Ecker, G. Studies on Propafenone-type Modulators of Multidrug Resistance III: Variations on the Nitrogen. *Quant. Stmet.-Act. Relat* **1997**, *16*, 361-6.
35. Ecker, G.; Huber, M.; Schmid, D.; Chiba, P. The importance of a nitrogen atom in modulators of multidrug resistance. *Mol Pharmacol* **1999**, *56*, 791-6.
36. Hiessbock, R.; Wolf, C.; Richter, E.; Hitzler, M.; Chiba, P.; Kratzel, M.; Ecker, G. Synthesis and in vitro multidrug resistance modulating activity of a series of dihydrobenzopyrans and tetrahydroquinolines. *J Med Chem* **1999**, *42*, 1921-6.
37. Chiba, P.; Holzer, W.; Landau, M.; Bechmann, G.; Lorenz, K.; Plagens, B.; Hitzler, M.; Richter, E.; Ecker, G. Substituted 4-acylpyrazoles and 4-acylpyrazolones: synthesis and multidrug resistance-modulating activity. *J Med Chem* **1998**, *41*, 4001-11.

38. Seelig, A.; Landwojtowicz, E. Structure-activity relationship of P-glycoprotein substrates and modifiers. *Eur J Pharm Sci* **2000**, *12*, 31-40.
39. *Bio-Loom program, trial version, by BioByte Co.*
40. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. *Proc Natl Acad Sci U S A* **1999**, *96*, 9997-10002.
41. Verdonk, M. L.; Rees, D. C. Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem* **2008**, *3*, 1179-80.
42. <https://www.ebi.ac.uk/chembl/db/>.
43. Thai, K. M.; Ecker, G. F. A binary QSAR model for classification of hERG potassium channel blockers. *Bioorg Med Chem* **2008**, *16*, 4107-19.
44. Wellfelt, K.; Skold, A. C.; Wallin, A.; Danielsson, B. R. Teratogenicity of the class III antiarrhythmic drug almokalant. Role of hypoxia and reactive oxygen species. *Reprod Toxicol* **1999**, *13*, 93-101.
45. Houltz, B.; Darpo, B.; Swedberg, K.; Blomstrom, P.; Brachmann, J.; Crijns, H. J.; Jensen, S. M.; Svernhage, E.; Vallin, H.; Edvardsson, N. Effects of the Ikr-blocker almokalant and predictors of conversion of chronic atrial tachyarrhythmias to sinus rhythm. A prospective study. *Cardiovasc Drugs Ther* **1999**, *13*, 329-38.
46. Jabeen, I.; Wetwitayaklung, P.; Klepsch, F.; Parveen, Z.; Chiba, P.; Ecker, G. F. Probing the stereoselectivity of P-glycoprotein-synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4-b][1,4]oxazines. *Chem Commun (Camb)* **2011**, *47*, 2586-8.
47. Pajeva, I.; Wiese, M. Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: a comparative molecular field analysis study. *J Med Chem* **1998**, *41*, 1815-26.
48. Ecker, G. F.; Pleban, K.; Kopp, S.; Csaszar, E.; Poelarends, G. J.; Putman, M.; Kaiser, D.; Konings, W. N.; Chiba, P. A three-dimensional model for the substrate binding domain of the multidrug ATP binding cassette transporter LmrA. *Mol Pharmacol* **2004**, *66*, 1169-79.
49. Pleban, K.; Kopp, S.; Csaszar, E.; Peer, M.; Hrebicek, T.; Rizzi, A.; Ecker, G. F.; Chiba, P. P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. *Mol Pharmacol* **2005**, *67*, 365-74.
50. Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430-431.

51. Pleban, K.; Hoffer, C.; Kopp, S.; Peer, M.; Chiba, P.; Ecker, G. F. Intramolecular distribution of hydrophobicity influences pharmacological activity of propafenone-type MDR modulators. *Arch Pharm (Weinheim)* **2004**, 337, 328-34.
52. Roe, M.; Folkes, A.; Ashworth, P.; Brumwell, J.; Chima, L.; Hunjan, S.; Pretswell, I.; Dangerfield, W.; Ryder, H.; Charlton, P. Reversal of P-glycoprotein mediated multidrug resistance by novel anthranilamide derivatives. *Bioorg Med Chem Lett* **1999**, 9, 595-600.
53. Dodic, N.; Dumaitre, B.; Daugan, A.; Pianetti, P. Synthesis and activity against multidrug resistance in Chinese hamster ovary cells of new acridone-4-carboxamides. *J Med Chem* **1995**, 38, 2418-26.
54. Bachmeier, C. J.; Miller, D. W. A fluorometric screening assay for drug efflux transporter activity in the blood-brain barrier. *Pharm Res* **2005**, 22, 113-21.
55. Wang, J. S.; Zhu, H. J.; Markowitz, J. S.; Donovan, J. L.; DeVane, C. L. Evaluation of antipsychotic drugs as inhibitors of multidrug resistance transporter P-glycoprotein. *Psychopharmacology (Berl)* **2006**, 187, 415-23.
56. Shepard, R. L.; Cao, J.; Starling, J. J.; Dantzig, A. H. Modulation of P-glycoprotein but not MRP1- or BCRP-mediated drug resistance by LY335979. *Int J Cancer* **2003**, 103, 121-5.
57. Dantzig, A. H.; Shepard, R. L.; Cao, J.; Law, K. L.; Ehlhardt, W. J.; Baughman, T. M.; Bumol, T. F.; Starling, J. J. Reversal of P-glycoprotein-mediated multidrug resistance by a potent cyclopropyldibenzosuberane modulator, LY335979. *Cancer Res* **1996**, 56, 4171-9.
58. Newman, M. J.; Rodarte, J. C.; Benbatoul, K. D.; Romano, S. J.; Zhang, C.; Krane, S.; Moran, E. J.; Uyeda, R. T.; Dixon, R.; Guns, E. S.; Mayer, L. D. Discovery and characterization of OC144-093, a novel inhibitor of P-glycoprotein-mediated multidrug resistance. *Cancer Res* **2000**, 60, 2964-72.
59. Wandel, C.; Kim, R. B.; Kajiji, S.; Guengerich, P.; Wilkinson, G. R.; Wood, A. J. P-glycoprotein and cytochrome P-450 3A inhibition: dissociation of inhibitory potencies. *Cancer Res* **1999**, 59, 3944-8.

### Table of Contents Graphic



**Contents**

**Introduction**

**Ligand Based Approaches**

**Structure Based Approaches**

**Importance of ABC-Transporter for ADME**

**Predicting Substrates for ABCB1**

**Outlook**

**Acknowledgement**

**References**

❶ **Information:** This chapter was published in *Current Pharmaceutical Design* (2010), Volume 16, 1742-1752 by Klepsch Freya, **Jabeen Ishrat** and Ecker Gerhard F.

# Pharmacoinformatic Approaches to Design Natural Product Type Ligands of ABC-Transporters

F. Klepsch<sup>1</sup>, I. Jabeen<sup>1</sup>, P. Chiba<sup>2</sup> and G. F. Ecker<sup>1,\*</sup>

<sup>1</sup>University of Vienna, Department of Medicinal Chemistry, Althanstrasse 14, 1090 Vienna, Austria, <sup>2</sup>Medical University of Vienna, Institute of Medical Chemistry, Währinger Straße 10, 1090 Vienna, Austria

**Abstract:** ABC-transporters have been recognized as being responsible for multiple drug resistance in tumor therapy, for decreased brain uptake and low oral bioavailability of drug candidates, and for drug-drug interactions and drug induced cholestasis. P-glycoprotein (ABCB1), the paradigm protein in the field, is mainly effluxing natural product toxins and shows very broad substrate specificity. Within this article we will highlight SAR and QSAR approaches for designing natural product type inhibitors of ABCB1 and related proteins as well as *in silico* strategies to predict ABCB1 substrates and inhibitors in order to design out undesirable drug/protein interaction.

**Keywords:** Natural products, ABC transporter, P-glycoprotein, *in silico* methods.

## INTRODUCTION

More than 30 years ago P-glycoprotein (P-gp, ABCB1), the paradigm ABC-transporter, has been discovered as being responsible for decreased accumulation of natural product toxins in tumor cells. [1] It soon became evident that P-gp has a remarkably broad substrate pattern transporting numerous structurally and functionally diverse natural products across cell membranes. The multispecific nature of this drug efflux transporter and its potential role in clinical drug resistance raised high expectations and initiated development of inhibitors that would re-establish sensitivity to standard therapeutic regimens [2]. However, since the identification of the P-gp inhibitory potential of verapamil [3] almost 3 decades have passed and still no P-gp inhibitor entered the market. Furthermore, since the discovery of P-gp in 1976 [4], additional 47 human ABC-transporters have been identified of which several have been related to either human disease or drug resistance [5]. Within the past decade considerable progress has been made in unravelling the physiological function of P-gp and other ABC-transporters. Results clearly demonstrated the multiple involvement of several members of the ABC-transporter family in drug-uptake, -disposition and -elimination [6] rendering them antitargets rather than classical targets suited for drug therapy. Within this article we will highlight ligand- and structure-based approaches targeting P-gp and some of its homologues by natural products and related compounds. In addition, we will also summarise recent attempts for predicting P-gp substrates, a topic which is becoming more and more important in the ABC-transporter field.

## LIGAND BASED APPROACHES

P-glycoprotein and its congeners are membrane-spanning proteins and thus until very recently only little structural information was available. Therefore, in lead optimization programs, mainly ligand-based approaches have been pursued. These include QSAR studies on structurally homologous series of compounds, such as verapamil analogues, triazines, acridonecarboxamides, phenothiazines, thioxanthenes, flavones, dihydropyridines, propafenones and cyclosporine derivatives [7, 8]. These studies pinpoint the importance of H-bond acceptors and their strength, of the distance between aromatic moieties and H-bond acceptors as well as the influence of global physicochemical parameters, such as lipophilicity and molar refractivity. In the quest for designing more

potent inhibitors of ABC-transporter with high selectivity, also natural products served as basic scaffolds for lead optimization programs. In the following section we will highlight selected studies dealing with flavonoids, steroids and sesquiterpenes.

## Flavonoids

Flavonoids represent a major class of natural compounds widely present in foods and herbal products (Fig. (1)). They have been shown to block both the breast cancer resistance protein (BCRP, ABCG2) [9, 10] and P-glycoprotein (P-gp) [11]. In order to develop more potent inhibitors of ABCG2, a set of flavonoids covering five flavonoid subclasses (flavones, isoflavones, chalcones, flavonols and flavanones) (Fig. (2)), were selected for quantitative structure activity (QSAR) relationship studies [9].

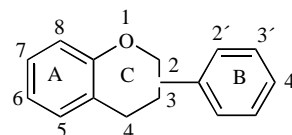
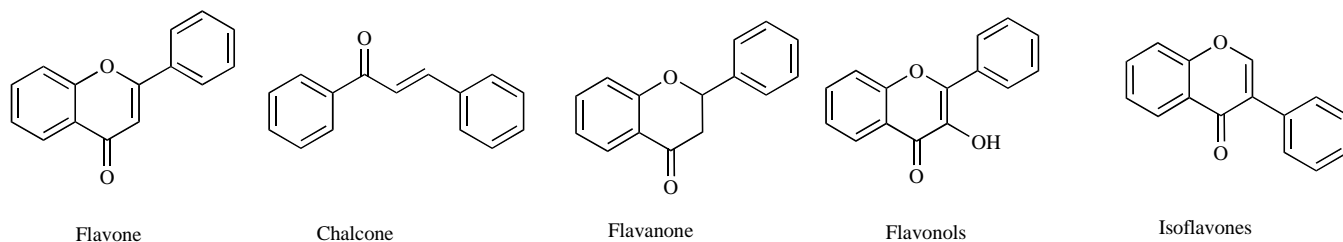


Fig. (1). Basic Structure of flavonoids (taken from [9]).

Systematic structure activity relationship studies showed that the presence of a 2, 3-double bond in ring C, ring B attached at position 2, hydroxylation at position 5, lack of hydroxylation at position 3 and hydrophobic substituents at positions 6, 7, 8 or 4', are the structural requirements for potent flavonoid-type BCRP inhibitors. Remarkably, although both ABCB1 and ABCG2 are polyspecific in ligand recognition, flavonoids show a different SAR pattern for the two transporters. A notable difference is that 3-hydroxylation was shown to increase flavonoid-P-gp interaction, whereas O-methylation of this hydroxyl group markedly decreased the interaction. Furthermore, hydroxylation at position 7 did not alter flavonoid-Pgp interaction [12], but moderately increased the flavonoid-BCRP interaction. Also in the series of propafenone-type inhibitors, subtle differences in ABCB1 and ABCG2 inhibitory activity could be observed within the same chemical scaffold [13]. In a study on tariquidar analogs, Wiese and co-workers performed Free- Wilson [14] analyses to identify the structural elements which significantly influence the inhibitory effect on ABCB1 and ABCG2 [15]. It was shown that methoxy groups in positions 6 and 7 of the tetrahydroisoquinolinylamide substructure contribute statistically significant to ABCB1 inhibition. In contrast, the elimination of methoxy groups in positions 6 and 7 of the tetrahydroisoquinoline substructure strengthened the interaction with ABCG2. Moreover, it

\*Address correspondence to this author at the Department of Medicinal Chemistry, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria; Tel: +43-1-4277-55110; Fax: +43-1-4277-9551; E-mail: Gerhard.f.ecker@univie.ac.at



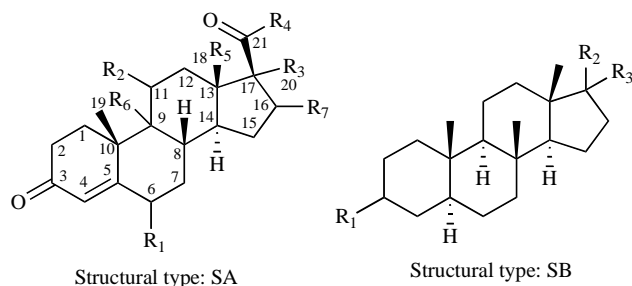
**Fig. (2).** Basic structures of five flavonoid subclasses (flavones, isoflavones, chalcones, flavanols and flavanones) used for QSAR study (taken from [9])

was demonstrated that the introduction of an electrophilic substituent, such as a nitro group, increases ABCG2 inhibitory potency relative to that for ABCB1.

However, in contrast to propafenones, flavonoids are supposed to interact with the nucleotide binding domain of the transporter. Thus, these differences in the SAR pattern may reflect the distinct structural requirements for binding to the NBDs of ABCG2 and ABCB1. Based on the QSAR model derived, logP makes a positive contribution to the ABCG2 inhibition activity. These findings were considered useful for developing potent flavonoid type inhibitors of ABCG2 (e.g. 7, 8-benzoflavone) with potential clinical applicability [9].

### Steroids

Steroids have been shown in numerous experiments to exhibit typical properties of MDR-reversing agents [16, 17]. Steroids are perfectly suited for 3D-QSAR studies such as CoMFA and CoMSIA, as they are rather rigid and small differences in structure give rise to considerable changes in biological activity. [18] Remarkably, in the class of steroids CoMSIA models were built for distinguishing which characteristic features are important for a steroid to be a substrate or an inhibitor of ABCB1. [19] Twenty steroids were selected from the literature [20] and divided into two groups: the substrate group contained 13 compounds, while the inhibitor group comprised all 20 compounds (Table 1). The overall chemical structures are shown in Fig. (3).



**Fig. (3).** Template structures of two different types of steroidal Compounds (taken from [19]).

The authors conclude that the requirement for strong hydrophobicity is more essential for inhibitors than for substrates. Another major difference is that for steroid substrates bulky substitutions surrounding C-6 are not well tolerated, whereas electronegative charged groups in position C-11  $\beta$  are favorable. Moreover, for steroid inhibitors bulky groups around C-3 decrease the activity, while there is no specific requirement at C-3 for steroid substrates. Any substituents around C-17 $\alpha$  and C-21  $\beta$  favor inhibitory potency, but disfavor or have little impact on substrates properties (Fig. (4) and (5)).

**Table 1.** Steroidal Data Set Used in 3D-QSAR Analysis

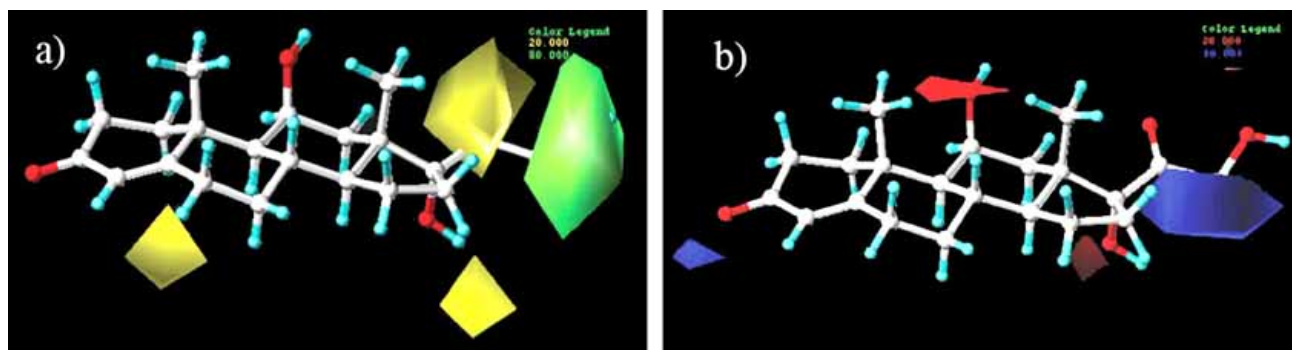
| No | Steroid Compound                              | Structural Type | Substrate (S)/Inhibitor (I) |
|----|---|-----------------|-----------------------------|
| 1  | Cortisol                                      | SA              | S + I                       |
| 2  | 17 $\alpha$ -Hydroxyprogesterone              | SA              | S + I                       |
| 3  | Progesterone                                  | SA              | S + I                       |
| 4  | Corticosterone                                | SA              | S + I                       |
| 5  | 11-Deoxycortisol                              | SA              | S + I                       |
| 6  | Medroxyprogesterone Acetate                   | SA              | S + I                       |
| 7  | Aldosterone                                   | SA              | S + I                       |
| 8  | Dexamethasone $\ddagger$                      | SA              | S + I                       |
| 9  | Dehydroepiandrosterone $\dagger$              | SB              | S + I                       |
| 10 | Pregnenolone $\dagger$                        | SB              | S + I                       |
| 11 | Testosterone $\ddagger$                       | SB              | S + I                       |
| 12 | Androstenedione $\ddagger$                    | SB              | S + I                       |
| 13 | Dihydrotestosterone                           | SB              | S + I                       |
| 14 | Deoxycorticosterone                           | SA              | I Only                      |
| 15 | Medroxyprogesterone                           | SA              | I Only                      |
| 16 | 16 $\alpha$ -Methylprogesterone               | SA              | I Only                      |
| 17 | 17 $\alpha$ -Hydroxypregnenolone $\dagger$    | SB              | I Only                      |
| 18 | Androsterone                                  | SB              | I Only                      |
| 19 | Pregnanedione                                 | SB              | I Only                      |
| 20 | 6, 16- $\alpha$ -Methylpregnenolone $\dagger$ | SB              | I Only                      |

$\ddagger$ 1, 2-Double bond,  $\dagger$ 5, 6-Double bond,  $\ddagger$ 4, 5-Double bond

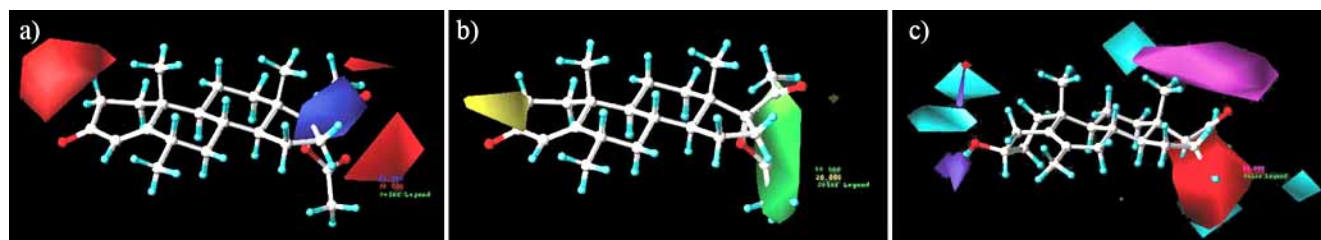
### Sesquiterpenes

Sesquiterpenes have been isolated from the extracts of the *Celastraceae* family and have been used for centuries in traditional medicine. Furthermore, they have shown clinical potential as anti-cancer drugs [21]. In a comprehensive study, 76 Dihydro- $\beta$ -agaro-furan derivatives were used to inhibit P-gp-mediated daunorubicin (DNR) efflux from intact cells [22] (Fig. (6)).

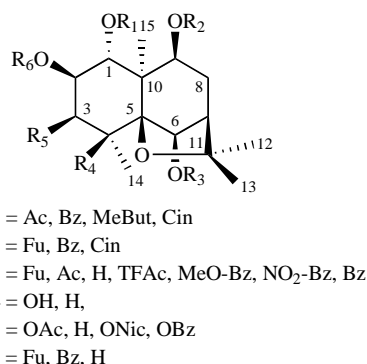
Structure-activity relationship studies [22] of compounds varied at the A-ring of sesquiterpenes suggest that an ester group at position C-2 seems essential for the inhibition of ABCB1.



**Fig. (4).** (a) shows steric contour maps of steroid substrates, the green contours suggest that the larger substituent around the C-21 $\alpha$  position is sterically favorable while substitutions at C-6, C-17 $\alpha$  and C-21 $\beta$  positions are sterically unfavorable. (b) shows the electrostatic contour maps of steroid substrates, red and blue contours describe the electrostatic regions, which are favorable and unfavorable to a negative charge, respectively. A negatively charged substituent at C-11 $\beta$  and electrostatic groups around C-3, C-17 $\alpha$  and C-21 are favorable for interaction between steroid substrates and Pgp (taken from [19]).



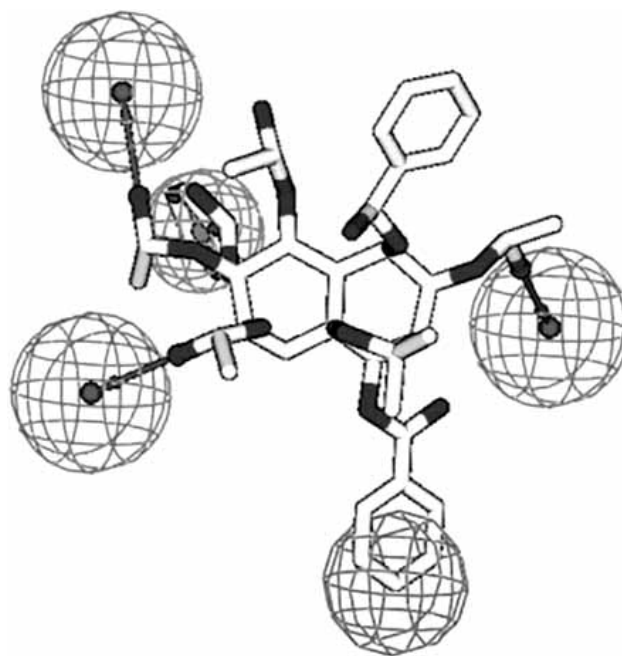
**Fig. (5).** (a) shows electrostatic contour maps of steroid inhibitors, negative charge favored red regions were found near C-3, C-17 and C-21 positions while positive charge favored or negative charge unfavored blue region is found around C-16 $\beta$  position. (b) shows steric contour plots of steroid inhibitors. Bulky groups in the vicinity of C3 are not tolerated, whereas a bulky substituent like -C(O)CH<sub>3</sub> around C-21 may greatly enhance the binding affinity to P-gp. (c) Representation of H-bond donor and acceptor contour maps of steroid inhibitors. The cyan and purple contours indicate regions, where an H-bond donor group increases or decreases activity, respectively. The magenta and red contours indicate regions, in which an H-bond acceptor group increases or decreases activity, respectively. Small purple contours around C-3 suggest that a hydrogen-bond acceptor such as a carbonyl group may increase the inhibitory effect. Large cyan contours around the first hexagonal ring (constituted by C-1–C-6 with the exception of C-3), and the C-21 $\alpha$  positions reveal that hydrogen-bond donors such as a methyl or hydroxyl group may enhance the inhibitory potency. Red and magenta contours around C-17 and C-21 indicate that these regions are very sensitive to hydrogen-bond donor or acceptor strength with respect to interaction with P-gp (taken from [19]).



**Fig. (6).** Common Scaffold of sesquiterpenes assayed for the inhibition of the human P-gp.

Sesquiterpenes with the OAc substituent at position C-3 were found to be more potent than the compounds with a hydroxyl or hydrogen group at the same position. It seems that the presence of an H-bond acceptor at C-3 is important for activity.

CoMSIA and CoMFA studies demonstrated that the carbonyl groups at the C-2, C-3, and C-8 position, act as acceptors for H-bond donors in the binding site (Fig. (7)). In addition, the models also point towards the importance of a bulky hydrophobic substituent at the C-2 $\beta$  position (depicted as a green sphere) and a



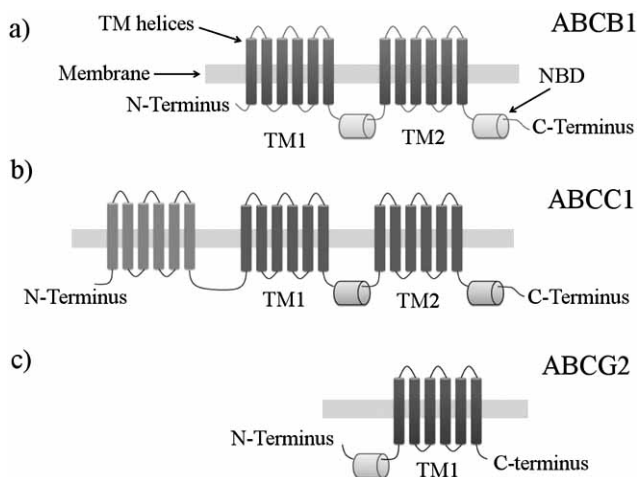
**Fig. (7).** Summary of the most prominent structural elements of ligands that are important for high P-gp activity obtained by 3D-QSAR/CoMFA (taken from [22]).



hydrophobic substituent at the C-6 position (depicted as a blue sphere). In general, the important features rendering sesquiterpenes highly active are the overall esterification level of the compounds, the presence of at least two aromatic-ester moieties (such as a benzoate-nicotinate or benzoate-benzoate), and the size of the molecule. Tetra- or penta-substituted sesquiterpenes show the highest potency, whereas additional ester moieties in the molecule lead to inactive compounds.

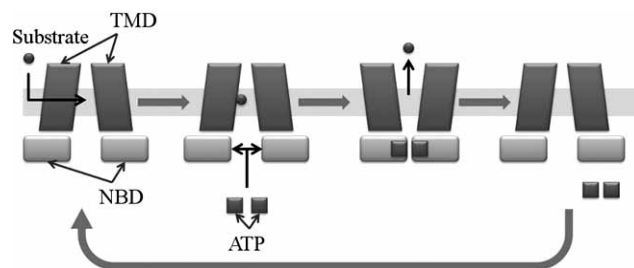
### STRUCTURE-BASED STUDIES

The general architecture of ABC transporters are more or less the same throughout this superfamily (Fig. (8)). Two transmembrane (TM) and two nucleotide binding (NB) domains are necessary to yield a functional efflux pump which can export its substrates. Since the NB domains harbor the hallmark ABC motifs they are highly conserved among all ABC transporters. Much less sequence identity can be found in the two transmembrane domains (TMD) which are generally responsible for drug binding and therefore the reason for diverse substrate/inhibitor profiles of representatives of this protein family. The structures of majorly prokaryotic ABC transporters were recently reviewed by Rees et al. [23], so we will concentrate on the three main human ABC transporters that are involved in multidrug resistance, ABCB1, ABCC1 and ABCG2. In the case of ABCB1 and ABCC1 all four domains are fused into a single polypeptide chain with the first TMD containing the N-terminus and the second NBD representing the C-terminus of the proteins. By contrast ABCG2 is a half transporter which has to homodimerize to be functional [24]. In addition, an inverse topology with respect to ABCB1 and ABCC1 can be observed, indicating that the NBD lies N-terminal of the TMD [25]. The hallmark of the ABCC1 transporter is a third TMD at the N-terminus referred to as TMD0.



**Fig. (8).** Comparison of different domain architecture of the ABC transporters ABCB1, ABCC1 and ABCG2.

ABC efflux pumps are flexible proteins that in association with drug binding and subsequent ATP hydrolysis undergo conformational changes. ABCB1 adopts at least three different states following ATP-binding and subsequent hydrolysis (reviewed in [26]). The apo or “open-inward” conformation is considered the ground state. In this conformation the protein shows an inverted “V” open towards the cytosolic environment of the cell. Substrates are considered to bind to this state with higher affinity. The second conformation that can be captured by ABCB1 is the nucleotide-bound form which is open to the extracellular space. After hydrolysis of two ATP molecules ABCB1 returns to the initial state (Fig. (9)).

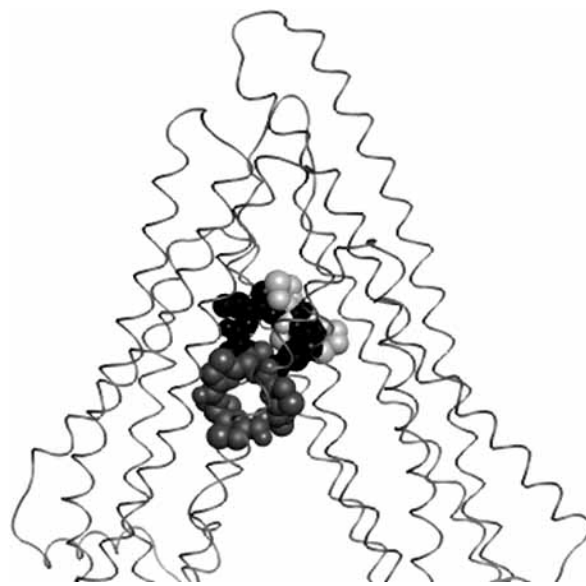


**Fig. (9).** Schematic illustration of the catalytic cycle of ABC transporters on the basis of ABCB1. The two different conformations are depicted before and after drug binding.

### Homology Models

The fact that ABC transporters are embedded in the membrane complicates the crystallization process of such proteins. Therefore, protein homology modeling based on templates of bacterial homologues representing different catalytic states, was the method of choice for structure-based studies. Table 2 gives an overview of current available homology models of selected ABC transporters. Due to its high resolution the crystal structure of the *Staphylococcus aureus* transporter SAV1866 (PDB code: 2HYD, resolution: 3.00 Å) [27] in the ADP bound “outward-facing” form often served as modeling template [28-33]. Interestingly, the same transporter crystallized in the AMP.PNP bound state [34] did not serve as modeling template. Several high resolution structures of different catalytic states of ABC-proteins were also obtained with the bacterial transporter MsbA [35] as template. This information gave new insights into the transport cycle and the associated conformational change of ABC proteins (Table 3).

Since March 2009 the first X-ray structure of a eukaryotic ABC efflux pump, ABCB1 (mouse) is available [36] (PDB code: 3G5U, resolution: 3.8 Å). With 87% sequence identity to human ABCB1 and moderate resolution (3.80) it serves as a good template for homology modeling [37]. Additionally the structure was published together with two co-crystallized enantiomeric cyclic peptide inhibitors (CPPIs; QZ59-RRR and QZ59-SSS) (Fig. (10)). This new information sheds light on possible ligand binding areas of ABCB1.



**Fig. (10).** Cocrystallized ABCB1 with cyclic P-gp inhibitors (CPPIs) QZ59-RRR (black) QZ59-SSS (dark and light grey).

**Table 2. Homology Models of the ABC Transporters ABCB1, ABCC1 and ABCG2**

| ABC Transporter | Template    | Sequence Identity / Homology | Catalytic State | References          |
|-----------------|-------------|------------------------------|-----------------|---------------------|
| ABCB1           | Mouse ABCB1 | 87 % / 93 %                  | Apo             | [36]                |
| ABCB1           | SAV1866     | 34 % / 52 %                  | ADP-bound       | [28-30, 33, 37, 67] |
| ABCB1           | MsbA        | 37 % / 57 %                  | AMP-PNP         | [28]                |
| ABCB1           | MalK        | 31 % / 50 %                  | Apo             | [33]                |
| ABCC1           | SAV1866     | 28 % / 49 %                  | ADP             | [31]                |
| ABCG2           | SAV1866     | 27 % / 49 %                  | ADP             | [32]                |

**Table 3. Structures of Whole ABC Transporters that are Available Until Now**

| ABC Transporter | Organism               | Catalytic State | Resolution [Å] | PDB Code | References |
|-----------------|------------------------|-----------------|----------------|----------|------------|
| ABCB1           | Mouse                  | Apo             | 3.80           | 3G5U     | [36]       |
| ABCB1           | Mouse                  | Apo             | 4.40           | 3G60     | [36]       |
| ABCB1           | Mouse                  | Apo             | 4.35           | 3G61     | [36]       |
| ABCB1           | Hamster                | Apo             | ~20            | -        | [68]       |
| ABCB1           | Hamster                | AMP-PNP         | ~20            | -        | [68]       |
| ABCB1           | Hamster                | ATP             | 8              | -        | [69]       |
| ABCC1           | Human                  | ATP             | ~22            | -        | [70]       |
| ABCG2           | Insect                 | ATP             | ~18            | -        | [71]       |
| SAV1866         | Staphylococcus aureus  | ADP             | 3.00           | 2HYD     | [27]       |
| SAV1866         | Staphylococcus aureus  | AMP-PNP         | 3.40           | 2ONJ     | [34]       |
| MsbA            | Escherichia coli       | Apo             | 5.30           | 3B5W     | [35]       |
| MsbA            | Vibrio cholerae        | Apo             | 5.50           | 3B5X     | [35]       |
| MsbA            | Salmonella typhimurium | AMP-PNP         | 4.50           | 3B5Y     | [35]       |
| MsbA            | Salmonella typhimurium | ADP-OV          | 4.20           | 3B5Z     | [35]       |
| MsbA            | Salmonella typhimurium | AMP-PNP         | 3.70           | 3B60     | [35]       |

### Binding Sites

It was shown that a functional unit of ABC-transporters has to consist of two TM and two NB domains. Only with this architecture a functional transporter can be obtained. Nevertheless, mutational studies showed that ABC transporters consisting of just two TMD regions without NBDs were able to bind ligands [38]. This led to the assumption that drug binding occurs in the TMD region.

Numerous experimental studies were performed trying to determine the different drug binding sites of P-glycoprotein, comprising among others cysteine and arginine scanning and photoaffinity labeling (reviewed in [26, 39, 40]). The overall assumption in this case is that P-glycoprotein possesses a huge binding pocket with at least four distinct binding sites, with TM 6 as main interaction helix. Well characterized are the binding sites of Rhodamine and Hoechst 33342, the so called R- and the H-site [41, 42]. Additionally, there is evidence for an allosteric regulatory site as well as a region where progesterone and prazosin may bind [43,

44]. These conclusions go hand in hand with the previously mentioned co-crystal structure of ABCB1 together with isomeric CPPIs [36]. The structure shows a huge binding pocket where the rather large cyclopeptides bind on different sites with partially overlapping interacting amino acid residues. Some of these residues are identical with the ones that are involved in rhodamine or verapamil binding [45, 46]. These data are also consistent with drug binding studies with the ABC transporter ABCG2. Also for ABCG2 at least four different binding sites, one H-site, a prazosin area and probably two different R-sites on each monomer have been postulated. [47]. The involvement of both monomers in rhodamine 123 binding can also be observed with ABCC1 where TMD1 and TMD2 are interacting [48].

Nature derived substrates, especially cytotoxins, are supposed to bind to a certain region in the binding pocket of the transmembrane domains of ABC transporters. However, large compounds with a steroidal architecture tend to bind to the ATP-binding site in the NBD region of the protein. As competitors of

ATP they are also able to inhibit the function of the MDR transporter.

### Ligand Docking

The computational method of ligand docking is a good way to validate experimentally derived binding pockets or even to propose new areas of binding. Several docking studies of natural compounds have been performed. Recently published docking results show quinazolinones binding at the same site like the CPPIs [37]. The docking poses are in accordance with pharmacophore modeling, which suggests a hydrogen bond between the ligand and the amino acid residue Tyr307 (TM5). In addition, protein-ligand interaction fingerprints (PLIF) were calculated, resulting in the residues Phe336 (TM6), Tyr953 (TM11) and Phe957 (TM11) performing contact interactions (Fig. (11)). The binding pocket was described as highly hydrophobic which excludes ionic interactions with tertiary amines. Therefore it was suggested that such interactions can be built after the conformational change of the protein and thus has to be validated with an outward facing model.

Similar results were also obtained in our group when performing docking studies with a homology model of ABCB1 and propafenone derivatives. Our results also showed interactions with the transmembrane helices mentioned above. This confirms the assumption of a large binding pocket and indicates overlapping quinazolinone and propafenone binding sites. In Fig. (12) an overview of interactions of drugs with certain TM helices is depicted. As can be noticed, TM 6 plays a crucial role in ligand binding.

The assumption that certain ABC transporter inhibitors of natural origin compete with ATP at the NBDs could also be confirmed by docking [49]. A screening of 122 compounds against the three MDR related proteins ABCB1, ABCC2 and ABCG2, revealed that several compounds showed multi-specificity. Since the highest sequence identity among these proteins can be found in the NBDs these compounds were docked into the crystal structure of the NBD1 of ABCC1 [50]. The results showed that the most hydrophilic natural products quercetin and silymarin together with the potent compound MK571 were able to bind to the structure with

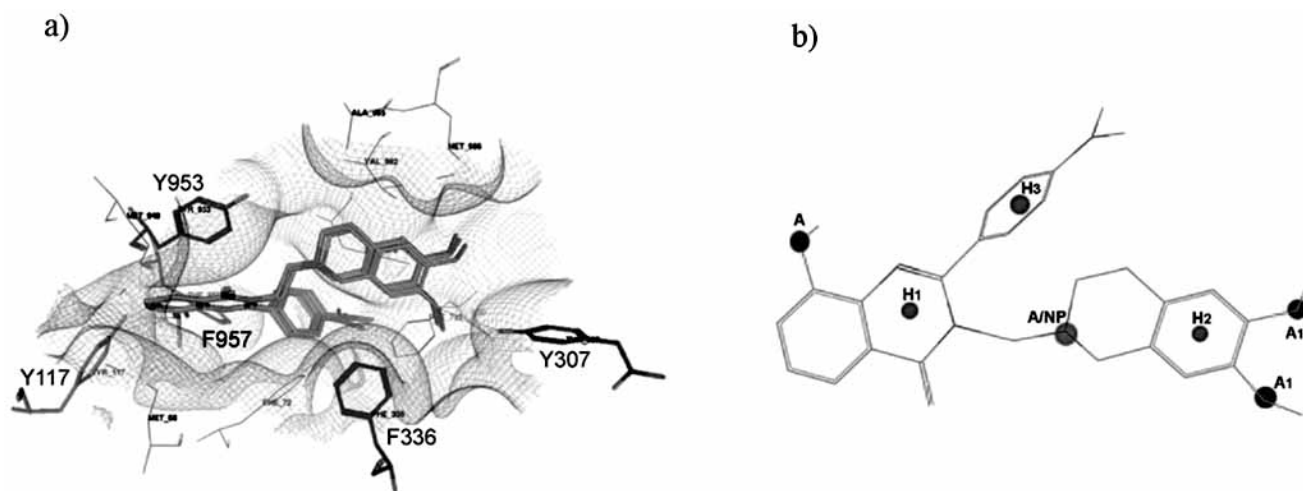


Fig. (11). a) Docking poses of quinazolinones in an ABCB1 homology model, b) Pharmacophore model (taken from [37]).

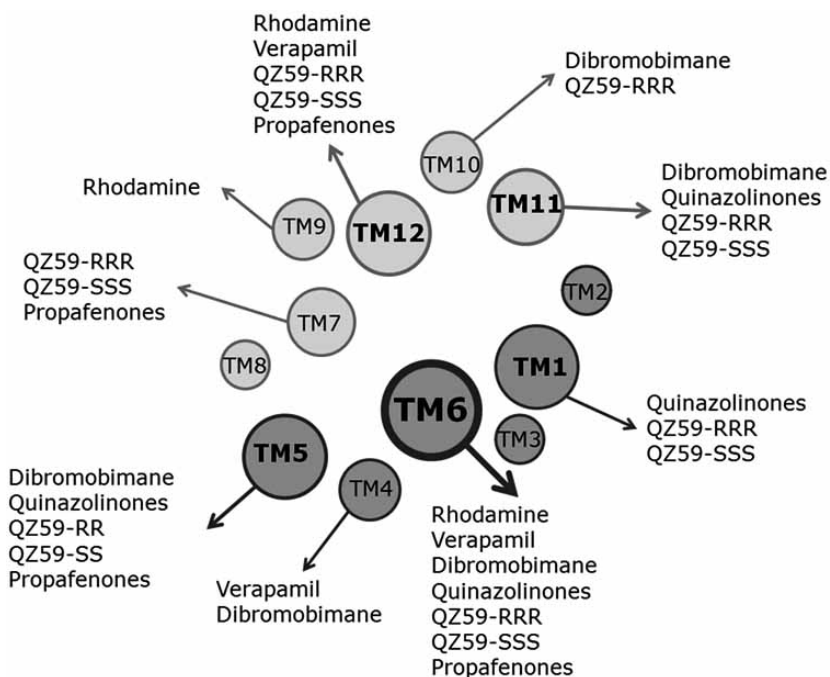


Fig. (12). Transmembrane (TM) helix interactions with investigated ABCB1 ligands. The circle size depends on the level of interaction.

high scores. More lipophilic inhibitors were not able to provide reasonable scoring values. Regarding the docking poses obtained it is noteworthy to mention that the negatively charged MK571 extends into the catalytic site and its aromatic rings are placed similar to the adenosine base ring of ATP. By contrast the poses of the lipophilic inhibitors showed no interaction with the catalytic site (Fig. (13)).

Also steroids and flavonoids were examined with respect to their binding affinity to the ATP binding site [51]. In this study docking of eleven different steroids, one flavonoid, ATP and MANT-ATP into ABCB1 and ABCG2 was performed. The results, which were rather the same for both transporters, suggest overlapping steroid and ATP binding sites near the P-loop of the nucleotide binding domains (Fig. (14) and (15)). The P-loop (or Walker A) is one of the three characteristic motifs of the NBDs of ABC transporters (Walker A, Walker B and signature motif C) and interacts with the phosphates of the nucleotides. The flavone kaempferide showed amino acid residue interactions similar to ATP. On the other hand the hydrophobic steroid RU-486 bound to a different area than the other steroids and ATP, but overlapped with the kaempferide and the MANT-ATP binding site. RU-486 and MANT-ATP share a highly hydrophobic moiety and both bind

within the hydrophobic cleft around I1050 (Fig. (14e)). Additionally the binding free energy of the complexes was calculated. According to this study the steroids investigated bind with the same affinity as ATP, which renders them potential competitors of ATP (Table 4).

Similar findings were published in a docking study that concentrated on flavonoids, including flavones, flavonols, flavanones and chalcones [52] (Fig. (16)). Calculated binding free energies were compared to experimentally derived  $K_d$ -values and a good correlation could be obtained. This study also showed that flavonoids preferably bind to the P-loop of the NBD, especially interacting with residues L1076 and S1077. In addition, the B-ring of flavonoids was supposed to build hydrophobic interactions with Y1044, which originally interacts with the adenosine base of ATP [52]. Comparing the different flavonoid derivatives showed that the additional hydroxyl-group at position 3, which is the only difference between flavonols and flavones, decreases the predicted docking energy because an additional hydrogen bond could be formed. Additional hydrophobic substituents added to flavones and flavonols at positions 6 or 8 also had a positive effect on binding. Chalcones, which show higher flexibility due to the open C-ring structure, also showed reduced docking energy. Especially with

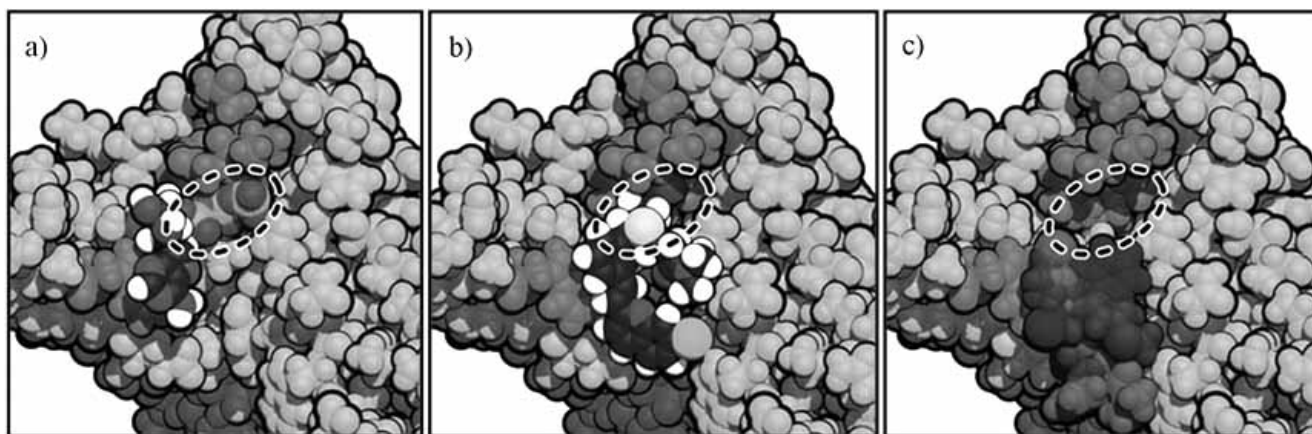
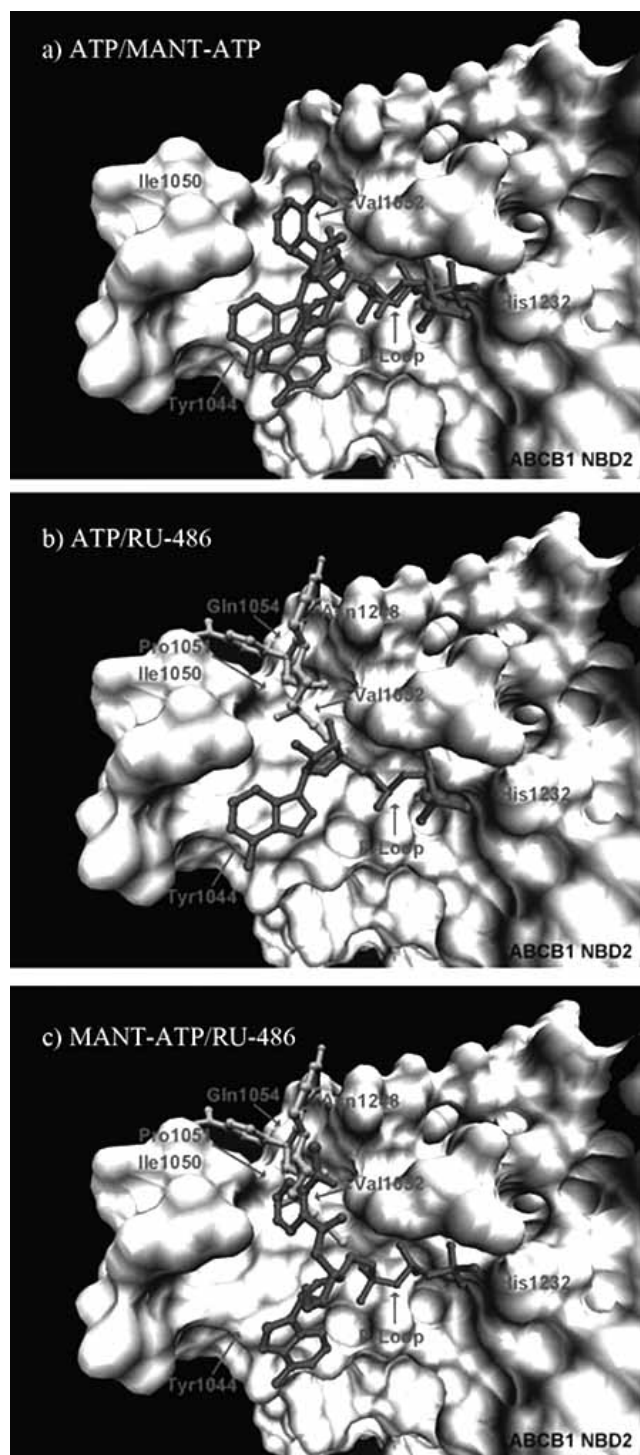


Fig. (13). a) MRP NBD1 cocrystallized with ATP. b) MRP NBD1 with MK-571. c) MRP with lipophilic inhibitors (taken from [50]).

Table 4. Amino Acid Interactions Observed With Docking Studies of Steroids and Flavonoids Into the NBD

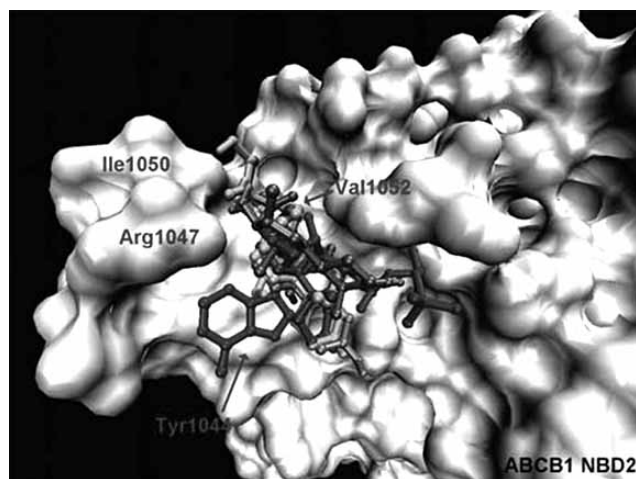
| Compound              | ABCB2 NBD2<br>Hydrophobic Interaction         | Hydrophilic Interaction                         | References |
|-----------------------|---|---|------------|
| Steroids              | Y1044, I1050, V1052, G1075, S1077             | L1076, R1047, Q1085, P1051                      | [51]       |
| RU-486                | I1050, P1051, V1052, Q1054, N1248             | None predicted                                  | [51]       |
| Kaempferide           | Y1044, V1052, G1073, G1075                    | R1047, S1077, T1078                             | [51]       |
| ATP                   | Y1044, G1073                                  | G1073, C1074, G1075, S1077, T1078               | [51]       |
| MANT-ATP              | Y1044, I1050, V1052                           | G1073, C1074, G1075, L1076, S1077, T1078, Y1087 | [51]       |
| Flavones              | G1070-T1077                                   | L1076   | [52]       |
| substituted Flavones  | Y1044, V1052, G1072-T1075, G1073, G1075,      | S1072, L1076, S1077, T1078                      | [52]       |
| Flavonols             | G1071-T1076, E1201, D1200, H1232              | S1071, C1074, G1075, S1077                      | [52]       |
| Chalcones             | Y1044, V1052, G1071-T1078,                    | S1077, T1078                                    | [52]       |
| Substituted Chalcones | Y1044, P1048, I1050-V1052, H1232, Q1247-E1249 | S1072, G1073, C1074, G1075, L1076, S1077        | [52]       |



**Fig. (14).** Docking poses of MANT-ATP, ATP and RU-486 into the homology models of ABCB1 NBD2 and ABCG2 NBD. (taken from [51]).

substituted chalcone derivatives, such as *O-n*-C<sub>10</sub>H<sub>21</sub> chalcone, very low docking energy values were predicted.

Until now the number of docking studies into ABC transporters is still low. As outlined above, most docking studies are restricted to the nucleotide binding domain. This can be explained by the lack of crystal structures of the transmembrane domain, which is the part



**Fig. (15).** Docking poses of steroids in a homology model of ABCB1 NBD2 (taken from [51]).

of the protein with quite low sequence similarity. However, this trend will probably change due to the recent publication of the structure of mouse P-gp.

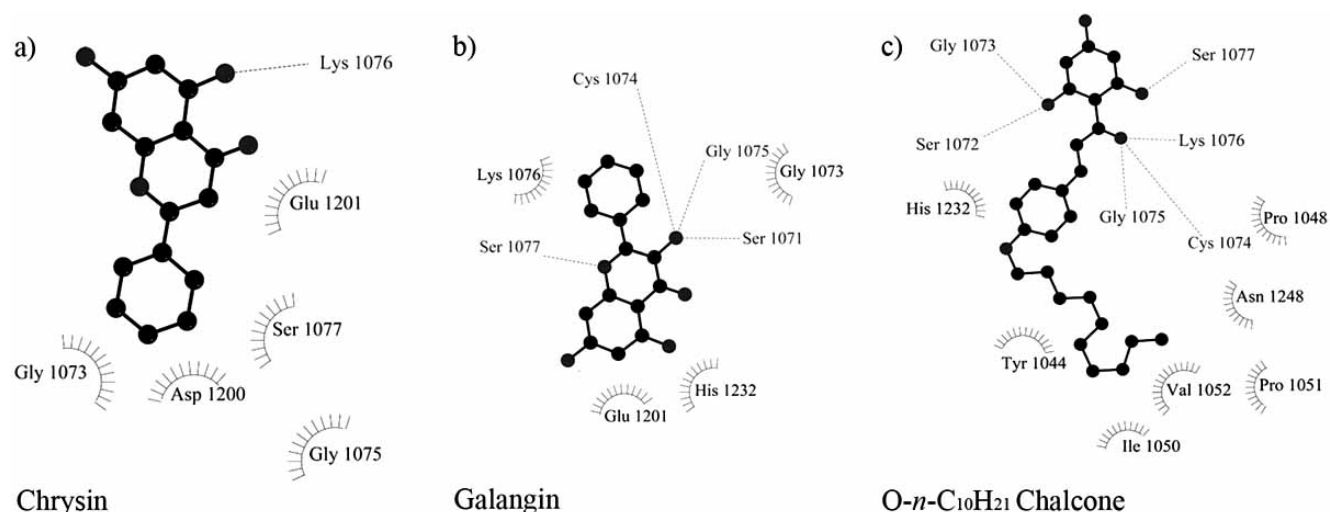
#### IMPORTANCE OF ABC-TRANSPORTER FOR ADMET

With our increasing knowledge on the physiological role of ABC transporter it became evident that there are several distinct transporters which are responsible for severe side effects of drugs and for drug/drug interactions. In these cases the focus shifts from the design of inhibitors to the design of “non-ligands”. Thus, the major challenge is to establish models for prediction of substrate properties with the ultimate goal to avoid interaction with these proteins.

ABCB1 is constitutively expressed at several diffusion barriers, such as the blood-brain barrier, the kidney, the liver and the intestine. At the latter it plays an important role in limiting the intestinal absorption of a wide variety of orally administered drugs. One paradigm example is the quinidine-digoxin interaction, where the P-gp inhibitor quinidine increases the digoxin absorption rate by about 30%. But it is not only drug/drug interaction playing a role, there is also proven evidence for drug/nutrient interaction [53]. These include mainly flavonoids found in fruit juices, vegetables, flowers and tea. Especially grapefruit juice has been shown to interfere with plasma levels of colchicines [54], paracetamol [55], and cyclosporine [56].

Thus, the importance of drug transporters for uptake and disposition is now widely accepted and Benet and co-workers proposed a biopharmaceutics Classification System (BCS) which allows prediction of *in vivo* pharmacokinetic performance of drug candidates based on measurements of their permeability and solubility [57]. Subsequently, this classification system was modified in order to allow prediction of overall drug disposition, including routes of drug elimination and the effects of efflux and absorptive transporters on oral drug absorption [58]. The overall message is that compounds with low water solubility being substrates of P-glycoprotein bear the inherent risk of low bioavailability.

Also at the blood-brain barrier (BBB) the important role of ABCB1 and ABCG2 is increasingly recognised. *In vitro* studies demonstrated that the uptake of vincristine was reduced in primary cultured bovine capillary endothelial cells expressing P-gp at the luminal side and that this decreased accumulation was due to active efflux. Steady state uptake was significantly increased in the presence of the P-gp blocking agent verapamil [59]. Additionally, *mdr1a* double knock out mice show hypersensitivity to a range of



**Fig. (16).** Observed interaction of ABCB1 NBD2 with chrysin (flavones) (a), galangin (flavonol) (b) and O-*n*-C<sub>10</sub>H<sub>21</sub>-chalcone (c) [52].

drugs known to be transported by P-gp [60]. Undoubtedly, selected ABC-transporter are an important impediment for the entry of hydrophobic drugs into the brain.

#### PREDICTING SUBSTRATE PROPERTIES FOR ABCB1

As already outlined above, ABCB1 is constitutively expressed in several organs, such as kidney, liver, intestine and also at the blood brain barrier (BBB). P-gp substrates therefore show poor oral absorption, enhanced renal and biliary excretion and usually do not enter the brain [61]. Furthermore, they are likely to be affected by the MDR phenotype and are thus not suitable as anticancer agents. This spurred the development of medium- and high-throughput systems addressing the P-gp substrate properties of compounds of interest.

However, data sets for *in silico* classification studies are rather small and sometimes also inconsistent [62]. Recently the group of Gottesman published a comprehensive study analysing data from the NCI60 screen [63], which comprises mostly natural product toxins. m-RNA levels of all 48 human ABC-transporter in 60 human tumour cell lines of the NCI60 anticancer drug screening panel were evaluated and correlated with cellular toxicity values of 1400 selected compounds. An inverse correlation between transporter mRNA levels and compound toxicity indicates that a compound is a substrate for the respective transporter. Undoubtedly, this is by far the largest consistent data set available by now. It is almost exclusively built of natural products, and studies from our group indicate that it might be successfully used as basis for P-gp substrate prediction models.

Based both on this data set as well as on a set of 259 compounds compiled from the literature we explored the performance of several classification methods combined with different descriptor sets. These include simple ADME-type descriptors (such as logP, number of rotatable bonds, number of H-bond donors and acceptors), VSA descriptors as described by Labute [64] and 2D auto-correlation vectors. The latter have already been successfully applied for prediction of P-gp inhibitors [65]. When comparing binary QSAR and support vector machines, the latter gave more robust models with total accuracies in the range of 80%. Generally, the prediction of non-substrates performs better than those for substrates [66]. However, more detailed studies are necessary to fully explore the potential and limits of this data set. If successful, this approach might be useful for *in silico* screening of natural product libraries in order to identify hitherto unknown drug/nutrient

interactions at P-gp and related ABC-transporter involved in ADMET.

#### OUTLOOK

Although P-glycoprotein and its prominent role in tumour multidrug resistance is known since 1976, up to now no P-gp inhibitor has reached the market. Thus, there is still need for development of new, specific P-gp inhibitors. As P-gp is mainly addressing natural product toxins as substrates, compounds from natural origin are versatile starting points for design of new ligands. Due to the polyspecificity of the protein, complex methods such as self organising maps or random forest classification might pave the way for successful *in silico* screening approaches, targeted at natural compound libraries. However, within the past decade the focus of interest shifted towards the role of ABC-transporters for ADMET and drug/drug interactions. Several pharmaceutical companies established high throughput screening systems for measuring P-gp substrate properties of their compound libraries and *in silico* methods have been developed which reach classification accuracies in the range of 80%. In this case the most comprehensive data set available up to now uses data from the NCI60 screening library, which is mostly composed of natural product related toxins. Finally, the publication of the structure of mouse P-glycoprotein will aid in the understanding of the molecular principles underlying the ligand-polyspecificity of these transporters and pave the way for structure-based drug design approaches.

#### ACKNOWLEDGMENT

We gratefully acknowledge financial support from the Austrian Science Fund (grant # F3502 and F3509). Ishrat Jabeen is grateful to the Higher Education Commission Pakistan (HEC) for financial support.

#### REFERENCES

- [1] Gottesman MM, Ling V. The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research. *FEBS Lett* 2006; 580: 998-1009.
- [2] Szakacs G, Paterson JK, Ludwig JA, Booth-Genthe C, Gottesman MM. Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* 2006; 5: 219-34.
- [3] Tsuruo T, Iida H, Tsukagoshi S, Sakurai Y. Overcoming of vincristine resistance in P388 leukemia *in vivo* and *in vitro* through enhanced cytotoxicity of vincristine and vinblastine by verapamil. *Cancer Res* 1981; 41: 1967-72.

- [4] Juliano RL, Ling V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta* 1976; 455: 152-62.
- [5] Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. *J Lipid Res* 2001; 42: 1007-17.
- [6] Szakacs G, Varadi A, Ozvegy-Laczka C, Sarkadi B. The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). *Drug Discov Today* 2008; 13: 379-93.
- [7] Raub TJ. P-glycoprotein recognition of substrates and circumvention through rational drug design. *Mol Pharm* 2006; 3: 3-25.
- [8] Pleban K, Ecker GF. Inhibitors of p-glycoprotein--lead identification and optimisation. *Mini Rev Med Chem* 2005; 5: 153-63.
- [9] Zhang S, Yang X, Coburn RA, Morris ME. Structure activity relationships and quantitative structure activity relationships for the flavonoid-mediated inhibition of breast cancer resistance protein. *Biochem Pharmacol* 2005; 70: 627-39.
- [10] Zhang S, Yang X, Morris ME. Flavonoids are inhibitors of breast cancer resistance protein (ABCG2)-mediated transport. *Mol Pharmacol* 2004; 65: 1208-16.
- [11] Bansal T, Jaggi M, Khar RK, Talegaonkar S. Emerging significance of flavonoids as P-glycoprotein inhibitors in cancer chemotherapy. *J Pharm Pharm Sci* 2009; 12: 46-78.
- [12] 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 USA* 1998; 95: 9831-6.
- [13] Cramer J, Kopp S, Bates SE, Chiba P, Ecker GF. Multispecificity of drug transporters: probing inhibitor selectivity for the human drug efflux transporters ABCB1 and ABCG2. *ChemMedChem* 2007; 2: 1783-8.
- [14] Free SM, Jr., Wilson JW. A Mathematical contribution to structure-activity studies. *J Med Chem* 1964; 7: 395-9.
- [15] Pick A, Muller H, Wiese M. Structure-activity relationships of new inhibitors of breast cancer resistance protein (ABCG2). *Bioorg Med Chem* 2008; 16: 8224-36.
- [16] Ueda K, Saeki T, Hirai M, Tanigawara Y, Tanaka K, Okamura M, *et al.* Human P-glycoprotein as a multi-drug transporter analyzed by using transepithelial transport system. *Jpn J Physiol* 1994; 44 (Suppl 2): S67-71.
- [17] Yang CP, DePinho SG, Greenberger LM, Arceci RJ, Horwitz SB. Progesterone interacts with P-glycoprotein in multidrug-resistant cells and in the endometrium of gravid uterus. *J Biol Chem* 1989; 264: 782-8.
- [18] Hamilton KO, Yazdani MA, Audus KL. Modulation of P-glycoprotein activity in Calu-3 cells using steroids and beta-ligands. *Int J Pharm* 2001; 228: 171-9.
- [19] Li Y, Wang YH, Yang L, Zhang SW, Liu CH, Yang SL. Comparison of steroid substrates and inhibitors of P-glycoprotein by 3D-QSAR analysis. *J Mol Structure* 2005; 733: 111-8.
- [20] Barnes KM, Dickstein B, Cutler GB, Jr., Fojo T, Bates SE. Steroid treatment, accumulation, and antagonism of P-glycoprotein in multidrug-resistant cells. *Biochemistry* 1996; 35: 4820-7.
- [21] Muñoz-Martínez F, Reyes CP, Pérez-Lomas AL, Jiménez IA, Gamarro F, Castans S. Insights into the molecular mechanism of action of Celastraceae sesquiterpenes as specific, non-transported inhibitors of human P-glycoprotein. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 2006; 1758: 98-110.
- [22] Reyes CP, Muñoz-Martínez F, Torrecillas IR, Mendoza CR, Gamarro F, Bazzocchi IL, *et al.* Biological evaluation, structure-activity relationships, and three-dimensional quantitative structure-activity relationship studies of dihydro-beta-agarofuran sesquiterpenes as modulators of P-glycoprotein-dependent multidrug resistance. *J Med Chem* 2007; 50: 4808-17.
- [23] Rees DC, Johnson E, Lewinson O. ABC transporters: the power to change. *Nat Rev Mol Cell Biol* 2009; 10: 218-27.
- [24] Kage K, Tsukahara S, Sugiyama T, Asada S, Ishikawa E, Tsuruo T, *et al.* Dominant-negative inhibition of breast cancer resistance protein as drug efflux pump through the inhibition of S-S dependent homodimerization. *Int J Cancer* 2002; 97: 626-30.
- [25] Allikmets R, Schriml LM, Hutchinson A, Romano-Spica V, Dean M. A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res* 1998; 58: 5337-9.
- [26] Seeger MA, van Veen HW. Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta* 2009; 1794: 725-37.
- [27] Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. *Nature* 2006; 443: 180-5.
- [28] Becker JP, Depret G, Van Bambeke F, Tulkens PM, Prevost M. Molecular models of human P-glycoprotein in two different catalytic states. *BMC Struct Biol* 2009; 9: 3.
- [29] Globisch C, Pajeva IK, Wiese M. Identification of putative binding sites of P-glycoprotein based on its homology model. *Chem Med Chem* 2008; 3: 280-95.
- [30] Stockner T, de Vries SJ, Bonvin AM, Ecker GF, Chiba P. Data-driven homology modelling of P-glycoprotein in the ATP-bound state indicates flexibility of the transmembrane domains. *FEBS J* 2009; 276: 964-72.
- [31] DeGorter MK, Conseil G, Deeley RG, Campbell RL, Cole SP. Molecular modeling of the human multidrug resistance protein 1 (MRP1/ABCC1). *Biochem Biophys Res Commun* 2008; 365: 29-34.
- [32] Hazai E, Bikadi Z. Homology modeling of breast cancer resistance protein (ABCG2). *J Struct Biol* 2008; 162: 63-74.
- [33] O'Mara ML, Tieleman DP. P-glycoprotein models of the apo and ATP-bound states based on homology with Sav1866 and MalK. *FEBS Lett* 2007; 581: 4217-22.
- [34] Dawson RJ, Locher KP. Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett* 2007; 581: 935-8.
- [35] Ward A, Reyes CL, Yu J, Roth CB, Chang G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci USA* 2007; 104: 19005-10.
- [36] Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, *et al.* Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 2009; 323: 1718-22.
- [37] Pajeva IK, Globisch C, Wiese M. Combined pharmacophore modeling, Docking, and 3D QSAR studies of ABCB1 and ABCC1 transporter inhibitors. *Chem Med Chem* 2009; 4(11): 1883-96.
- [38] Loo TW, Clarke DM. The transmembrane domains of the human multidrug resistance P-glycoprotein are sufficient to mediate drug binding and trafficking to the cell surface. *J Biol Chem* 1999; 274: 24759-65.
- [39] Loo TW, Clarke DM. Mutational analysis of ABC proteins. *Arch Biochem Biophys* 2008; 476: 51-64.
- [40] Shilling RA, Venter H, Velamakanni S, Bapna A, Woecking B, Shahi S, *et al.* New light on multidrug binding by an ATP-binding-cassette transporter. *Trends Pharmacol Sci* 2006; 27: 195-203.
- [41] Loo TW, Clarke DM. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem* 2002; 277: 44332-8.
- [42] Qu Q, Sharom FJ. Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet. *Biochemistry* 2002; 41: 4744-52.
- [43] Martin C, Berridge G, Higgins CF, Mistry P, Charlton P, Callaghan R. Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* 2000; 58: 624-32.
- [44] Shapiro AB, Fox K, Lam P, Ling V. Stimulation of P-glycoprotein-mediated drug transport by prozasin and progesterone. Evidence for a third drug-binding site. *Eur J Biochem* 1999; 259: 841-50.
- [45] Loo TW, Bartlett MC, Clarke DM. Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. *Biochem J* 2006; 399: 351-9.
- [46] Loo TW, Clarke DM. Identification of residues in the drug-binding site of human P-glycoprotein using a thiol-reactive substrate. *J Biol Chem* 1997; 272: 31945-8.
- [47] Clark R, Kerr ID, Callaghan R. Multiple drugbinding sites on the R482G isoform of the ABCG2 transporter. *Br J Pharmacol* 2006; 149: 506-15.
- [48] Deeley RG, Cole SP. Substrate recognition and transport by multidrug resistance protein 1 (ABCC1). *FEBS Lett* 2006; 580: 1103-11.
- [49] Matsson P, Pedersen JM, Norinder U, Bergstrom CA, Artursson P. Identification of novel specific and general inhibitors of the three major human ATP-binding cassette transporters P-gp, BCRP and MRP2 among registered drugs. *Pharm Res* 2009; 26: 1816-31.
- [50] Ramaen O, Leulliot N, Sizun C, Ulryck N, Pamard O, Lallemand JY, *et al.* Structure of the human multidrug resistance protein 1 nucleotide binding domain 1 bound to Mg<sup>2+</sup>/ATP reveals a non-productive catalytic site. *J Mol Biol* 2006; 359: 940-9.

- [51] Mares-Samano S, Badhan R, Penny J. Identification of putative steroid-binding sites in human ABCB1 and ABCG2. *Eur J Med Chem* 2009; 44: 3601-11.
- [52] Badhan R, Penny J. *In silico* modelling of the interaction of flavonoids with human P-glycoprotein nucleotide-binding domain. *Eur J Med Chem* 2006; 41: 285-95.
- [53] Aszalos A. Role of ATP-binding cassette (ABC) transporters in interactions between natural products and drugs. *Curr Drug Metab* 2008; 9: 1010-8.
- [54] Dahan A, Amidon GL. Grapefruit juice and its constituents augment colchicine intestinal absorption: potential hazardous interaction and the role of p-glycoprotein. *Pharm Res* 2009; 26: 883-92.
- [55] Dasgupta A, Reyes MA, Risin SA, Actor JK. Interaction of white and pink grapefruit juice with acetaminophen (paracetamol) *in vivo* in mice. *J Med Food* 2008; 11: 795-8.
- [56] Paine MF, Widmer WW, Pusek SN, Beavers KL, Criss AB, Snyder J, *et al.* Further characterization of a furanocoumarin-free grapefruit juice on drug disposition: studies with cyclosporine. *Am J Clin Nutr* 2008; 87: 863-71.
- [57] Wu CY, Benet LZ. Predicting drug disposition *via* application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* 2005; 22: 11-23.
- [58] Custodio JM, Wu CY, Benet LZ. Predicting drug disposition, absorption/elimination/transporter interplay and the role of food on drug absorption. *Adv Drug Deliv Rev* 2008; 60: 717-33.
- [59] Tsuji A, Terasaki T, Takabatake Y, Tenda Y, Tamai I, Yamashita T, *et al.* P-glycoprotein as the drug efflux pump in primary cultured bovine brain capillary endothelial cells. *Life Sci* 1992; 51: 1427-37.
- [60] Borst P, Schinkel AH. Introduction to the blood-brain barrier. In: Partridge WM, Ed. Cambridge: Cambridge University Press 1998; pp. 198-206.
- [61] Chan LM, Lowes S, Hirst BH. The ABCs of drug transport in intestine and liver: efflux proteins limiting drug absorption and bioavailability. *Eur J Pharm Sci* 2004; 21: 25-51.
- [62] Chiba P, Ecker GF, Eds. Transporters as Drug Carriers. Weinheim: Wiley-VCH 2009; pp. 349-59.
- [63] Szakacs G, Annereau JP, Lababidi S, Shankavaram U, Arciello A, Bussey KJ, *et al.* Predicting drug sensitivity and resistance: profiling ABC transporter genes in cancer cells. *Cancer Cell* 2004; 6: 129-37.
- [64] Labute P. A widely applicable set of descriptors. *J Mol Graph Model* 2000; 18: 464-77.
- [65] Kaiser D, Terfloth L, Kopp S, Schulz J, de Laet R, Chiba P, *et al.* Self-organizing maps for identification of new inhibitors of P-glycoprotein. *J Med Chem* 2007; 50: 1698-702.
- [66] Zdrzil B, Prakashvudhisam C, Ecker GF. NCI60 screening data - a versatile tool for *in silico* models predicting substrate properties for ABC-transporter. ACS 234th National Meeting & Exposition. Boston 2007.
- [67] Sakurai A, Onishi Y, Hirano H, Seigneuret M, Obanayama K, Kim G, *et al.* Quantitative structure-activity relationship analysis and molecular dynamics simulation to functionally validate nonsynonymous polymorphisms of human ABC transporter ABCB1 (P-glycoprotein/MDR1). *Biochemistry* 2007; 46: 7678-93.
- [68] Rosenberg MF, Kamis AB, Callaghan R, Higgins CF, Ford RC. Three-dimensional structures of the mammalian multidrug resistance P-glycoprotein demonstrate major conformational changes in the transmembrane domains upon nucleotide binding. *J Biol Chem* 2003; 278: 8294-9.
- [69] Rosenberg MF, Callaghan R, Modok S, Higgins CF, Ford RC. Three-dimensional structure of P-glycoprotein: the transmembrane regions adopt an asymmetric configuration in the nucleotide-bound state. *J Biol Chem* 2005; 280: 2857-62.
- [70] Rosenberg MF, Mao Q, Holzenburg A, Ford RC, Deeley RG, Cole SP. The structure of the multidrug resistance protein 1 (MRP1/ABCC1), crystallization and single-particle analysis. *J Biol Chem* 2001; 276: 16076-82.
- [71] McDevitt CA, Collins RF, Conway M, Modok S, Storm J, Kerr ID, *et al.* Purification and 3D structural analysis of oligomeric human multidrug transporter ABCG2. *Structure* 2006; 14: 1623-32.

---

Received: February 8, 2010

Accepted: February 26, 2010



## APPENDIX

### Page 169-226

|                                      |              |
|--------------------------------------|--------------|
| A1. Supplementary data for chapter 2 | Page 170-172 |
| A2. Supplementary data for chapter 5 | Page 173-193 |
| A3. Supplementary data for chapter 6 | Page 194-226 |

## Supporting Information

# Synthesis, ABCB1 inhibitory activity and 3D- QSAR studies of a series of new chalcone derivatives

Brunhofer Gerda<sup>1#</sup>, Jabeen Ishrat<sup>1#</sup>, Parveen Zahida<sup>2#</sup>, Berner Heinz<sup>1</sup>, Manuel Pastor,  
Chiba Peter<sup>2</sup>, Erker Thomas, and Ecker Gerhard F<sup>1\*</sup>

<sup>1</sup>University of Vienna, Department of Medicinal Chemistry, Althanstrasse 14, 1090

Wien, Austria <sup>2</sup>Medical University of Vienna, Institute of Medical Chemistry,

Währinger Straße 10, 1090 Wien, Austria

|                       |                |
|-----------------------|----------------|
| <b>Chemistry</b>      | <b>171-172</b> |
| <b>Suppl Figure 1</b> | <b>172</b>     |
| <b>References</b>     | <b>172</b>     |

## APPENDIX A1

**(E)-2-Methylthiocinnamaldehyde (1).** Yield: 0.660 g (90%) yellow crystals; mp: 74-76 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 9.75 (AB-system, 1H, *J*<sub>AB</sub>=7.8 Hz), 8.03 (AB-system, 1H, *J*<sub>AB</sub>=15.8 Hz), 7.69-7.51 (m, 1H), 7.49-7.15 (m, 3H), 6.67 (2x AB-system, 1H, *J*<sub>AB</sub>=15.8 Hz, *J*<sub>AB</sub>=7.8 Hz), 2.51 (s, 3H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 193.9, 149.2, 139.8, 133.0, 131.2, 129.9, 127.6, 127.3, 125.8, 16.7. MS *m/z*: 178 (9%, M<sup>+</sup>), 149 (75%), 134 (100%), 131 (95%), 116 (12%). Anal. Calcd for C<sub>10</sub>H<sub>10</sub>OS: C 67.38; H 5.66. Found: C 67.60; H 5.45.

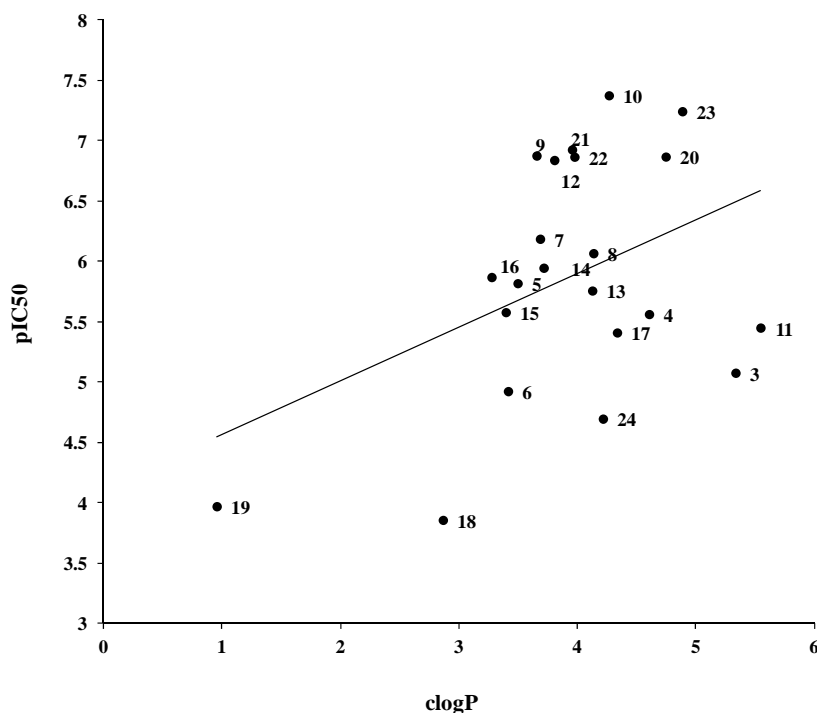
**3,4,5-Trimethoxycinnamaldehyde (2).**<sup>1</sup> Yield: 1.022 g (92%) pale yellow crystals; mp: 112-114°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 9.69 (AB-system, 1H, *J*<sub>AB</sub>=7.7 Hz), 7.40 (AB-system, 1H, *J*<sub>AB</sub>=15.9 Hz), 6.80 (s, 2H), 6.64 (2x AB-system, 1H, *J*<sub>AB</sub>=15.9 Hz, *J*<sub>AB</sub>=7.7 Hz), 3.91 (s, 9H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 193.4, 153.5, 152.7, 129.4, 127.9, 105.6, 61.0, 56.2. MS *m/z*: 178 (9%, M<sup>+</sup>), 149 (75%), 134 (100%), 131 (95%), 116 (12%).

**(E)-3-(2,4,6-Trimethylphenyl)-1-(2'-methoxyphenyl)-2-propen-1-one (3).**<sup>2</sup> Yield: 1.562 g (98%) white crystals; mp: 116-117 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 7.80 (AB-system, 1H, *J*<sub>AB</sub>=16.3 Hz), 7.71-7.60 (m, 1H), 7.53-7.40 (m, 1H), 7.12-6.88 (m, 5H), 3.87 (s, 3H), 2.37 (s, 6H), 2.29 (s, 3H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 192.9, 158.1, 141.5, 138.2, 137.1 (2C), 132.9, 132.0, 131.7, 130.5, 129.1 (2C), 120.7, 111.4, 55.6, 21.1 (2C), 21.0. MS *m/z*: 280 (21%, M<sup>+</sup>), 265 (100%), 249 (14%), 135 (64%), 129 (30%). Anal. Calcd for C<sub>19</sub>H<sub>20</sub>O<sub>2</sub>·0.1H<sub>2</sub>O: C 80.88; H 7.22. Found: C 80.82; H 7.11.

**6-Methoxy-2-[1-naphthalen-2-yl-methylidene]-3,4-dihydro-2H-naphthalen-1-one (11).**<sup>3</sup> The compound was synthesized by using 2.5 mmol (0.441 g) 6-methoxy-1-tetralone and 2.5 mmol (0.391 g) 2-naphthaldehyde. Yield: 0.130 g (17%) brown crystals. mp: 145-148°C. <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 8.15 (d, *J*=8.7 Hz, 1H), 8.06-7.79 (m, 5H), 7.63-7.42 (m, 3H), 6.90 (dd, *J*=8.7 Hz, *J*=2.4 Hz, 1H), 6.72 (d, *J*=2.4 Hz, 1H), 3.88 (s, 3H), 3.30-3.11 (m, 2H), 3.03-2.85 (m, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>): δ 190.4, 163.6,

## APPENDIX A1

149.2, 145.7, 136.1, 135.9, 133.6, 133.1, 133.0, 130.8, 129.4, 128.3, 128.0, 127.7, 127.3, 126.7, 126.4, 113.4, 112.3, 55.5, 29.3, 27.4. MS:  $m/z$  314 ( $M^+$ , 58%), 313 (100%), 165 (25%), 141 (16%), 120 (18%). Anal. Calcd for  $C_{22}H_{18}O_2$ : C 84.05; H 5.77. Found: C 83.77; H 5.90.



**Suppl figure 1.** Correlation of P-gp inhibitory activity of compounds **3-24** (expressed as  $\log(1/IC_{50})$  vs  $\log P$  of the ligands.

## References

1. Joshi, B. P.; Sharma, A.; Sinha, A. K. Ultrasound-assisted convenient synthesis of hypolipidemic active natural methoxylated (E)-arylalkenes and arylalkanones. *Tetrahedron* **2005**, 61, 3075-3080.
2. Barnes, R. P.; Cochrane, C. C. The Properties of o-Methoxybenzoylmesitylmethane. *Journal of the American Chemical Society* **1942**, 64, 2262-2262.
3. Xu, W.-Z.; Huang, Z.-T.; Zheng, Q.-Y. Synthesis of Benzo[c]xanthenes from 2-Benzylidene-1-tetralones by the Ultraviolet Radiation-Mediated Tandem Reaction. *The Journal of Organic Chemistry* **2008**, 73, 5606-5608.

# Probing the stereoselectivity of P-glycoprotein – synthesis, biological activity and ligand docking studies of a set of enantiopure benzopyrano[3,4b][1,4]oxazines

Ishrat Jabeen,<sup>a</sup> Penpun Wetwitayaklung,<sup>a,b</sup> Freya Klepsch,<sup>a</sup> Zahida Parveen,<sup>c</sup> Peter Chiba<sup>c</sup> and Gerhard F. Ecker<sup>\*a</sup>

<sup>a</sup>Department of Medicinal Chemistry, University of Vienna, Althanstraße 14, 1090, Vienna, Austria

<sup>b</sup>Department of Pharmacognosy, Faculty of Pharmacy, Silpakorn University, Nakhon Pathom, Thailand

<sup>c</sup>Medical University of Vienna, Institute of Medical Chemistry Waehringerstraße 10, 1090, Vienna, Austria  
E-mail: gerhard.f.ecker@univie.ac.at

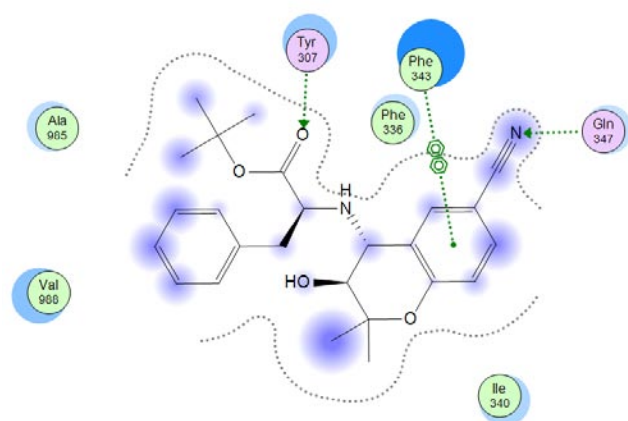
## Supporting Information

|   |   |
|---|---|
| 1. Table1   | 2 |
| 2. Figure1  | 2 |
| 3. Figure2  | 3 |
| 4. Homology Modeling and Docking  | 3 |
| 5. Biological Assay   | 3 |
| 6. General procedure and spectroscopic data of enantiomeric pure (S,S), and (R,R)-epioxides (4) | 4 |
| 7. General procedure and spectroscopic data of L-amino acid-tert-Butyl Esters (5-7)             | 4 |
| 8. General procedure for cyclisation and spectral data of target compounds (11-13)              | 5 |
| 9. <sup>1</sup> H- and <sup>13</sup> C-NMR spectras of target compounds (11a-13b)               | 7 |

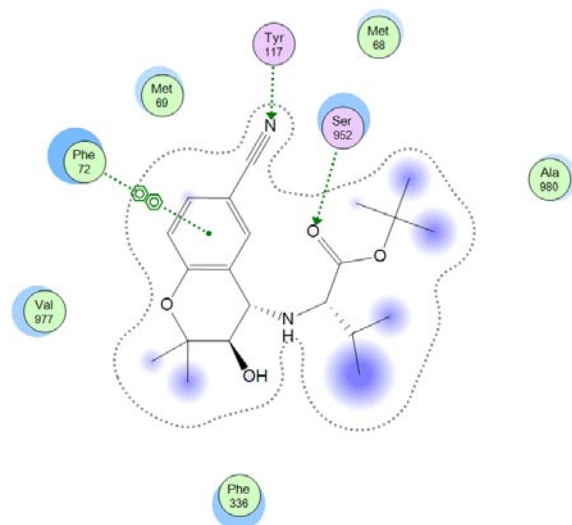
1. Table1 . Number of clusters obtained in common scaffold clustering in one run, in separate runs and the interacting amino acid residues

| # Compounds    | (# of clusters) Clustering in one run | (# of clusters) clustering in separate runs | Interacting amino acid residue       |
|----------------|---------------------------------------|---|--------------------------------------|
| 5a; 6a; 7a     | 2                                     | 4   | Try307,Tyr310,Phe343, Phe336, Gln347 |
| 5a; 6a; 7a     | 1                                     | 1   | Phe951, Ser952, Cys956, Met69        |
| 5b; 6b; 7b     | 4                                     | 3   | Try307,Tyr310,Phe343, Phe336, Gln347 |
| 5b; 6b; 7b     | 1                                     | 1   | Tyr117, Ser952, Phe72, Met69         |
| 5a,b; 6a,b; 7b | 1                                     | 1   | Try307,Tyr310,Phe343, Phe336, Gln347 |
| 11a; 12a; 13a  | 4                                     | 7   | Tyr307, Phe343, Ala342, Phe303       |
| 11b; 12b; 13b  | 7                                     | 7   | Tyr307, Phe343, Ala342, Phe303       |
| 11b; 12b; 13b  | 1                                     | 1   | Ala985, Ile765, Leu724               |

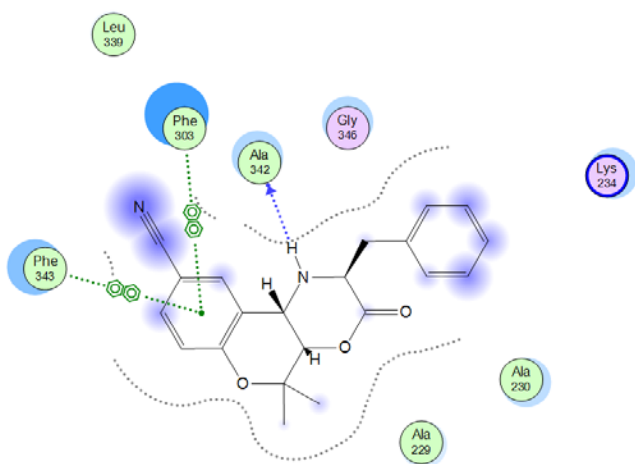
2. Figure 1. Ligand protein interaction of the selected docking poses in different positions.



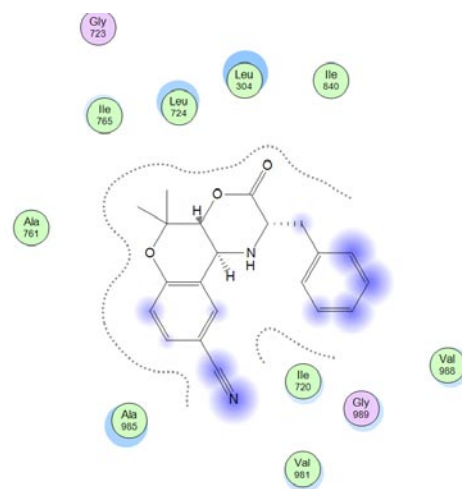
7a



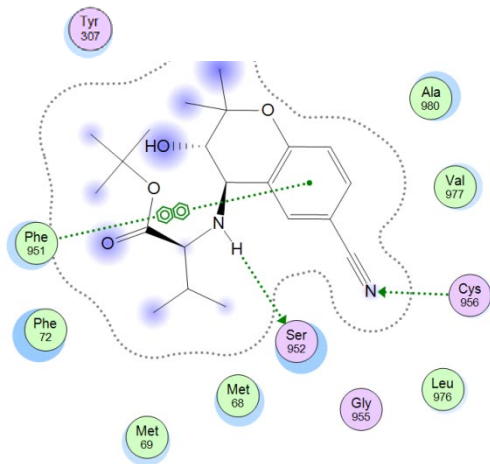
6b



13b (Near entry gate)

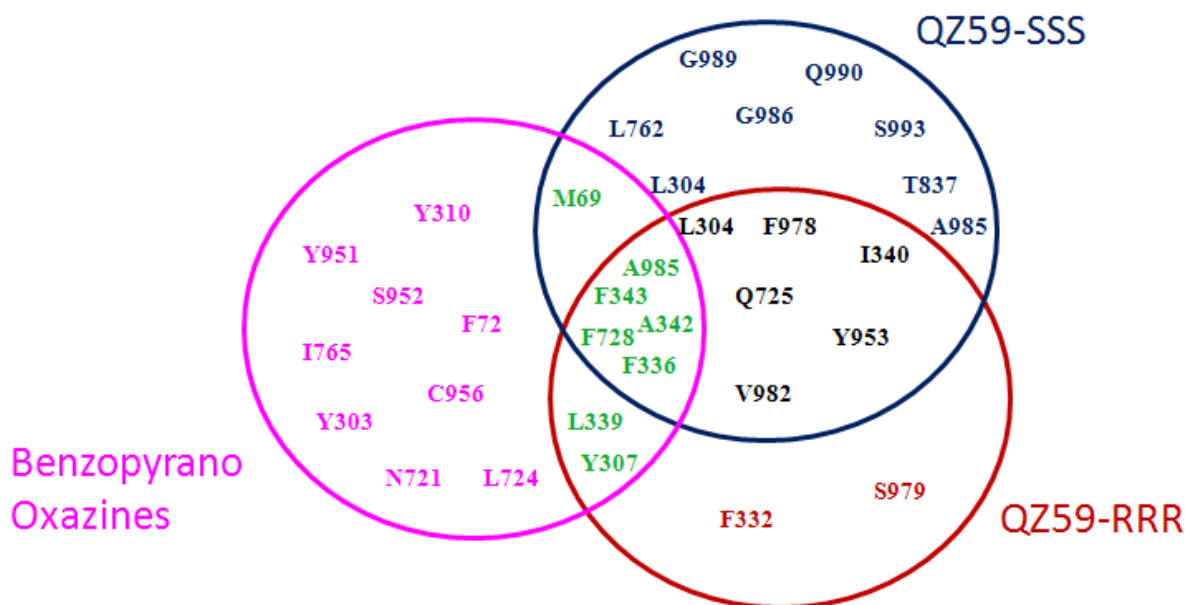


13b (Deeper inside the membrane)



6a

3. Figure 2 Comparison of the main positioning of the benzopyrano[3,4-b][1,4]oxazines with those of the both stereoisomers of cocrystallised tetrapeptides.



#### 4. Homology Modeling and Docking

The homology model was generated with the modeling program Modeller 9v7<sup>1</sup> based on the sequence alignment suggested by Aller et al<sup>2</sup>. Using the automodel procedure 100 different homology models were created and refined. For a slight correction of the distorted TM helix 12 a secondary structure constraint was put on residues 885 – 918. The final model was selected on basis of the smallest number of outliers and a high DOPE score and was evaluated with the program PROCHECK<sup>3</sup>. The Ramachandran plot showed that 84.6 % of the residues lie in most favored, 12.5 % in additional allowed, 2.1 % in generously allowed regions and 0.8 % in disallowed regions. The 2.9 % of the residues that are in generously allowed or disallowed regions are located in the nucleotide binding domains (NBD) or extracellular loops (ECL) and are therefore not involved in drug binding.

Compounds were docked into the homology model of human P-gp by using the software package GOLD, creating 100 poses per ligand. The binding site was defined as covering the complete transmembrane region, which leads to distribution of poses in a large area. Ligand protein complexes were minimized by the LigX graphical interface implemented in MOE by using the MMFF94 force field.

5. **Biological Assay.** The human T-lymphoblast cell line CCRF-CEM and the multidrug resistant CEM/vcr1000 cell line were provided by V. Gekeler (Byk Gulden, Konstanz, Germany). The resistant CEM/vcr1000 line was obtained by stepwise selection in vincristine containing medium. Cells were kept under standard culture conditions (RPMI1640 medium supplemented with 10% fetal calf serum). The P-gp-expressing resistant cell line was cultured in presence of 1000ng/ml vincristine. One week prior to the experiments cells were transferred into medium without selective agents or antibiotics.

Briefly, cells were pelleted, the supernatant was removed by aspiration and cells were resuspended at a density of 1 x 10<sup>6</sup>/ml in RPMI1640 medium containing 3µmol/l daunomycin. Cell suspensions were incubated at 37°C for 30min. After this time a steady state of daunorubicin accumulation was reached. Tubes were chilled on ice and cells were pelleted at 500 x g. Cells were washed once in RPMI1640 medium to remove extracellular daunorubicin. Subsequently, cells were

resuspended in medium prewarmed to 37°C, containing either no modulator or chemosensitizer at various concentrations ranging from 3nM to 500 µM, depending on solubility and expected potency of the modifier. Generally, 8 serial dilutions were tested for each modulator. After 1, 2, 3 and 4 min aliquots of the incubation mixture were drawn and pipetted into 4 volumes of ice cold stop solution (RPMI1640 medium containing verapamil at a final concentration of 100µM). Parental CCRF-CEM cells were used to correct for simple membrane diffusion, which was less than 3% of the efflux rates observed in resistant cells. Samples drawn at the respective time points were kept in an ice water bath and measured within one hour on a Becton Dickinson FACSCalibur (Becton Dickinson, Heidelberg, Germany) flow cytometer as described. Dose response curves were fitted to the data points using non-linear least squares and EC50 values were calculated as described<sup>1</sup>. EC<sub>50</sub> values of individual compounds are the average of at least triplicate determinations. A cv of below 20% was obtained in all determinations.

6. **General procedure for the enantiomerically pure (S,S) (4a) and (R,R)- epoxide (4b).** Commercial household bleach (DanKlorix<sup>®</sup>) was buffered to pH 11.3 with 0.05 N Na<sub>2</sub>HPO<sub>4</sub> and 1N NaOH and then cooled to 0°C. To 1000 mL of this solution a solution of **3** (75.58 mmol) and Mn(III) Salen catalyst (2.74x10<sup>-3</sup> mmol) in 76 mL of CH<sub>2</sub>Cl<sub>2</sub> was added, stirred at 0°C for 5 hr and then at room temperature overnight. The mixture was filtered through Celite and the organic phase was separated, brined once, dried (Na<sub>2</sub>SO<sub>4</sub>) and brought to dryness. Purification by flash chromatography (petroleum ether-ethylacetate; 8:2) yield 76.9% of (S,S)-**4a** and 78.9% of (R,R)-**4b** and as colourless crystals; mp 133-135°C; IR (KBr): 2227 (CN) cm<sup>-1</sup>, 1280 (epoxide) cm<sup>-1</sup>; δ<sub>H</sub> (200MHz; CDCl<sub>3</sub>) 1.28 (s, 3H, CH<sub>3</sub>), 1.57(s, 3H, CH<sub>3</sub>), 1.57 (d, 1H, J = 4.52 Hz, 3-H/4-H), 3.89 (d, 1H, J = 4.52 Hz, 3-H/4-H), 6.84 (d, 1H, J = 8.53 Hz, 8-H), 7.51 (dd, 1H, J = 2.00 Hz, J = 8.41 Hz, 7-H), 7.63 (d, 1H, J = 2.01 Hz, 5-H); δ<sub>C</sub> (CDCl<sub>3</sub>) 22.99(CH<sub>3</sub>), 25.46 (CH<sub>3</sub>), 49.34 (3-C), 62.27 (4-C), 74.64 (2-C), 104.27 (6-C), 118.70 (CN), 119.00 (8-C), 121.67 (4a-C), 133.77, 134.38 (5-C, 7-C), 156.45 (8a-C).
7. **General procedure for L-amino acid-tert-Butyl Ester (5-7).** A solution of enantiomeric pure epoxide **4a** or **4b** (4.97 mmol) and corresponding L-amino acid-tert-butyl ester (5.47 mmol) in 50 mL 96% ethanol was stirred at 80°C for 5 days, then evaporated in vacuo. Purification by flash chromatography (petroleum ether/ethylacetate = 8/2) yield respective L-amino acid t-butyl ester (**5-7**).

**(3S,4R)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-alanine-tert-butyl-Ester (5a).** (S,S)-epoxide **4a** and L-alanine-tert-Butyl ester gave **5a** yield 67% as yellowish oil; mp 145-147 °C; IR (KBr): 2225 (CN) cm<sup>-1</sup>, 1724 (COOR) cm<sup>-1</sup>; δ<sub>H</sub> (200MHz; CDCl<sub>3</sub>) 1.20 (s, 3H, 2-CH<sub>3</sub>), 1.37 (d, 3H, J = 7.03 Hz, CHCH<sub>3</sub>), 1.41 (s, 9H, C(CH<sub>3</sub>)<sub>3</sub>), 1.43 (s, 3H, 2-CH<sub>3</sub>), 1.85 (br, 1H, NH), 3.37 (d, 1H, J = 10.04 Hz, 3-H), 3.44 (d, 1H, J = 10.29 Hz, 4-H), 3.58 (q, 1H, J = 7.03 Hz, N-CH-COO), 4.60 (s, 1H, OH), 6.72 (d, 1H, J = 8.53 Hz, 8-H), 7.31 (dd, 1H, J = 2.01 Hz, J = 8.53 Hz, 7-H), 7.54 (d, 1H, J = 1.75 Hz, 5-H); δ<sub>C</sub> (CDCl<sub>3</sub>) 18.80, 20.31 (2x2-CH<sub>3</sub>), 26.72 (CHCH<sub>3</sub>), 27.82 (C(CH<sub>3</sub>)<sub>3</sub>), 56.75 (N-CH-CO, 4-C), 73.83 (3-C), 79.86 (C(CH<sub>3</sub>)<sub>3</sub>), 82.20 (2-C), 103.41 (6-C), 117.93 (8-C), 119.22 (CN), 126.31 (4a-C), 131.80, 132.38 (5-C, 7-C), 156.49 (8a-C), 176.66 (C=O); Ms m/z 347 (0.32, M<sup>+</sup>), 160 (29.7); [α]<sub>D</sub><sup>20</sup> +5.32 (c 0.141, in CH<sub>2</sub>Cl<sub>2</sub>); (Found: C, 65.69; H, 7.34; N, 8.20. C<sub>19</sub>H<sub>26</sub>N<sub>2</sub>O<sub>4</sub> requires C, 65.88; H, 7.56; N, 8.09).

**(3R,4S)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-alanine-tert-butyl-Ester (5b).** (R,R)-epoxide **4b** and L-alanine-tert-Butyl ester gave **5b** yield 56% as yellowish oil; mp 101-102 °C; IR (KBr) 2226 (CN) cm<sup>-1</sup>, 1722 (COOR) cm<sup>-1</sup>; δ<sub>H</sub> (200MHz; CDCl<sub>3</sub>) 1.18 (s, 3H, 2-CH<sub>3</sub>), 1.28 (d, 3H, J = 7.38 Hz, CHCH<sub>3</sub>), 1.43 (s, 9H, C(CH<sub>3</sub>)<sub>3</sub>), 1.46 (s, 3H, 2-CH<sub>3</sub>) 2.98 (q, 1H, J = 7.28 Hz, CHCH<sub>3</sub>), 3.13 (dd, 1H, J = 3.51 Hz, J = 9.92 Hz, 3-H), 3.78 (d, 1H, J = 9.79 Hz, 4-H), 4.03 (d, 1H, J = 3.51 Hz, OH), 6.81 (d, 1H, J = 8.54 Hz, 8-H), 7.40 (dd, 1H, J = 1.88 Hz, J = 8.47 Hz, 7-H), 7.87 (d, 1H, J = 1.88 Hz, 5-H); δ<sub>C</sub> (CDCl<sub>3</sub>) 19.11 (2-CH<sub>3</sub>), 21.25 (2-CH<sub>3</sub>), 27.19 (CHCH<sub>3</sub>), 27.78 (C(CH<sub>3</sub>)<sub>3</sub>), 51.18 (N-CH-CO), 56.42 (4-C), 70.86 (3-C), 79.68 (C(CH<sub>3</sub>)<sub>3</sub>), 82.35 (2-C), 103.78 (6-C), 118.21 (8-C), 119.48 (CN), 123.90 (4a-C), 132.38, 132.52 (5-C,7-C), 157.77 (8a-C), 178.79 (C=O); MS m/z 347 (0.12, M<sup>+</sup>), 160 (24.7); [α]<sub>D</sub><sup>20</sup> -188.03 (c 0.117, in CH<sub>2</sub>Cl<sub>2</sub>); (Found: C, 66.13; H, 7.78; N, 7.88. C<sub>19</sub>H<sub>26</sub>N<sub>2</sub>O<sub>3</sub> requires C, 88; H, 7.56; N, 8.09)

**(3S,4R)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-valine-tert-butyl-Ester (6a).** (S,S)-epoxide **4a** and L-valine-tert-Butyl ester gave **6a** yield 65% yellowish oil; mp 128-130 °C; IR (KBr): 2226 (CN) cm<sup>-1</sup>, 1721 (CO), 3478 (OH) cm<sup>-1</sup>; δ<sub>H</sub> (200MHz; CDCl<sub>3</sub>) 0.91, 1.05 (each d, each 3H, each J= 6.77 Hz, CH(CH<sub>3</sub>)<sub>2</sub>), 1.13 (s, 3H, 2-CH<sub>3</sub>), 1.43 (s, 12H, CH<sub>3</sub>, C(CH<sub>3</sub>)<sub>3</sub>), 1.85 (br, 1H, NH), 2.01-2.15 (m, 1H, CH(CH<sub>3</sub>)<sub>2</sub>), 3.36 (d, 1H, J = 4.27 Hz, N-CH-CO), 3.42 (br, 2H, 3-H, 4-H), 4.16 (s, 1H, OH), 6.74 (d, 1H, J= 8.53 Hz, 8-H), 7.33 (dd, 1H, J = 8.53/1.88 Hz, 7-H), 7.69 (d, 1H, J = 1.88 Hz, 5-H); δ<sub>C</sub> (CDCl<sub>3</sub>) 17.71 (CH-CH<sub>3</sub>), 18.82 (2-CH<sub>3</sub>), 19.51 (CH-CH<sub>3</sub>), 26.72 (2-CH<sub>3</sub>), 27.93 (C(CH<sub>3</sub>)<sub>3</sub>), 32.58 (CH(CH<sub>3</sub>)<sub>2</sub>), 56.72 (4-C), 66.58 (N-CH-CO), 74.28 (3-C), 79.74 (C(CH<sub>3</sub>)<sub>3</sub>), 82.16 (2-C), 103.38 (6-C), 117.92 (8-C), 119.28(CN), 126.40 (4a-C), 132.31 (5-C, 7-C), 156.49(8a-C) 176.02(C=O); Ms m/z 375 (0.79, M<sup>+</sup>), 273 (57.1%) 160 (40.9%); [α]<sub>D</sub><sup>20</sup> -30.00 (c 0.113, in CH<sub>2</sub>Cl<sub>2</sub>); (Found: C, 67.33; H, 7.98; N, 7.26. C<sub>21</sub>H<sub>30</sub>N<sub>2</sub>O<sub>4</sub> requires C, 67.35; H, 8.07; N, 7.48)



**(3R,4S)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-valine-tert-butyl-Ester (6b).** (R,R)-epoxide **4b** and L-valine-tert-Butyl ester gave **6b** yield 60% yellowish oil; mp 79-81°C; IR (KBr): 2226 (CN)  $\text{cm}^{-1}$ , 1703 (C=O), 3493 (OH)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 0.98, 0.99 (each d, each 3H, each  $J = 6.77$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.17 (s, 3H, 2- $\text{CH}_3$ ), 1.44 (s, 12H,  $\text{CH}_3$ ,  $\text{C}(\text{CH}_3)_3$ ), 1.86-2.01 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 2.38 (br, 1H, NH), 2.76 (d, 1H,  $J = 4.26$  Hz, N-CH-CO), 3.15 (dd, 1H,  $J = 9.54/3.76$  Hz, 3-H), 3.78 (d, 1H,  $J = 9.54$  Hz, 4-H), 3.83 (d, 1H,  $J = 3.76$  Hz, OH), 6.80 (d, 1H,  $J = 8.28$  Hz, 8-H), 7.39 (dd, 1H,  $J = 8.28/1.75$  Hz, 7-H), 7.94 (d, 1H,  $J = 1.75$  Hz, 5-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 17.66 ( $\text{CH}-\text{CH}_3$ ), 19.06 (2- $\text{CH}_3$ ), 19.76 ( $\text{CH}-\text{CH}_3$ ), 27.02 (2- $\text{CH}_3$ ), 27.82 ( $\text{C}(\text{CH}_3)_3$ ), 32.63 ( $\text{CH}(\text{CH}_3)_2$ ), 56.54 (4-C), 60.55 (N-CH-CO), 71.01 (3-C), 79.58 ( $\text{C}(\text{CH}_3)_3$ ), 82.31 (2-C), 103.58 (6-C), 118.18 (8-C), 119.32(CN), 123.82 (4a-C), 132.40, 133.39 (5-C, 7-C), 157.75(8a-C) 177.94(C=O); Ms  $m/z$  375 (0.41%,  $\text{M}^+$ ), 273 (16%) 72 (100%);  $[\alpha]_{\text{D}}^{20} -193.80$  (c 0.129, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 67.58; H, 8.01; N, 7.25.  $\text{C}_{21}\text{H}_{30}\text{N}_2\text{O}_4$  requires C, 67.35; H, 8.07; N, 7.48)

**(3S,4R)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-phenylalanine-tert-butyl-Ester (7a).** (S,S)-epoxide **4a** and L-phenylalanine-tert-Butyl ester gave **7a** yield 59% yellowish oil; mp 99-101°C; IR (KBr): 2225 (CN)  $\text{cm}^{-1}$ , 1722 (C=O), 3441 (OH)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.11 (s, 3H, 2- $\text{CH}_3$ ), 1.43 (s, 3H, 2- $\text{CH}_3$ ), 1.48 (s, 9H,  $\text{C}(\text{CH}_3)_3$ ), 1.71 (br, 1H, NH), 2.74 (dd, 1H,  $J = 12.70/10.36$  Hz, benzyl Ha), 3.22 (dd, 1H,  $J = 12.70/3.53$  Hz, benzyl Hb), 3.31 (br, 2H, N-CH-CO, 3-H), 3.77 (dd, 1H,  $J = 9.79/3.41$  Hz, 4-H), 4.34 (br, 1H, OH), 6.68 (d, 1H,  $J = 8.46$  Hz, 8-H), 6.88 (s, 1H, 5-H), 7.28-7.40 (m, 6-H, 7-H, phenyl-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 18.76 (2- $\text{CH}_3$ ), 26.74 (2- $\text{CH}_3$ ), 27.89 ( $\text{C}(\text{CH}_3)_3$ ), 40.26 (benzyl  $\text{CH}_2$ ), 56.90 (4-C), 63.40 (N-CH-CO), 74.09 (3-C), 79.74 ( $\text{C}(\text{CH}_3)_3$ ), 82.53 (2-C), 103.38 (6-C), 103.38 (6-C), 117.79 (8-C), 118.89 (CN), 126.00 (4a-C), 127.66 (arm. CH), 129.01 (2x arom.CH), 129.06 (2x arom. CH), 131.25, 132.43 (5-C, 7-C), 137.05(arom.C), 156.39 (8a-C), 175.50 (C=O); Ms  $m/z$  423 (1.95%,  $\text{M}^+$ ), 275 (47.6%), 160 (89%);  $[\alpha]_{\text{D}}^{20} +0.90$  (c 0.111, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 70.35; H, 7.09; N, 6.72.  $\text{C}_{25}\text{H}_{30}\text{N}_2\text{O}_4$  requires C, 71.07; H, 7.16; N, 6.63)

**(3R,4S)-N-(6-cyano-3-hydroxy-2,2-dimethyl-2H-3,4-dihydro-1-benzopyran-4yl)-L-phenylalanine-tert-butyl-Ester (7b).** (R, R)-epoxide **4b** and L-phenylalanine-tert-Butyl ester gave **7b** yield 53% yellowish oil; mp 94-96 °C; IR (KBr): 2229 (CN)  $\text{cm}^{-1}$ , 1695 (C=O)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.20 (s, 3H, 2- $\text{CH}_3$ ), 1.52 (s, 12H, 2- $\text{CH}_3$ ,  $\text{C}(\text{CH}_3)_3$ ), 2.40 (br, 1H, NH), 2.78 (dd, 1H,  $J = 13.27/8.84$  Hz, benzyl Ha), 3.05 (dd, 1H,  $J = 13.27/4.17$ Hz, benzyl Hb), 3.18 (m, 2H, N-CH-CO, 3-H), 3.79 (d, 1H,  $J = 9.98$  Hz, 4-H), 3.93 (d, 1H,  $J = 3.41$  Hz, OH), 6.84 (d, 1H,  $J = 8.59$ Hz, 8-H), 7.12(s, 1H, 5-H), 7.27(m,1H, 7-H), 7.33-7.46(m, 5H, phenyl-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 18.76 (2- $\text{CH}_3$ ), 27.12 (2- $\text{CH}_3$ ), 27.77 ( $\text{C}(\text{CH}_3)_3$ ), 41.03 (benzyl  $\text{CH}_2$ ), 56.44 (4-C), 57.27 (N-CH-CO), 71.00 (3-C), 79.54 ( $\text{C}(\text{CH}_3)_3$ ), 82.66 (2-C), 103.77 (6-C), 118.07 (8-C), 119.28 (CN), 123.32 (4a-C), 127.63 (arm. CH), 128.59 (2x arom. CH), 129.50 (2x arom.CH), 132.38, 132.63 (5-C, 7-C), 136.81(arom.C), 157.64 (8a-C), 177.66 (C=O); Ms  $m/z$  423 (2.44%,  $\text{M}^+$ ), 160 (59.5%), 120 (100%);  $[\alpha]_{\text{D}}^{20} -117.5$  (c 0.12, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 70.88; H, 7.21; N, 6.56.  $\text{C}_{25}\text{H}_{30}\text{N}_2\text{O}_4$  requires C, 71.07; H, 7.16; N, 6.63)

8. **General procedure for cyclisation.** (4.61 mmol) of L-amino acid-tert-butyl ester (**5-7**) was dissolved in a small amount of  $\text{CH}_2\text{Cl}_2$ , hydrolysed by 6 mL of 70%  $\text{HClO}_4$ , stirred overnight, and 4N  $\text{NH}_4\text{OH}$  solution was added slowly. The precipitate was dried and used in the next reaction step without further purification. A suspension of precipitates (2.76 mmol), 4-dimethylaminopyridine (0.69 mmol) and bis (2-oxo-3-oxazolidinyl) phosphinic chloride (4.12 mmol) in  $\text{CH}_2\text{Cl}_2$  (50 mL) was heated to reflux at 80°C for 10 min, then triethylamine (0.95 mL, 6.85 mmol) was added and the solution was refluxed at 70°C for 4 days. The suspension was filtered and evaporated to dryness. Purification was done by flash chromatography (petroleum ether/ethylacetate; 9:1) to yield the target compounds (**11-13**).

**(2S,4aS,10bR)-2,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (11a)** yield 52% as colourless crystals; mp 157-158°C; IR (KBr): 2224 (CN)  $\text{cm}^{-1}$ ; 1751 (lactone)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.31 (s, 3H, 5- $\text{CH}_3$ ), 1.52 (s, 3H, 5- $\text{CH}_3$ ), 1.56 (d, 3H,  $J = 7.08$  Hz, 2- $\text{CH}_3$ ), 3.86-4.03 (m, 2H, N-CH-CO, 4a-H), 4.14 (d, 1H,  $J = 10.30$  Hz, 10b-H), 6.87 (d, 1H,  $J = 8.53$  Hz, 7-H), 7.46 (dd, 1H,  $J = 8.53/2.02$  Hz, 8-H), 7.84 (d, 1H,  $J = 2.02$  Hz, 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 18.28 (5- $\text{CH}_3$ ), 19.86 (5-C), 25.88 (2- $\text{CH}_3$ ), 49.63 (10b-C), 54.21 (N-CH-CO), 78.07(5-C), 83.87 (4a-C), 104.25 (9-C), 118.24 (7-C), 118.83 (CN), 121.89 (10a-C), 131.22, 133.45 (8-C, 10C), 155.59 (6a-C), 169.98 (C=O); MS  $m/z$  272 (3.37%,  $\text{M}^+$ ), 185 (39.6%), 170 (100%);  $[\alpha]_{\text{D}}^{20} +95.34$  (c 0.118, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 65.94; H, 5.98; N,10.15.  $\text{C}_{15}\text{H}_{16}\text{N}_2\text{O}_3$  requires C, 66.16; H, 5.92; N, 10.29)

**(2S,4aR,10bS)-2,5,5-trimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (11b)** yield 45% as colourless crystals; mp 125-127.5°C; IR (KBr) 2224 (CN)  $\text{cm}^{-1}$ ; 1737 (lactone);  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.31 (s, 3H, 5- $\text{CH}_3$ ), 1.52 (d, 3H,  $J = 7.03$  Hz, 2- $\text{CH}_3$ ), 1.55 (s, 3H, 5- $\text{CH}_3$ ), 3.90 (q, 1H,  $J = 7.03$  Hz, N-CH-CO), 4.05 (d, 1H,  $J = 10.54$  Hz, 4a-H), 4.21 (d, 1H,  $J = 10.29$ , 10b-H), 6.86 (d, 1H,  $J = 8.54$  Hz, 7-H), 7.45 (dd, 1H,  $J = 1.88$  Hz,  $J = 8.54$  Hz, 8-H), 7.77 (d, 1H,  $J = 1.88$ , 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 18.64, 19.93 (2x2- $\text{CH}_3$ ), 26.19 (2- $\text{CH}_3$ ), 47.36 (10b-C), 50.88 (N-CH-CO), 77.93 (5-C), 81.08 (4a-C), 104.33 (9-C), 118.07 (7-C), 118.88 (CN), 122.62 (10a-C), 131.68, 133.25 (10-C, 8-C) 155.66 (6a-C), 171.52 (C=O); MS  $m/z$  273 (0.7,  $\text{M}^+$ ), 185 (35), 170 (100);  $[\alpha]_{\text{D}}^{20} -147.52$  (c 0.101, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 66.27; H, 6.04; N, 10.05  $\text{C}_{15}\text{H}_{16}\text{N}_2\text{O}_3$  requires C, 66.16; H, 5.92; N, 10.29)

**(2S,4aS,10bR)-2-isopropyl-5,5-dimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (12a)** yield 66% as yellowish solid; mp 145-146 °C; IR (KBr) 2220 (CN)  $\text{cm}^{-1}$ ; 1741 (lactone);  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.01, 1.13 (each d, each 3H, each  $J = 6.95$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.30 (s, 3H, 5- $\text{CH}_3$ ), 1.53 (s, 3H, 5- $\text{CH}_3$ ), 2.49-2.64 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 3.87 (d, 1H,  $J = 5.81$  Hz, N-CH-CO), 3.97 (d, 1H,  $J = 9.50$  Hz, 4a-H), 4.06 (d, 1H,  $J = 9.50$  Hz, 10b-H), 6.88 (d, 1H,  $J = 8.58$  Hz, 7-H), 7.48 (dd, 1H,  $J = 8.58/2.02$  Hz, 8-H), 7.83 (s, 1H, 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 17.46 (CH- $\text{CH}_3$ ), 19.05 (5- $\text{CH}_3$ ), 19.84 (CH- $\text{CH}_3$ ), 25.92 (5- $\text{CH}_3$ ), 31.00 ( $\text{CH}(\text{CH}_3)_2$ ), 49.34 (10b-C), 63.77 (N-CH-CO), 78.02 (5-C), 83.16 (4a-C), 104.37 (9-C), 118.32 (7-C), 118.87 (CN), 121.35 (10a-C), 131.31, 133.45 (8-C, 10-C), 156.00 (6a-C), 169.45 (C=O); Ms  $m/z$  300 (0.80%,  $\text{M}^+$ ), 170 (100%);  $[\alpha]_{\text{D}}^{20} +80$  (c 0.05, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 67.78; H, 6.67; N, 9.08.  $\text{C}_{17}\text{H}_{20}\text{N}_2\text{O}_3$  requires C, 68.71; H, 6.71; N, 9.33)

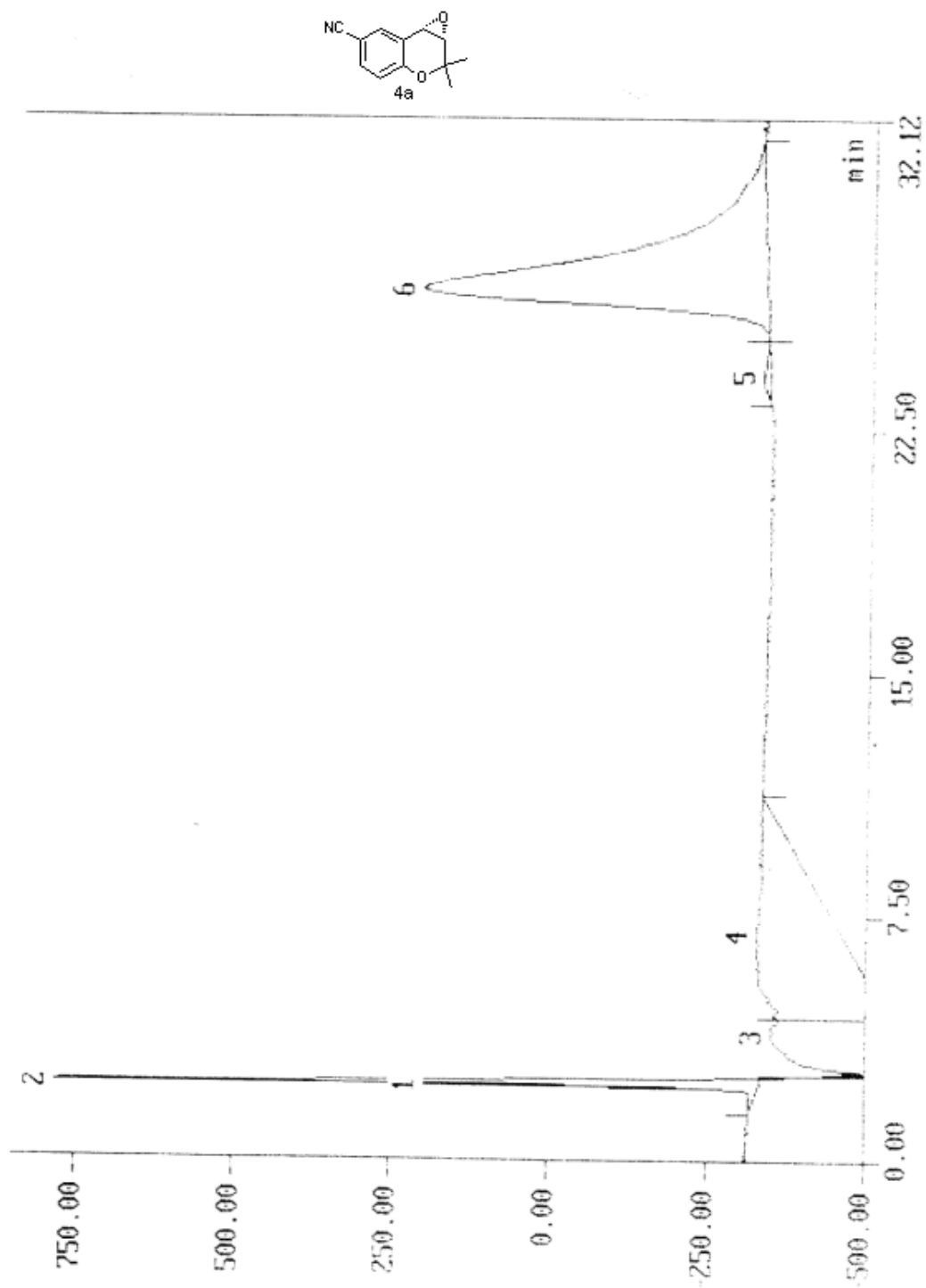
**(2S,4aR,10bS)-2-isopropyl-5,5-dimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (12b)** yield 80% as white crystals; mp 120-123 °C; IR (KBr) 2224 (CN)  $\text{cm}^{-1}$ ; 1745 (lactone), 1575 (NH);  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.02, 1.11 (each d, each 3H, each  $J = 6.82$  Hz,  $\text{CH}(\text{CH}_3)_2$ ), 1.28 (s, 3H, 5- $\text{CH}_3$ ), 1.53 (s, 3H, 5- $\text{CH}_3$ ), 1.66 (br, 1H, NH), 2.38-2.54 (m, 1H,  $\text{CH}(\text{CH}_3)_2$ ), 3.65 (br, 1H, 10b-H), 3.89 (t, 1H,  $J = 9.68$  Hz, N-CH-CO), 4.15 (d, 1H,  $J = 9.68$  Hz, 4a-H), 6.87 (d, 1H,  $J = 8.52$  Hz, 7-H), 7.47 (dd, 1H,  $J = 8.52/1.77$  Hz, 8-H), 7.90 (br, 1H, 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 17.47 (CH- $\text{CH}_3$ ), 19.62, 19.76 (CH- $\text{CH}_3$ , 5- $\text{CH}_3$ ), 26.02 (5- $\text{CH}_3$ ), 31.80 ( $\text{CH}(\text{CH}_3)_2$ ), 49.30 (10b-C), 59.87 (N-CH-CO), 77.89 (5-C), 82.45 (4a-C), 104.38 (9-C), 118.22 (7-C), 118.92 (CN), 122.11 (10a-C), 131.42, 133.39 (8-C, 10-C), 155.79 (6a-C), 170.31 (C=O); Ms,  $m/z$  300 (0.66%,  $\text{M}^+$ ), 185 (39.7%), 170 (92%);  $[\alpha]_{\text{D}}^{20} -116.97$  (c 0.109, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 67.94; H, 6.66; N, 9.32.  $\text{C}_{17}\text{H}_{20}\text{N}_2\text{O}_3$  requires C, 67.98; H, 6.71; N, 9.33)

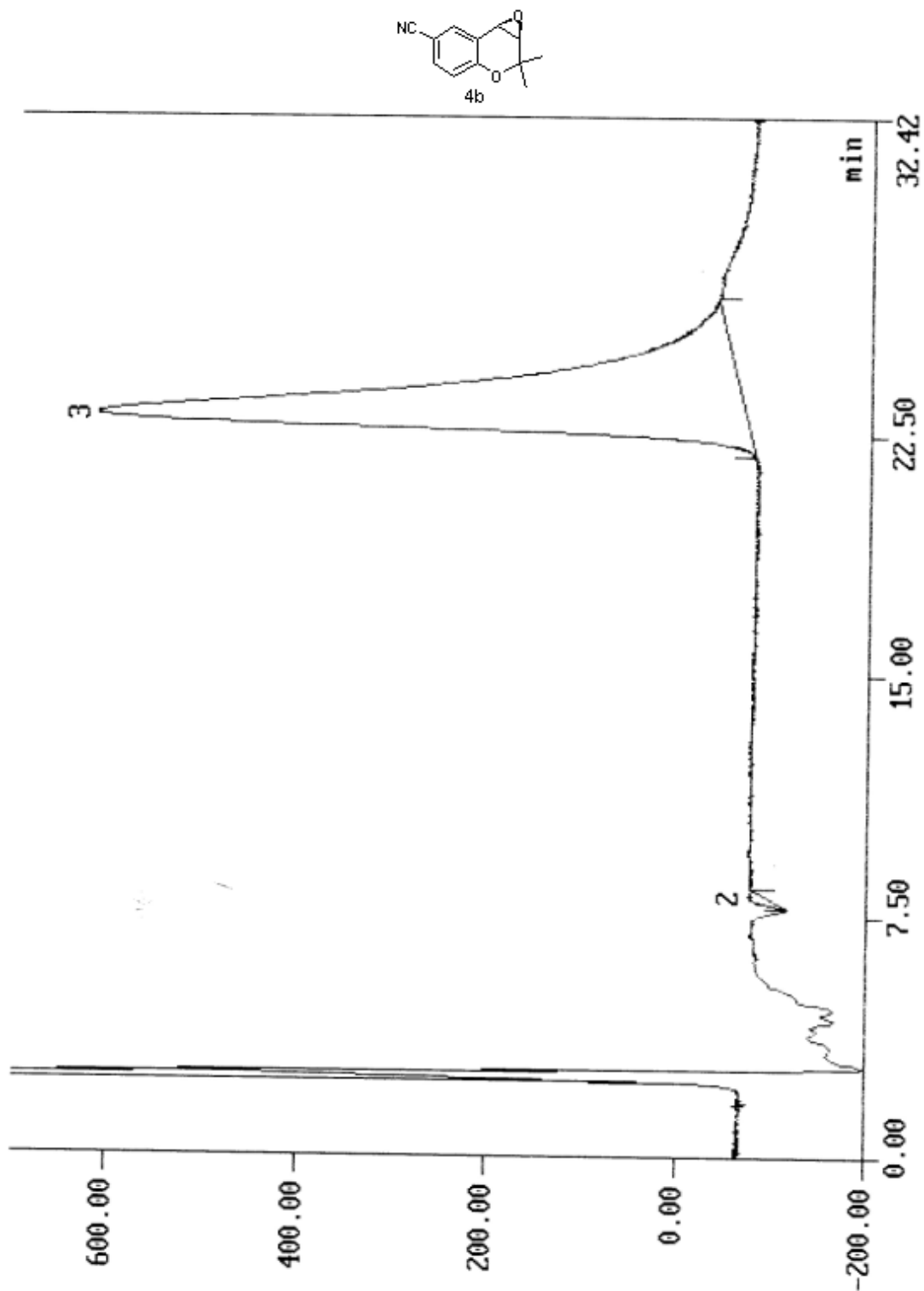
**(2S,4aS,10bR)-2-benzyl-5,5-dimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (13a)** yield (51%) as yellowish solid; mp 140-142 °C; IR (KBr): 2226 (CN)  $\text{cm}^{-1}$ , 1740 (lactone)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.29 (s, 3H, 5- $\text{CH}_3$ ), 1.46 (s, 3H, 5- $\text{CH}_3$ ), 3.23 (dd, 1H,  $J = 14.02/4.67$  Hz, benzyl Ha), 3.47 (dd, 1H,  $J = 14.02/5.69$  Hz, benzyl Hb), 3.69 (d, 1H,  $J = 10.11$  Hz, 4a-H), 3.97 (d, 1H,  $J = 10.11$  Hz, 10b-H), 4.21 (dd, 1H,  $J = 5.69/4.67$  Hz, N-CH-CO), 6.84 (d, 1H,  $J = 8.43$  Hz, 7-H), 7.27-7.43 (m, 5H, phenyl-H), 7.61 (dd, 1H,  $J = 8.43/1.96$  Hz, 8-H), 7.75 (d, 1H,  $J = 1.96$  Hz, 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 19.78 (5- $\text{CH}_3$ ), 25.75 (5- $\text{CH}_3$ ), 37.65 (benzyl  $\text{CH}_2$ ), 49.45 (10b-C), 59.38 (N-CH-CO), 77.85 (5-C), 83.22 (4a-C), 104.26 (9-C), 118.18 (7-C), 118.80 (CN), 121.00 (10a-C), 127.47 (arom.CH), 129.07 (2x arom.CH), 129.37 (2x arom.CH), 131.13, 133.39 (8-C, 10-C), 135.99 (arom.C), 155.84 (6a-C), 169.27 (C=O); Ms  $m/z$  348 (3.05%,  $\text{M}^+$ ), 257 (17%), 170 (70.8%), 91 (100%);  $[\alpha]_{\text{D}}^{20} +133.33$  (c 0.12, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 72.17; H, 5.86, N, 7.96.  $\text{C}_{21}\text{H}_{20}\text{N}_2\text{O}_3$  requires C, 72.40; H, 5.79; N, 8.04)

**(2S,4aR,10bS)-2-benzyl-5,5-dimethyl-3-oxo-1,4a,5,10b-tetrahydro-3-H[1]benzopyrano[3,4-b][1,4]oxazine-9-carbonitril (13b)** yield (45%) as yellowish oil; mp 202-201 °C; IR (KBr): 2223 (CN)  $\text{cm}^{-1}$ , 1744 (lactone)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (200MHz;  $\text{CDCl}_3$ ) 1.05 (s, 3H, 5- $\text{CH}_3$ ), 1.48 (s, 3H, 5- $\text{CH}_3$ ), 1.83 (br, 1H, NH), 3.18 (dd, 1H,  $J = 13.56/6.44$  Hz, benzyl Ha), 3.28 (dd, 1H,  $J = 13.56/4.99$  Hz, benzyl Hb), 3.45 (d, 1H,  $J = 10.11$  Hz, 4a-H), 4.09 (d, 1H,  $J = 10.11$  Hz, 10b-H), 4.11 (dd, 1H,  $J = 6.44/4.99$  Hz, N-CH-CO), 6.81 (d, 1H,  $J = 8.53$  Hz, 7-H), 7.27-7.32 (m, 5H, phenyl-H), 7.42 (dd, 1H,  $J = 8.53/2.02$  Hz, 8-H), 7.72 (br, 1H, 10-H);  $\delta_{\text{C}}$  ( $\text{CDCl}_3$ ) 19.35 (5- $\text{CH}_3$ ), 25.84 (5- $\text{CH}_3$ ), 39.82 (benzyl  $\text{CH}_2$ ), 47.10 (10b-C), 56.77 (N-CH-CO), 77.75 (5-C), 82.58 (4a-C), 104.26 (9-C), 118.03 (7-C), 118.91 (CN), 121.71 (10a-C), 127.08 (arom.CH), 128.59 (2x arom.CH), 129.64 (2x arom.CH), 131.02, 133.23 (8-C, 10-C), 137.26 (arom.C), 155.70 (6a-C), 170.11 (C=O); Ms  $m/z$  348 (2.38%,  $\text{M}^+$ ), 257 (20.5%), 170 (39.40%), 91 (100%);  $[\alpha]_{\text{D}}^{20} -33.65$  (c 0.104, in  $\text{CH}_2\text{Cl}_2$ ); (Found: C, 72.20; H, 5.74; N, 7.87.  $\text{C}_{21}\text{H}_{20}\text{N}_2\text{O}_3$  requires C, 72.40; H, 5.79; N, 8.04)

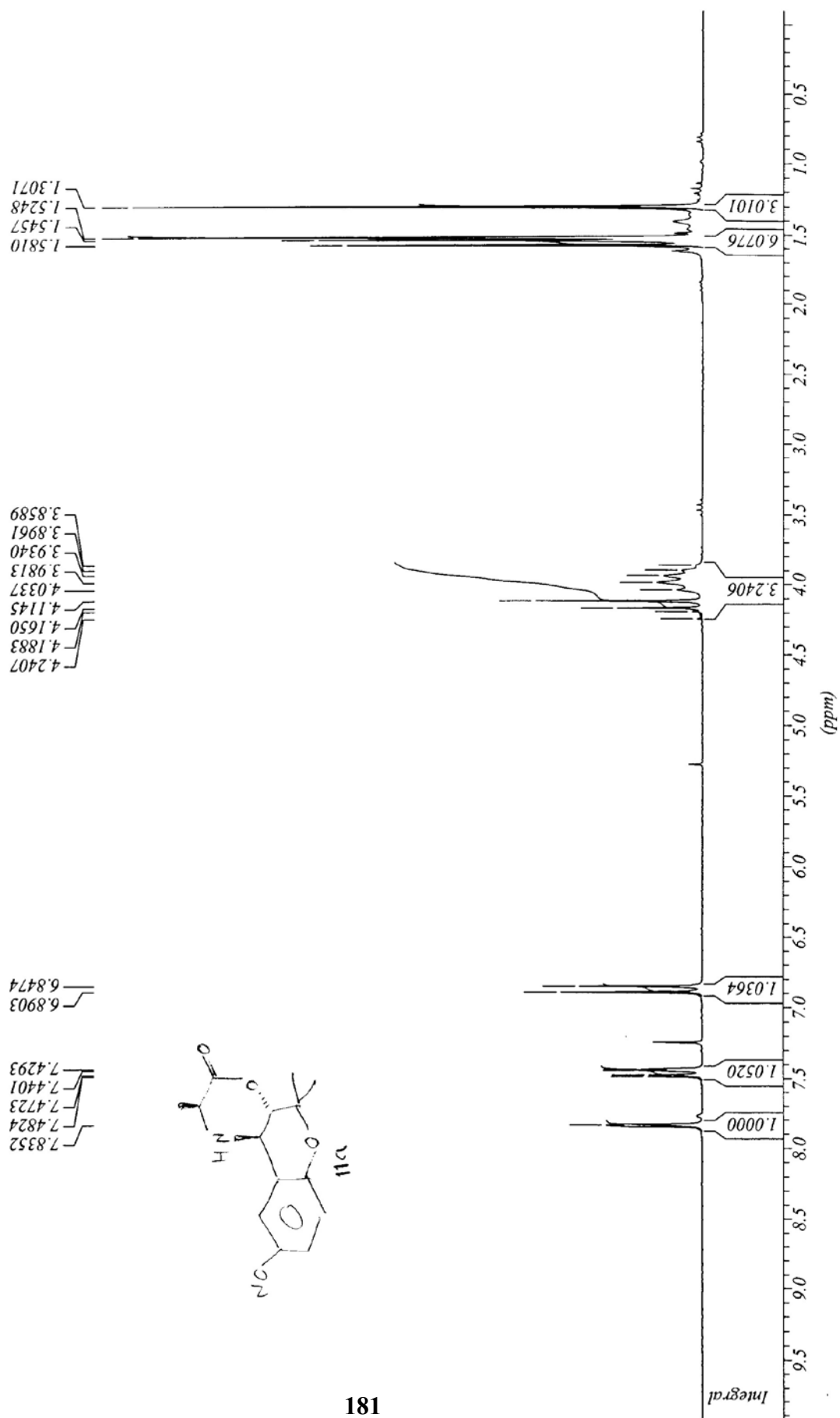
1. P. Chiba, G. Ecker, D. Schmid, J. Drach, B. Tell, S. Goldenberg and V. Gekeler, *Mol Pharmacol*, 1996, **49**, 1122-1130.

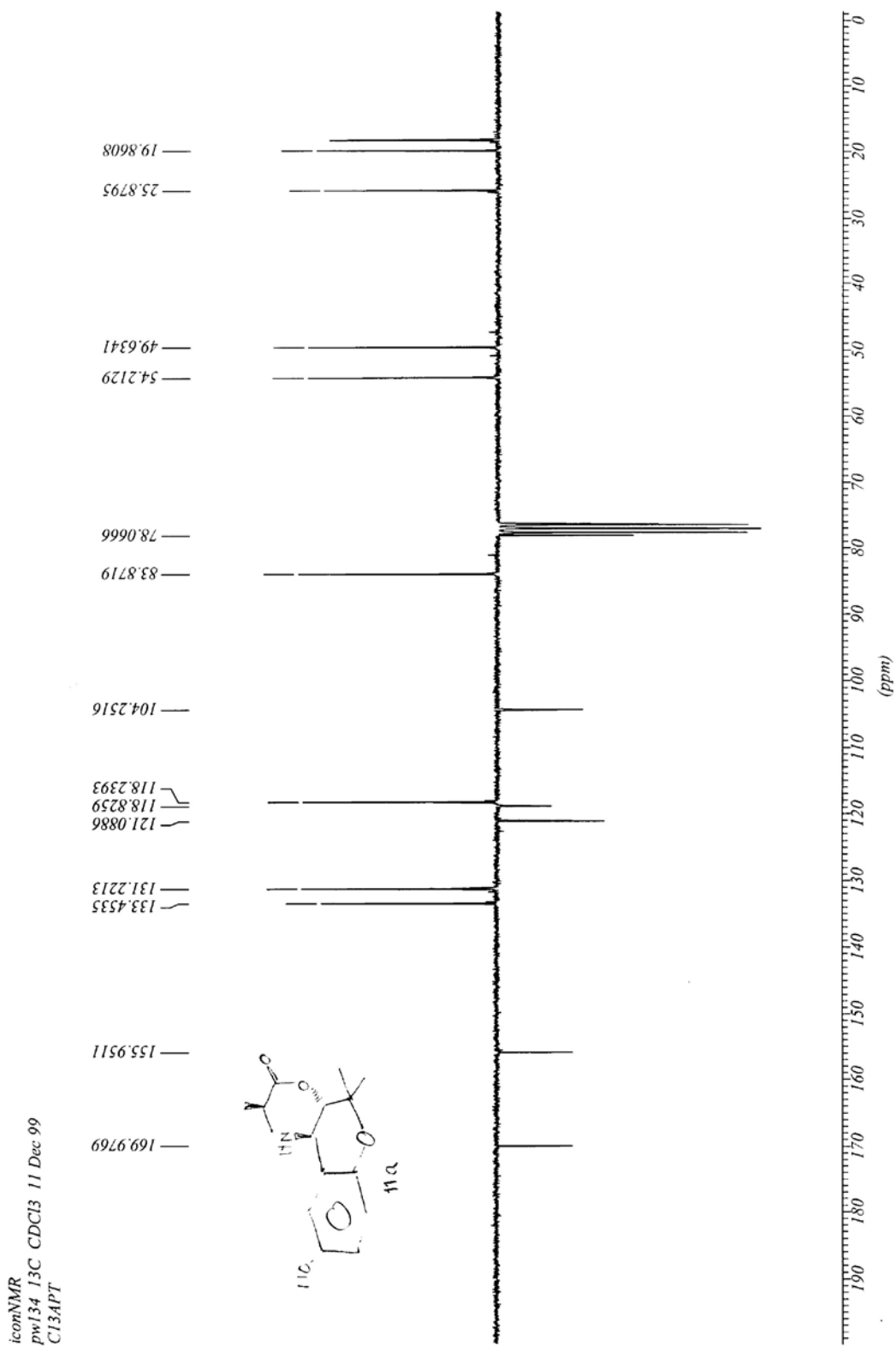
9.  $^1\text{H}$  and  $^{13}\text{C}$ -NMR spectras of target compounds (11a-13b)



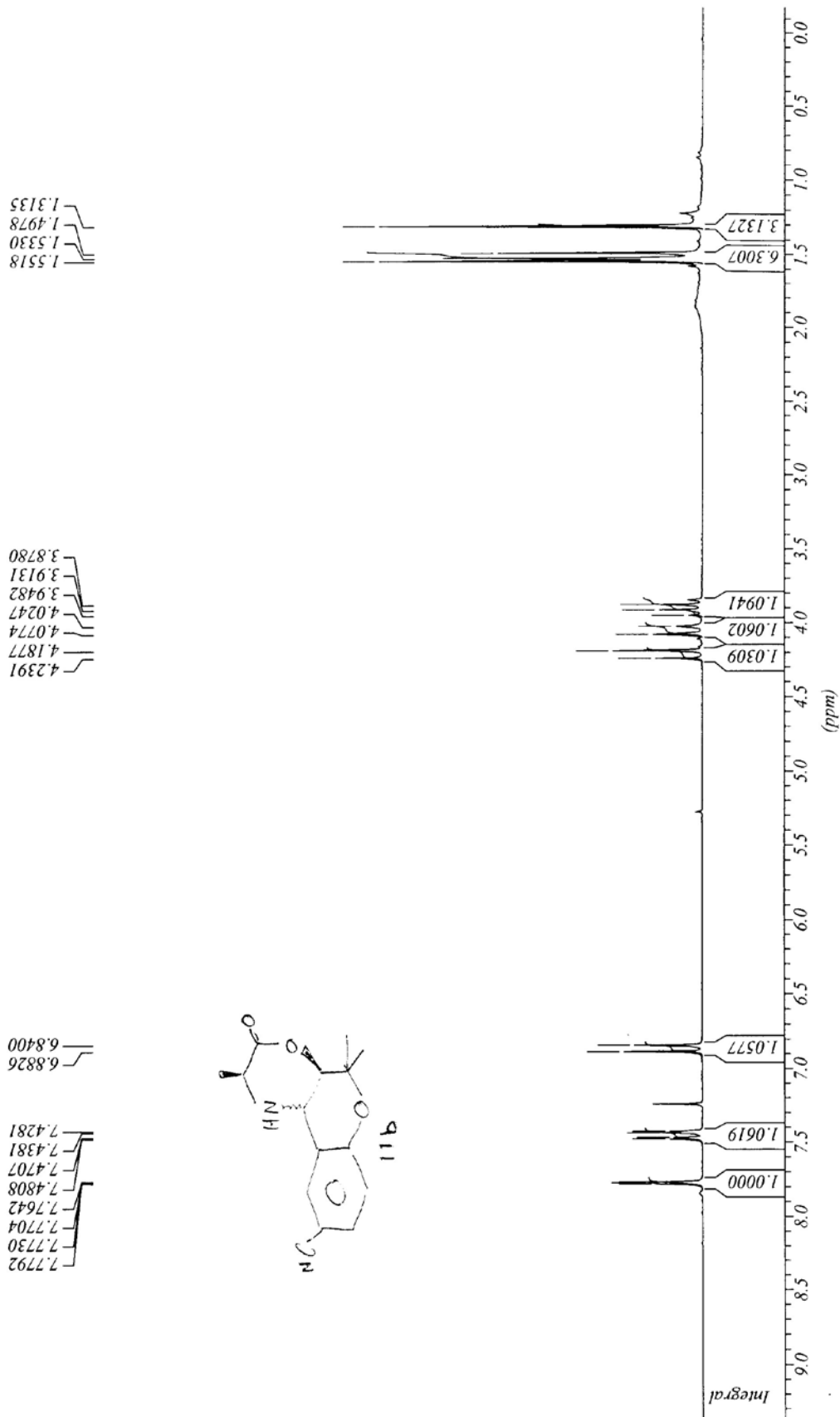


iconNMR  
pw134 1H CDCl3 11 Dec 99  
PROTON

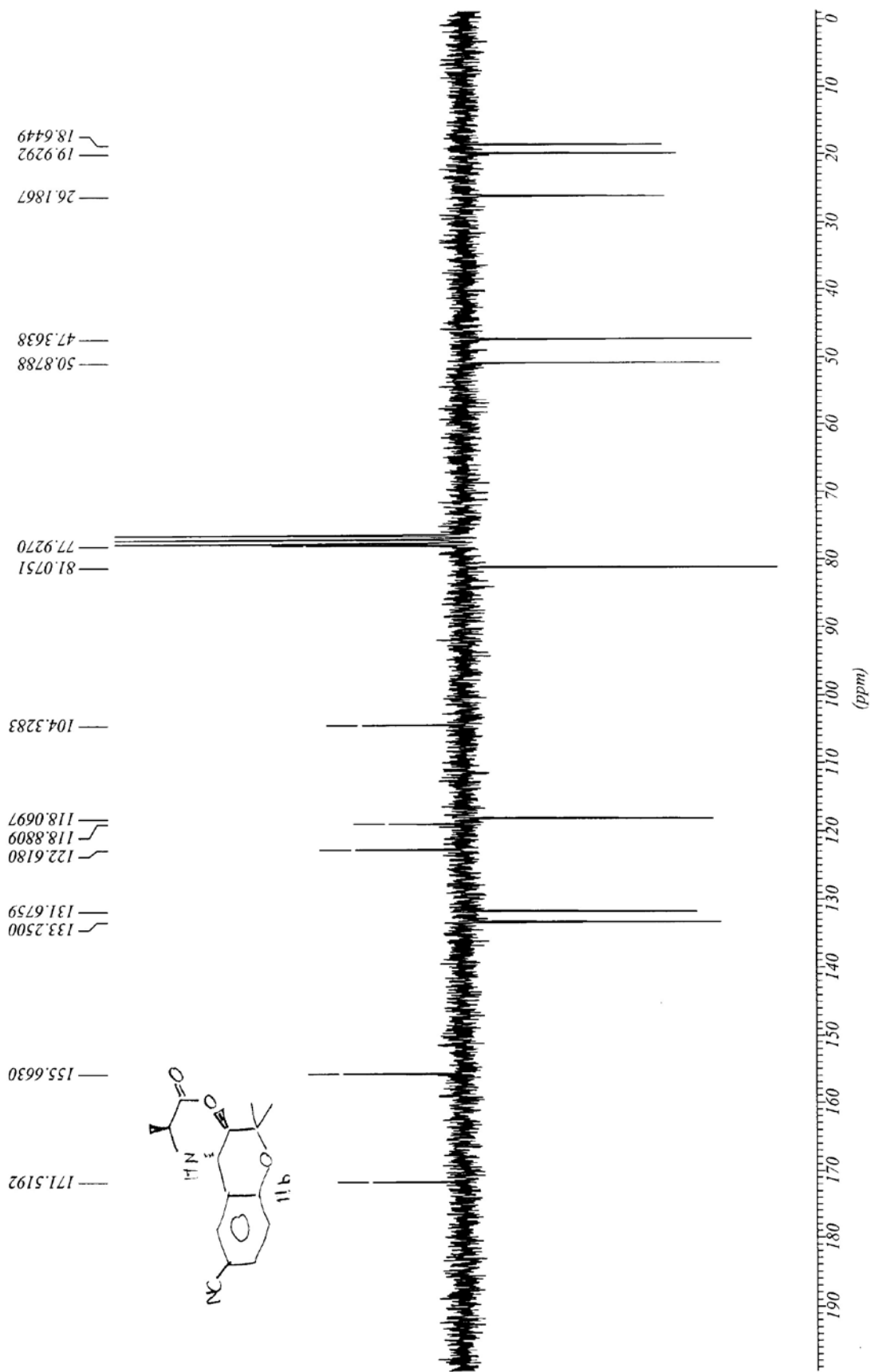




CDJL<sub>2</sub>  
pw106 1H DMSO-d<sub>6</sub> 11 Nov 99



pw106 13C CDC13 11 Nov 99





pw33 1H CDC13 15 Apr 99

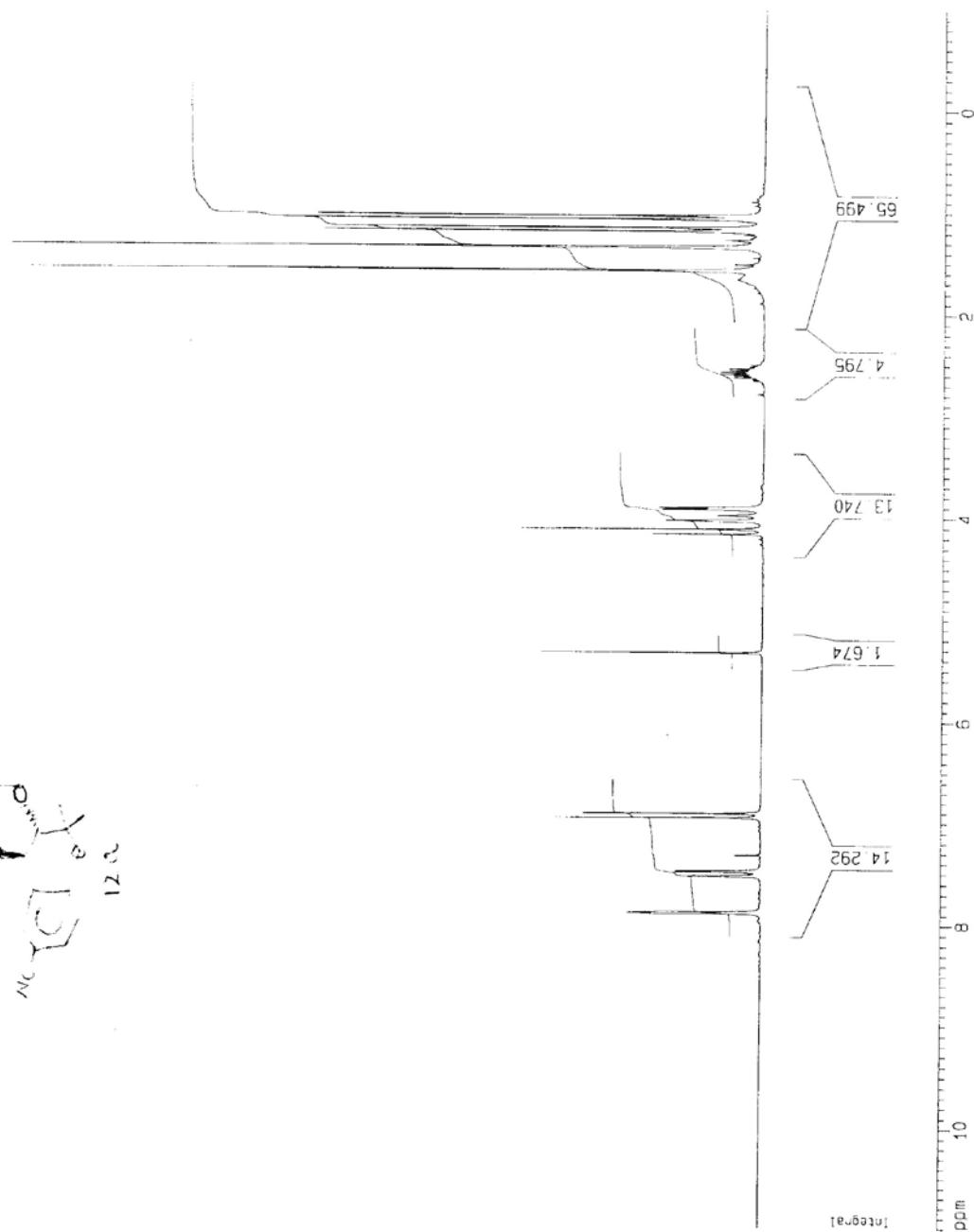


Current Data Parameters  
 NAME pw33  
 EXPR0 3  
 PHOCNO 1

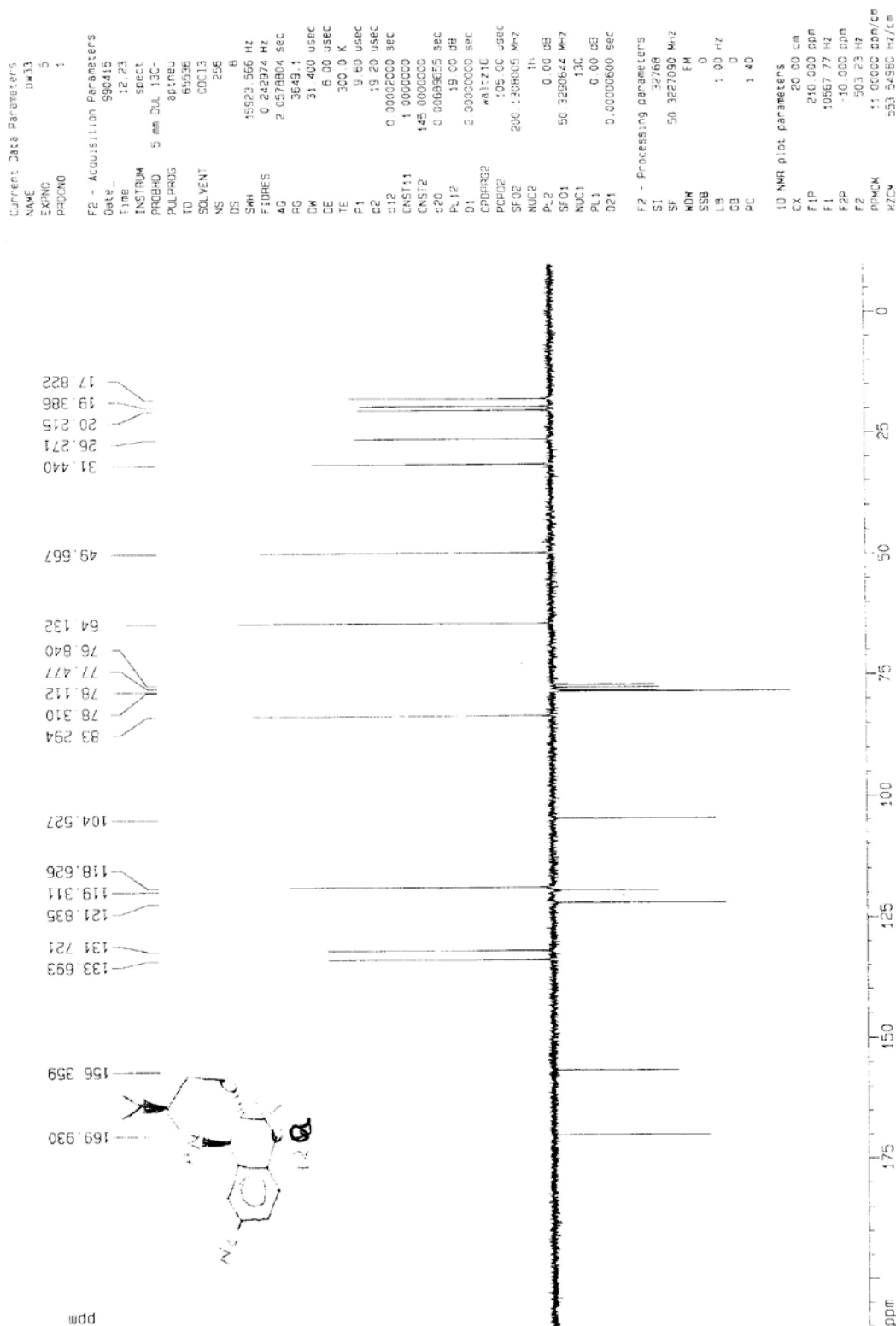
F2 - Acquisition Parameters  
 Date\_ 980415  
 Time 11:40  
 INSTRUM spect  
 PROBHD 5 mm QNP 13C-  
 PULPROG zg30  
 TO 32766  
 SOLVENT CDC13  
 NS 16  
 DS 2  
 SWH 4111.842 Hz  
 FIDRES 0.125403 Hz  
 AQ 3.9846387 sec  
 RG 114  
 DM 121.600 usec  
 DE 6.00 usec  
 TE 300.0 K  
 D1 1.0000000 sec  
 P1 10.70 usec  
 SFO1 200.1312359 MHz  
 NUC1 1H  
 PL: -2.00 dB

F2 - Processing parameters  
 S1 16384  
 SF 200.1299856 MHz  
 WDW EM  
 SSB 0  
 LB 0.30 Hz  
 GB 0  
 PC 1.00

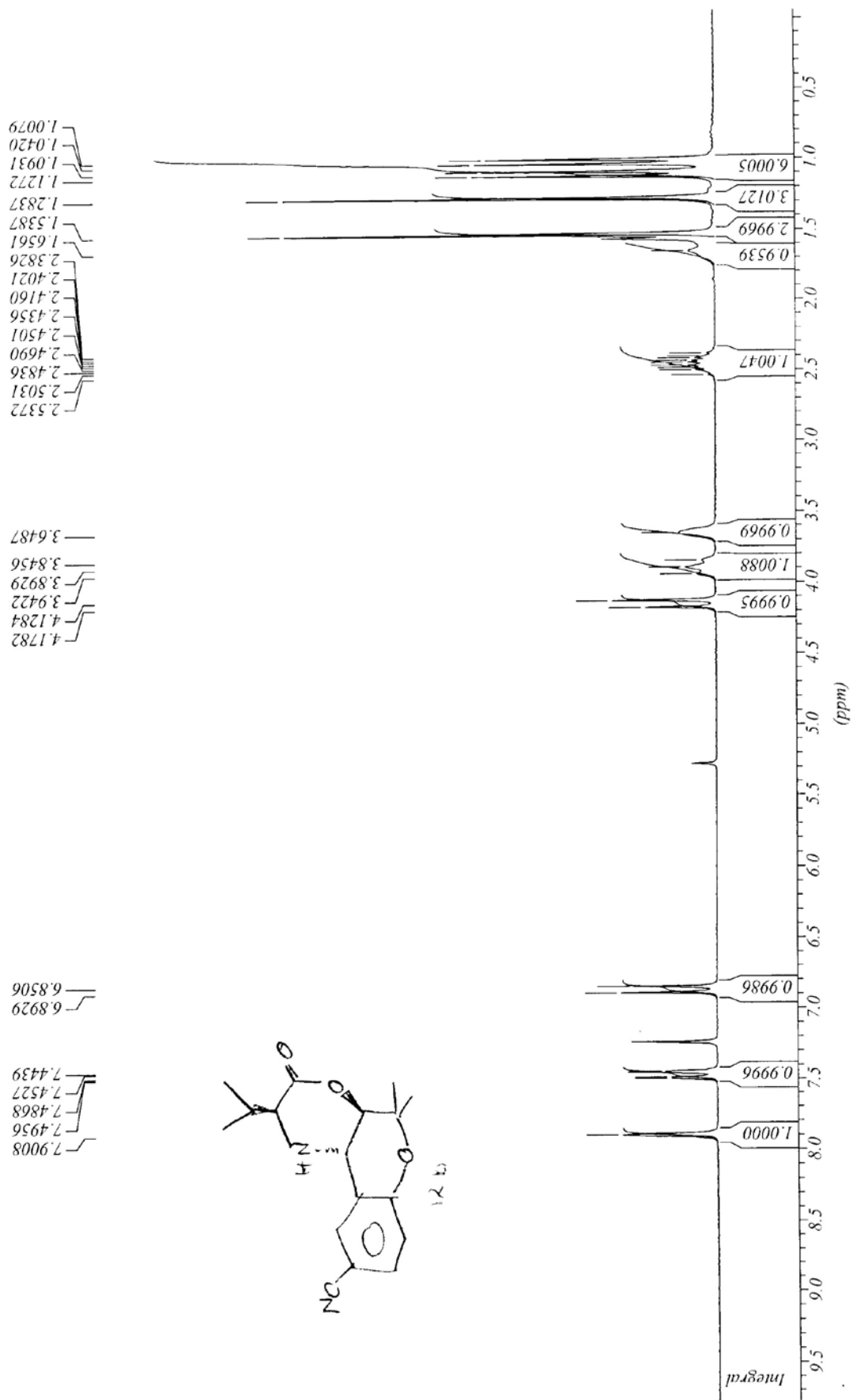
1D NMR plot parameters  
 CX 20.00 cm  
 F1P 11.000 ppm  
 F1 220.143 Hz  
 F2P -1.000 ppm  
 F2 -200.13 Hz  
 PPMCM 0.6000 ppm/cm  
 HZCM 120.07600 Hz/cm



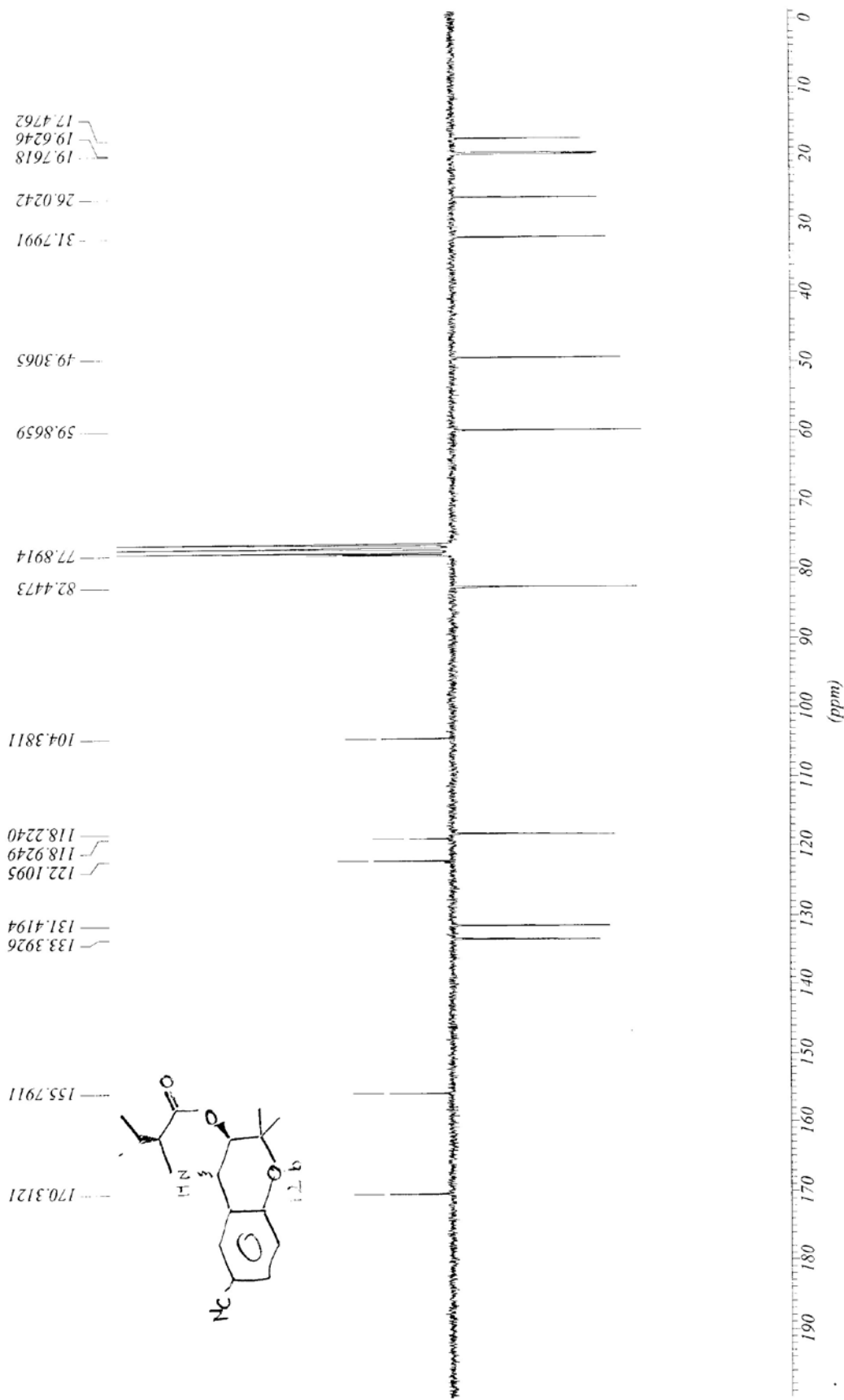
pw33 13C CDC13 15 Apr 99



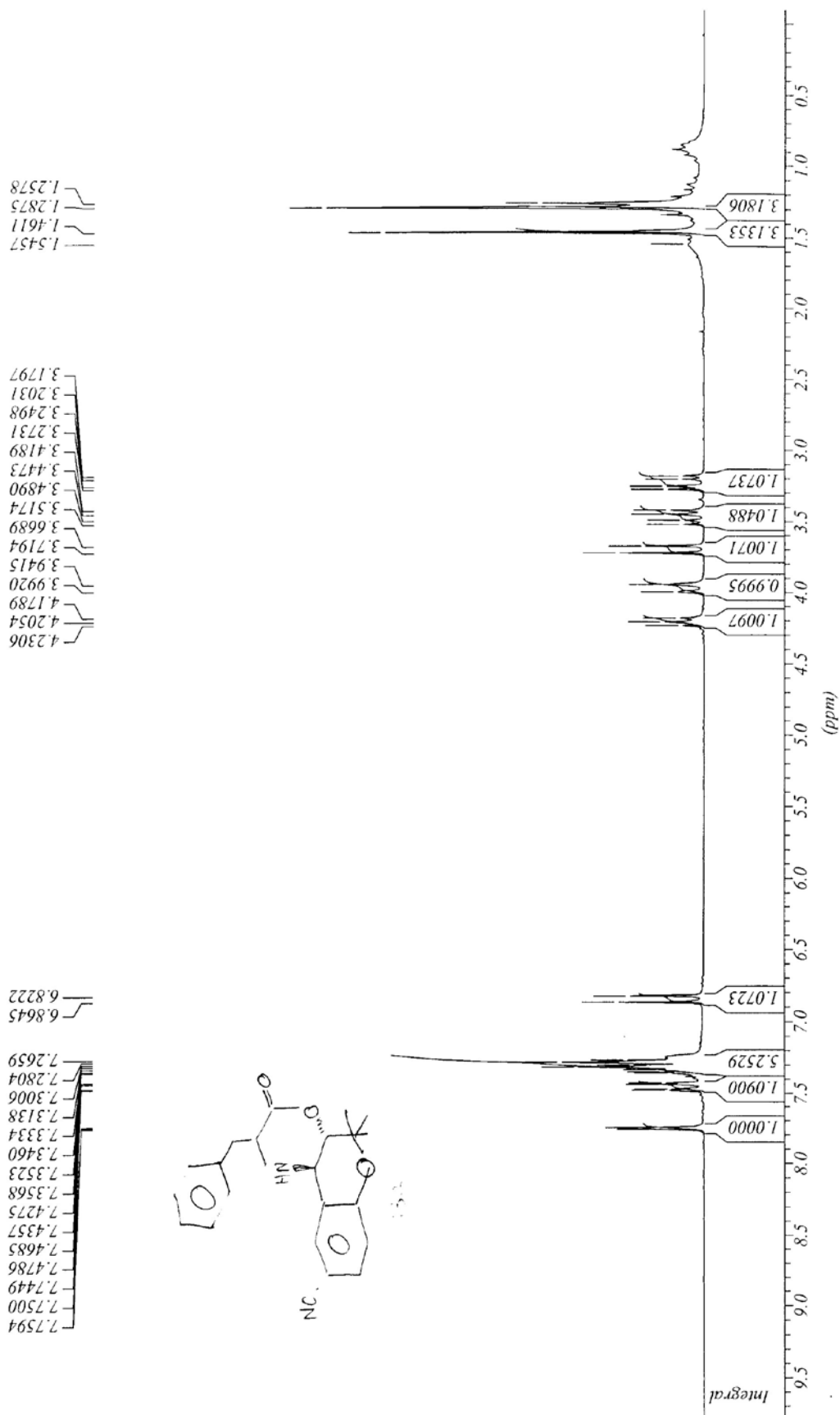
pw40f 1H CDCI3 10 Sept '00  
PROTON CDCI3 v pentium 17



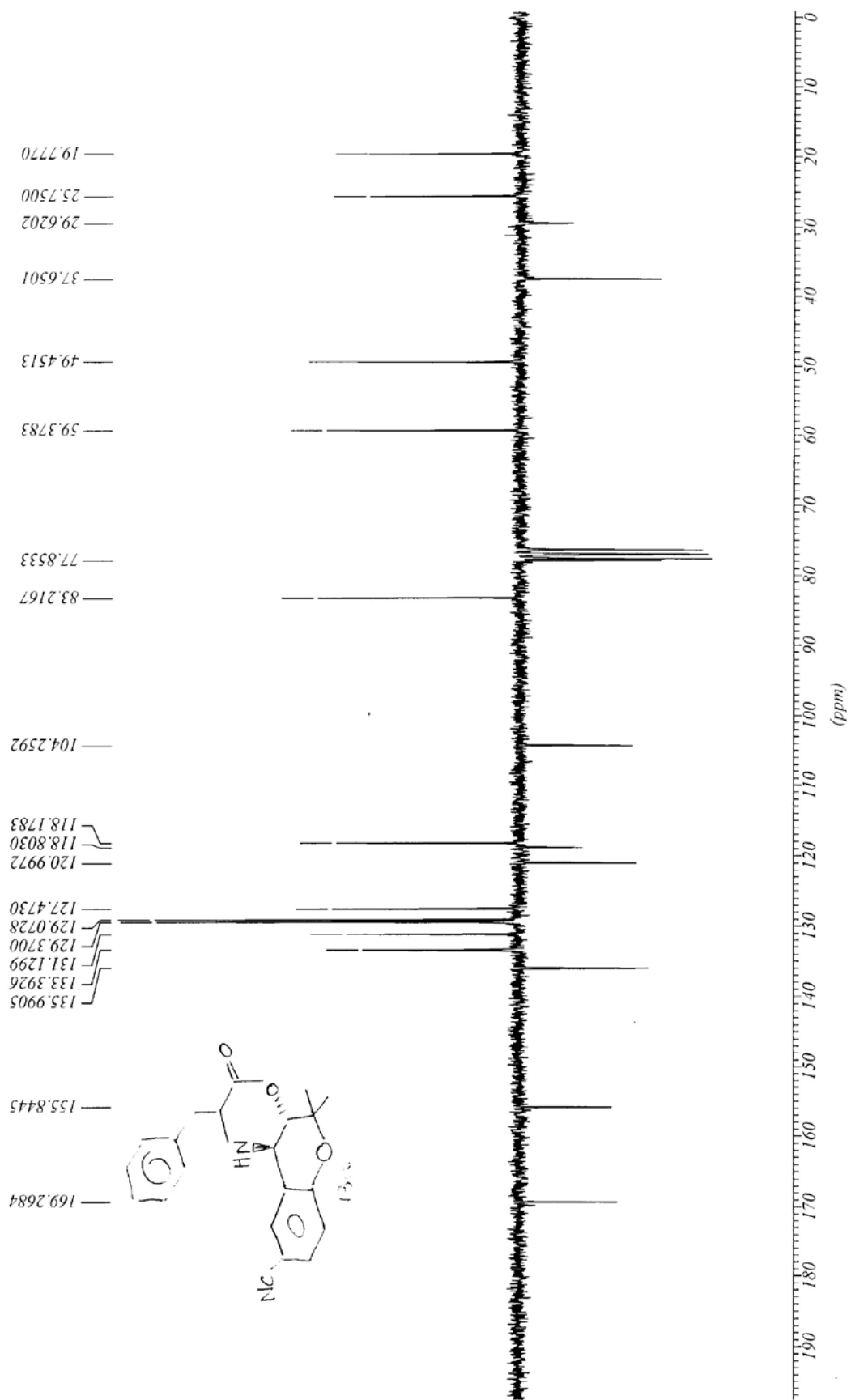
iconNMR  
pw40f 13C CDC13 10 Sept '00  
C13APT



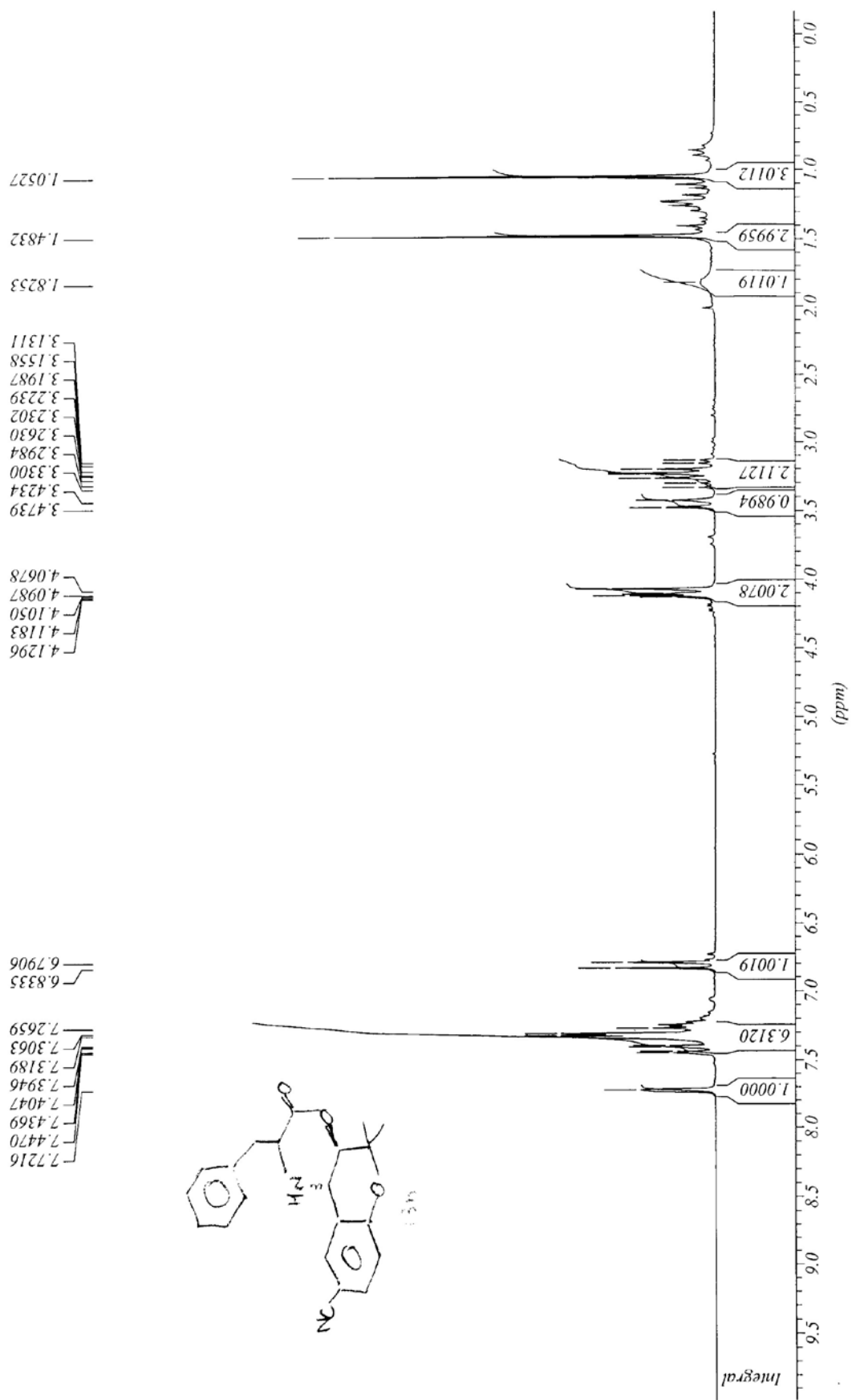
iconNMR  
pw180 1H CDCl3 28 May '00  
PROTON



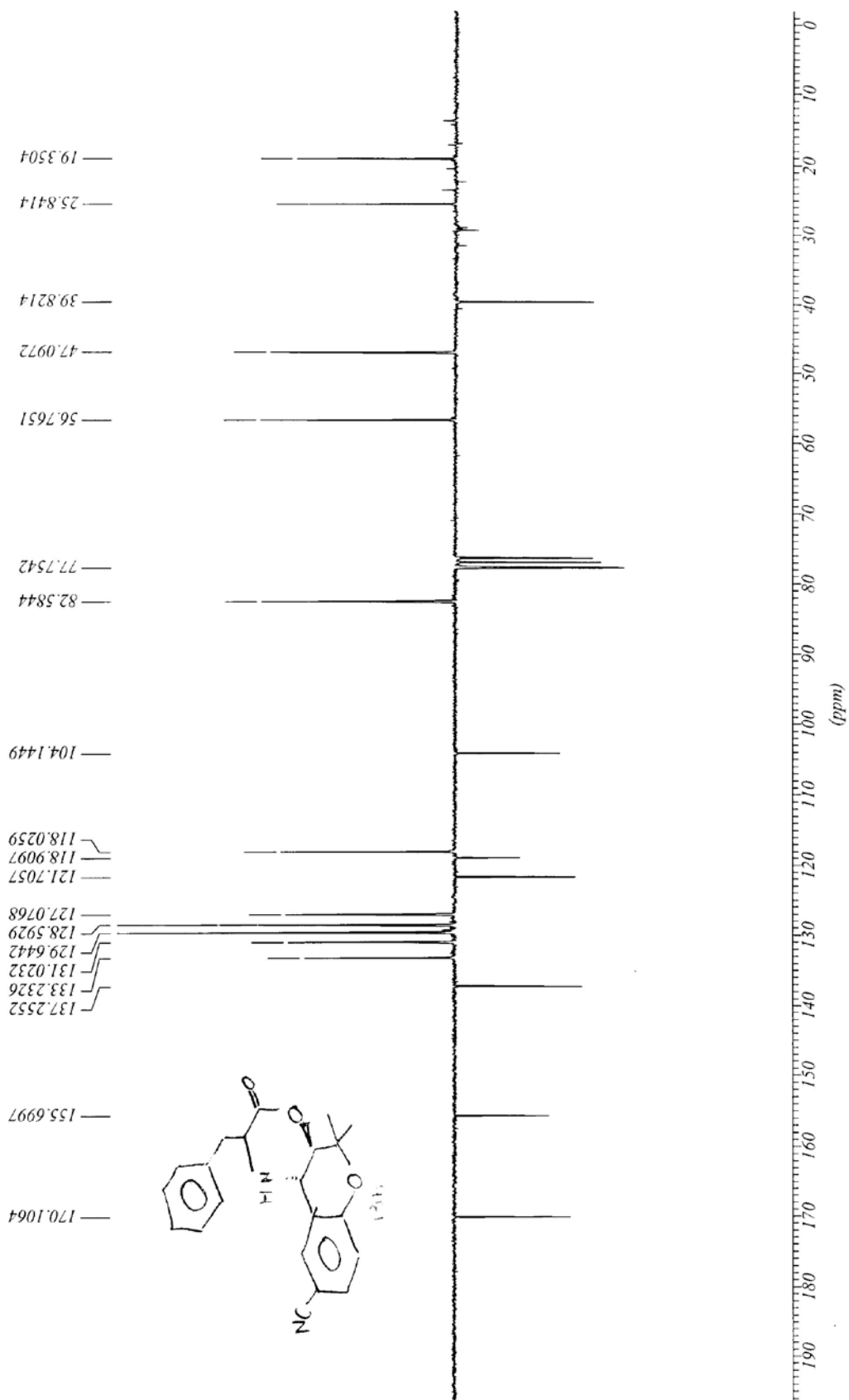
iconNMR  
pw180\_13C\_CDCl3\_28 May '00  
C13APT



pw181 1H CDCl3 6 Apr '00  
PROTON CDCl3 v penpin 4



iconNMR  
pw181 13C CDCl3 16 Apr '00  
C13APT





### References

1. .A. Sali, L. Potterton, F. Yuan, H. Van Vlijmen and M. Karplus, *Proteins*, 1995, **23**, 318-326.
2. S. G. Aller, J. Yu, A. Ward, Y. Weng, S. Chittaboina, R. Zhuo, P. M. Harrell, Y. T. Trinh, Q. Zhang, I. L. Urbatsch and G. Chang, *Science*, 2009, **323**, 1718-1722.
3. .R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. App. Cryst.* , 1993, **26**, 283-291.

## Supporting Information

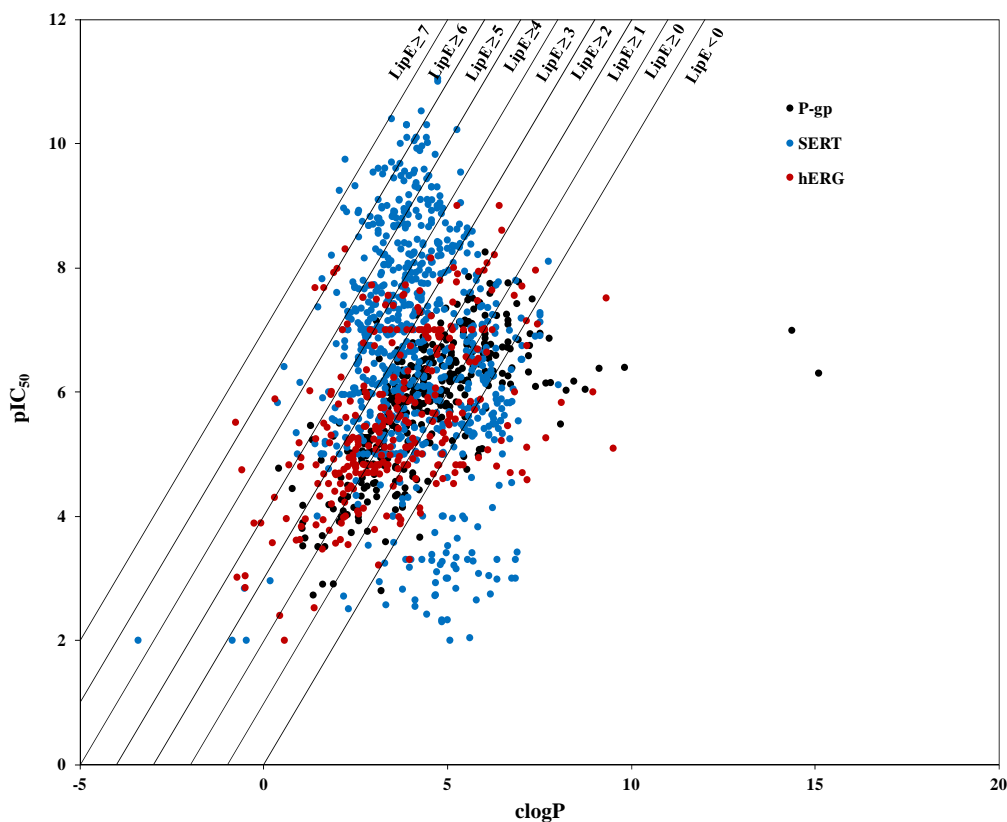
### Structure-activity relationships, ligand efficiency and lipophilic efficiency profiles of benzophenone-type inhibitors of the multidrug transporter P-glycoprotein

Ishrat Jabeen, Karin Pleban, Peter Chiba, ‡ Gerhard F. Ecker†\*

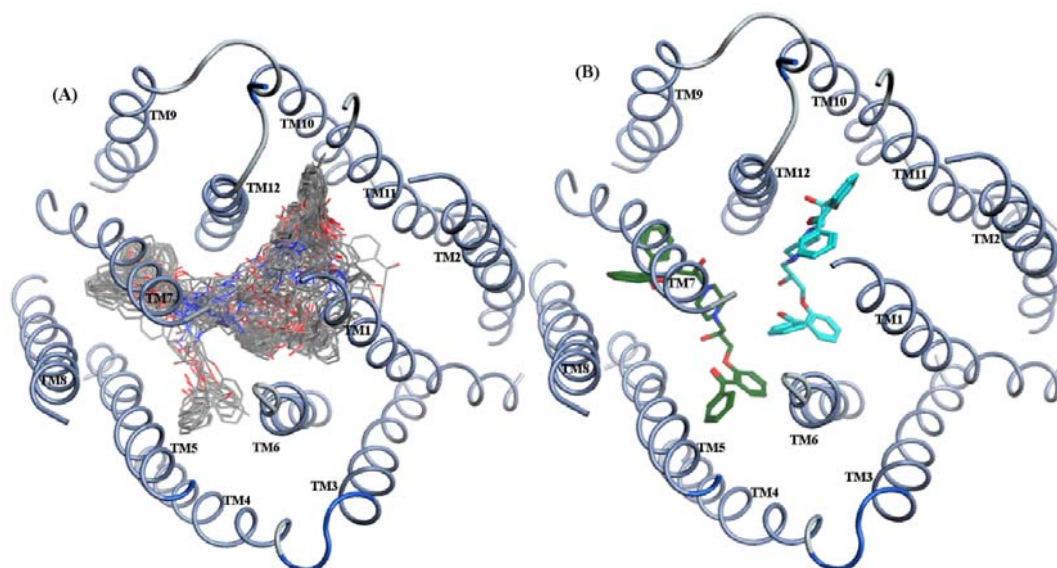
† University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, 1090,  
Vienna, Austria

‡ Medical University of Vienna, Institute of Medical Chemistry, Waehringerstraße 10,  
1090, Vienna, Austria

|                             |         |
|-----------------------------|---------|
| Figure 1                    | 195     |
| Figure 2                    | 195     |
| Figure 3                    | 196     |
| Figure 4                    | 196     |
| Data set of P-gp inhibitors | 197-207 |
| Data set of SERT ligands    | 208-225 |
| References                  | 226     |

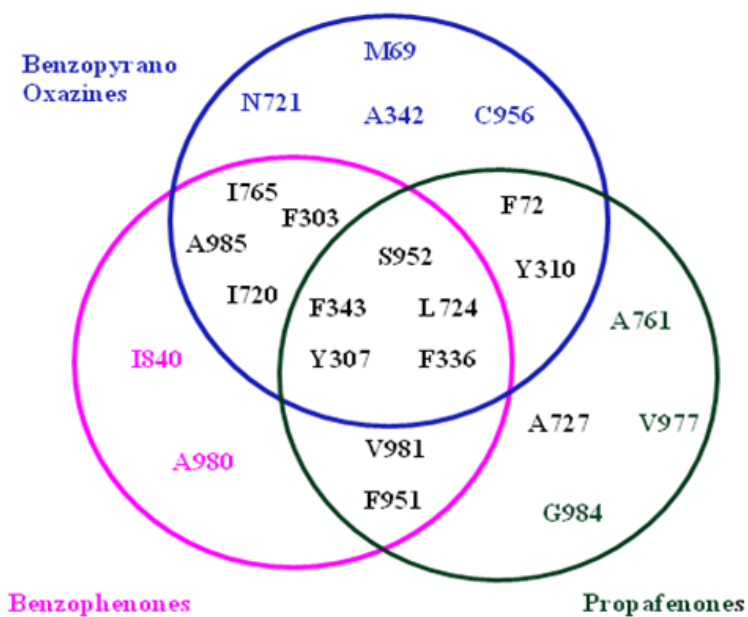


**SM Figure 1.** LipE distribution profiles of inhibitors of P-gp, SERT and hERG. A LipE value of greater than 5, clog ~2.5 and potency of ~10 nM are considered to be standard threshold of most promising ligands by Leeson *et al.*<sup>1</sup>

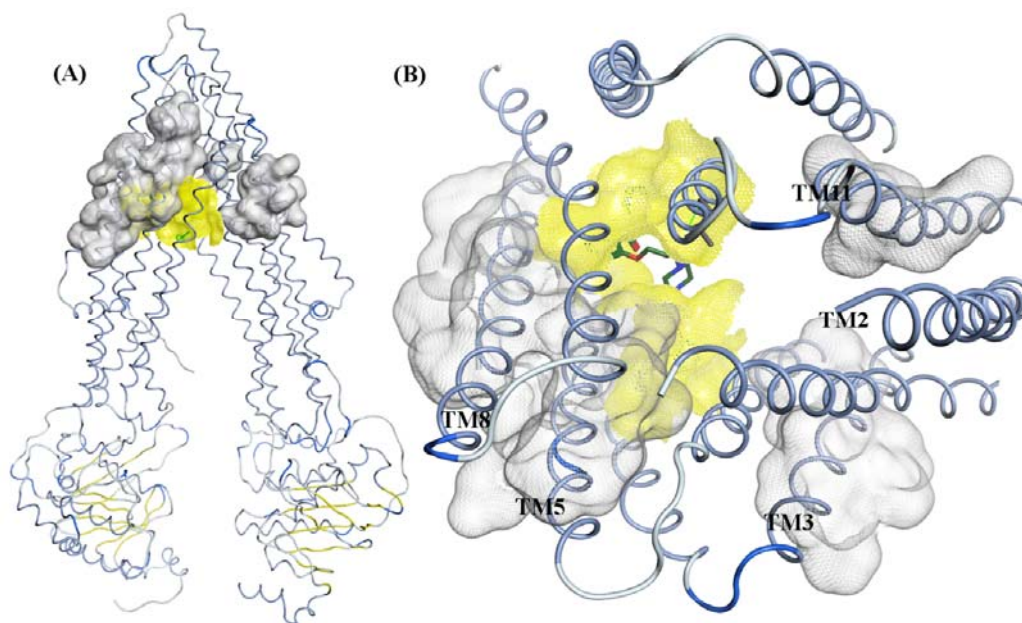


**SM Figure 2.** (A) showing docking poses in 7 clusters based on common scaffold of ligands. (B) Docking poses of 23 in two different clusters containing all four ligands, poses with green and blue color are representatives of cluster 1 and 2 respectively.

## APPENDIX A3



**SM Figure 3.** Overlap of interacting amino acid residues of propafenone type inhibitors of P-gp.



**SM Figure 4.** Photolabeled drug binding domains of propafenones analogues (TM3, 5, 8 and 11) represented by gray color looking from outside in the binding pocket of P-gp.<sup>1,2,3</sup> Yellow regions represent TM5, 6, 7, 8, 9 and 12 as proposed interaction positions of benzophenones in present studies.

## MDR DATA SET

| Code                 | IC50 $\mu$ M | SMILES   | Reference |
|----------------------|--------------|--|-----------|
| WIESE_E106           | 0.68         | <chem>S1c2cc(OCC)ccc2N(c2c1cccc2)CCCN1CCN(CC1)C</chem>                 |           |
| WIESE_E095           | 0.11         | <chem>S1c2c(N(c3c1cccc3)CCCNCC(c1cccc1)c1cccc1)ccc2</chem>             |           |
| WIESE_E085           | 1.48         | <chem>S1c2c(N(c3c1cccc3)CCCCCN1CCN(CC1)CCO)cccc2</chem>                |           |
| WIESE_E080           | 15.40        | <chem>S1c2c(N(c3c1cccc3)CCOCCN1CCN(CC1)C)cccc2</chem>                  |           |
| WIESE_E073           | 1.02         | <chem>S1c2c(N(c3c1cccc3)C(=O)CCCN1CCN(CC1)C)cc(cc2)C(F)(F)F</chem>     |           |
| WIESE_E051           | 3.28         | <chem>S1c2c3c(ccc2N(c2c1cccc2)CCN1CCCC1)cccc3</chem>                   |           |
| WIESE_E050           | 2.06         | <chem>S1c2c(N(c3c1cccc3)CCN1CCCC1)cccc2</chem>                         |           |
| WIESE_E013           | 5.95         | <chem>S1c2c(N(c3c1cccc3)C(=O)CCCN1CCN(CC1)CCO)ccc2</chem>              |           |
| WIESE_E008           | 5.20         | <chem>S1c2c(N(c3c1cccc3)C(=O)CCCN1CCN(CC1)C)cccc2</chem>               |           |
| WIESE_B015           | 0.45         | <chem>S1c2c(N(c3c1cccc3)CCCN1CCNCC1)cc(cc2)C(=O)C</chem>               |           |
| WIESE_B013           | 3.30         | <chem>Clc1cc(N2CCN(CC2)CCCN2c3cc(Cl)ccc3Sc3c2cccc3)ccc1</chem>         |           |
| WIESE_B011           | 2.19         | <chem>S1c2c(N(c3c1cccc3)CCCN1CCN(CC1)C)cc(cc2)C(=O)CCC</chem>          |           |
| WIESE_B008           | 16.96        | <chem>S1(=O)c2c(N(c3c1cccc3)CCCN1CCN(CC1)C)cc(cc2)C(=O)C</chem>        |           |
| WIESE_B007           | 0.99         | <chem>S1(=O)(=O)c2c(N(c3c1cccc3)CCCN1CCNCC1)cc(cc2)C(F)(F)F</chem>     |           |
| WIESE_B006           | 3.49         | <chem>S1(=O)(=O)c2c(N(c3c1cccc3)CCCN1CCNCC1)cc(cc2)C(=O)C</chem>       |           |
| WIESE_B005           | 0.55         | <chem>S1c2c(N(c3c1cccc3)CCCN1CCN(CC1)CCO)cc(cc2)C(=O)CCC</chem>        |           |
| WIESE_B001           | 36.09        | <chem>S1(=O)(=O)c2c(N(c3c1cccc3)CCCN1CCN(CC1)C)cc(cc2)C(=O)C</chem>    |           |
| UNT_REM 100          | 0.20         | <chem>s1c2c(ccc3c2cccc3)c(Cc2cccc2)c1C(O)CNC(C)C</chem>                |           |
| UNT_REM 097          | 0.46         | <chem>s1c2c(ccc3c2cccc3)c(Cc2cccc2)c1C(O)CNCCC</chem>                  |           |
| UNT_REM 091          | 0.29         | <chem>s1c2c(cccc2)c(Cc2cccc2)c1C(O)CNCCC</chem>                        |           |
| UNT_REM 044          | 0.31         | <chem>s1c2c(cccc2)c(Cc2ccc(F)cc2)c1C(O)CNCCC</chem>                    |           |
| UNT_REM 030          | 0.29         | <chem>s1c2c(cccc2)c(Cc2ccc(cc2)C)c1C(O)CNCCC</chem>                    |           |
| UNT_REM 025          | 0.64         | <chem>s1c2c(cccc2)c(Cc2ccc(cc2)C)c1C(O)CNC(C)C</chem>                  |           |
| SONST_Glybenclamide  | 219.50       | <chem>Clc1cc(C(=O)NCCc2ccc(S(=O)(=O)NC(=O)NC3CCCC3)cc2)c(OC)cc1</chem> |           |
| SONST_FumitremorginC | 21.41        | <chem>O(C)c1cc2[nH]c3c(C=C4N(C3C=C(C)C)C(=O)C3N(CC3)C4=O)c2cc1</chem>  |           |
| RICHT_P018           | 37.98        | <chem>O1C2C(=CN(CC)CC)C(=O)CCC2(c2cc(ccc12)C)C</chem>                  |           |
| RICHT_P006           | 27.69        | <chem>O1C2C(=CN3CCCC3)C(=O)CCC2(c2cc(ccc12)C)C</chem>                  |           |
| REC2252              | 0.42         | <chem>O(CCCCN(CC(CC(c1cccc1)c1cccc1)C)C)c1cccc1C(=O)CCc1cccc1</chem>   |           |
| REC2223              | 0.68         | <chem>O(CCCN(CCC(c1cccc1)c1cccc1)C)c1cccc1C(=O)CC=1C=CCCC=1</chem>     |           |
| REC2220              | 0.94         | <chem>O(CCN(C(CC(c1cccc1)c1cccc1)C)C)c1cccc1C(=O)CCc1cccc1</chem>      |           |
| REC2219              | 3.03         | <chem>O(CCNC)c1cccc1C(=O)CCc1cccc1</chem>                              |           |
| REC2218              | 0.56         | <chem>O(CCCN(Cc1cccc1)C)c1cccc1C(=O)CCc1cccc1</chem>                   |           |
| REC2203              | 0.91         | <chem>O(CCCN(C(CC(c1cccc1)c1cccc1)C)C)c1cccc1C(=O)CCc1cccc1</chem>     |           |

## APPENDIX A3

|             |         |  |   |
|-------------|---------|--|---|
|             |         | O)CCC=1C=CCCC=1  |   |
| PCO_GP794PW | 48.80   | O1c2c([C@H]3N[C@H](CO[C@@H]3C1(C)C(C)C)cc(cc2)C#N                      | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP790PW | 45.18   | ClC[C@@H](N(C)C1c2cc(ccc2OC(C)(C)C1O)C#N)C(C)C                         | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP788PW | 35.22   | ClC[C@@H](N(C)[C@@H]1c2cc(ccc2OC(C)(C)[C@H]1O)C#N)C(C)C                | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP786   | 6.84    | O1c2c(cc(cc2)C#N)C(N([C@@H](C(C)C)CO)C(O)C1(C)C                        | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP784   | 5.46    | O1c2c(cc(cc2)C#N)[C@@H](N([C@@H](C(C)C)CO)C)[C@H](O)C1(C)C             | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP782   | 54.06   | O1c2c(cc(cc2)C#N)[C@@H](N[C@@H](C(C)C)CO)[C@H](O)C1(C)C                | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP778   | 2.69    | O1c2c([C@H]3N[C@H](Cc4cccc4)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N             | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP777   | 259.78  | O1C2C(N[C@@H](Cc3cccc3)C1=O)c1cc(ccc1OC2(C)C)C#N                       | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP774   | 0.55    | O1c2c(cc(cc2)C#N)[C@@H](N[C@@H](Cc2cccc2)C(OC(C)(C)C)=O)[C@H](O)C1(C)C | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP770   | 0.77    | O1c2c(cc(cc2)C#N)C(N[C@@H](Cc2cccc2)C(OC(C)(C)C)=O)C(O)C1(C)C          | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP767   | 27.84   | O1C2C(N(C)C(C)C)C1=O)c1cc(ccc1OC2(C)C)C#N                              | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP766   | 79.27   | O1c2c([C@H]3N(C)[C@H](C(C)C)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N             | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP764   | 47.84   | O1c2c([C@H]3N(C)C(C)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N                     | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP763   | 1241.51 | O1c2c([C@H]3NC(C)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N                        | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP761   | 76.90   | O1C2C(N[C@@H](C)C1=O)c1cc(ccc1OC2(C)C)C#N                              | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP759   | 9.63    | O1C2C(NC(C)C)C1=O)c1cc(ccc1OC2(C)C)C#N                                 | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP758   | 59.33   | O1C2C(N[C@@H](C(C)C)C1=O)c1cc(ccc1OC2(C)C)C#N                          | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP757   | 15.32   | O1c2c([C@H]3N[C@H](C(C)C)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N                | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP754   | 3.96    | O1c2c(cc(cc2)C#N)[C@@H](N([C@H](C(OC(C)(C)C)=O)C)C)[C@H](O)C1(C)C      | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP750   | 3.72    | O1c2c(cc(cc2)C#N)C(N([C@H](C(OC(C)(C)C)=O)C)C(O)C1(C)C                 | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP746   | 14.55   | O1c2c(cc(cc2)C#N)C(N[C@H](C(OC(C)(C)C)=O)C(O)C1(C)C                    | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP742   | 29.85   | O1c2c(cc(cc2)C#N)[C@@H](N[C@H](C(OC(C)(C)C)=O)C)[C@H](O)C1(C)C         | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP730   | 1.35    | O1c2c(cc(cc2)C#N)C(N([C@@H](C(C)C)C(OC(C)(C)C)=O)C)C(O)C1(C)C          | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP726   | 0.96    | O1c2c(cc(cc2)C#N)[C@@H](N([C@@H](C(C)C)C(OC(C)(C)C)=O)C)[C@H](O)C1(C)C | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP723   | 4.63    | O1c2c(cc(cc2)C#N)[C@@H](N[C@H](C(C)C)C(OC(C)(C)C)=O)[C@H](O)C1(C)C     | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP719   | 1.34    | O1c2c(cc(cc2)C#N)C(NC(C)C)C(OC(C)(C)C)=O)C(O)C1(C)C                    | Submitted in Eur J Med Chem (Chapter 3) |
| PCO31       | 310.75  | O1c2c(cc(cc2)C#N)C(NC)C(O)C1(C)C                                       |   |
| PCO29       | 35.62   | O1c2c(cc(cc2)C#N)C(NCc2cccc2)C(O)C1(C)C                                |   |
| LACHM_Verb5 | 1874.00 | O1c2c(cc(N)c(O)c2)C(=CC1=O)C   |   |
| LACHM_Verb3 | 225.70  | O1c2c(cc(N)c(N3CCCC3CO)c2)C(=CC1=O)C                                   |   |

APPENDIX A3

|              |        |  |  |
|--------------|--------|--|--|
| LACHM_Verb2  | 65.70  | O1c2c(cc([N+](=O)[O-])c(N3CCCC3CO)c2)C(=CC1=O)C                |  |
| HOLZ_THPROP  | 0.82   | s1ccc(OCC(O)CNCCC)c1C(=O)CCc1cccc1                             |  |
| HOLZ_MB049   | 1.14   | lc1ccc(cc1)C(=O)c1cc(ccc1OCC(O)CNCCC)C                         |  |
| HOLZ_MB043   | 0.86   | lc1cccc1CN(CC(O)COc1cccc1C(=O)c1cccc1)C                        |  |
| HOLZ_MB041   | 6.14   | lc1ccc(cc1)C(=O)N(CCN(CC(O)COc1cccc1C(=O)c1cccc1)C)C           |  |
| HOLZ_LAN9    | 95.20  | s1cccc1C(=O)C=1C(=O)N(N(CC(O)CNCCC)C=1C)c1cccc1                |  |
| HOLZ_LAN8    | 81.11  | O=C1N(N(CC(O)CNC(C)(C)C)C(C)=C1C(=O)c1cccc1)c1cccc1            |  |
| HOLZ_LAN7    | 82.69  | O=C1N(N(CC(O)CNC(C)(C)C)C(C)=C1C(=O)c1cccc1)c1cccc1            |  |
| HOLZ_LAN6    | 122.53 | s1cccc1C(=O)C=1C(=O)N(N(CC(O)CNC(C)(C)C)C=1C)c1cccc1           |  |
| HOLZ_LAN5    | 72.83  | s1cccc1C(=O)C=1C(=O)N(N(CC(O)CNC(C)(C)C)C=1C)c1cccc1           |  |
| HOLZ_LAN4    | 2.70   | O=C1N(N(CC(O)CNCCC)C(C)=C1C(=O)CCc1cccc1)c1cccc1               |  |
| HOLZ_LAN3    | 7.73   | O=C1N(N(CC(O)CNC(C)(C)C)C(C)=C1C(=O)CCc1cccc1)c1cccc1          |  |
| HOLZ_LAN2    | 11.66  | O=C1N(N(CC(O)CNC(C)(C)C)C(C)=C1C(=O)CCc1cccc1)c1cccc1          |  |
| HOLZ_LAN10   | 184.72 | O=C1N(N(CC(O)CNCCC)C(C)=C1C(=O)c1cccc1)c1cccc1                 |  |
| HOLZ_LAN1    | 6.31   | O=C1N(N(CC(O)CNCCC)C(C)=C1C(=O)CCc1cccc1)c1cccc1               |  |
| HOLZ_Kar1    | 3.73   | O=C1N(N=C(C)C1C(=NCCC)CCc1cccc1)c1cccc1                        |  |
| HOLZ_KAR9a   | 11.03  | O=C(c1c(nn(c1NCCC)-c1cccc1)C)c1cccc1                           |  |
| HOLZ_GB010A  | 24.04  | O=C1N(N(CC(O)CN(C(C)C)C(C)C)C(C)=C1C(O)c1cccc1)c1cccc1         |  |
| HOLZ_GB007   | 1.21   | O(CC(O)CN(C(C)C)C(C)C)c1n(nc(C)c1C(=O)c1cccc1)C                |  |
| HOLZ_GB006   | 24.12  | O=C1N(N(CC(O)CN(C(C)C)C(C)C)C(C)=C1C(=O)c1cccc1)C              |  |
| HOLZ_GB005   | 0.90   | O=C1N(N(CC(O)CN2CCN(CC2)c2cccc2)C(C)=C1C(=O)c1cccc1)c1cccc1    |  |
| HOLZ_GB004   | 0.68   | O=C1N(N=C(C)C1=C(N1CCN(CC1)c1cccc1)C)c1cccc1                   |  |
| HOLZ_GB001   | 15.69  | O=C1N(N(CC(O)CN(C(C)C)C(C)C)C(C)=C1C(=O)c1cccc1)c1cccc1        |  |
| HOLZ_BRI47/1 | 8.22   | O=C(CCc1cccc1)c1c(nn(c1NCCC)-c1cccc1)C                         |  |
| HOLZ_BP037   | 5.50   | O(CC(O)CN1CCCC1)c1cccc1C(O)c1cccc1                             |  |
| HOLZ_BP030   | 0.40   | [Sj](OC(CN1CCCC1)COc1cccc1C(=O)c1cccc1)(C(C)C)(c1cccc1)c1cccc1 |  |
| HOLZ_BP029   | 1.42   | O(CC(OC(=O)C)CN1CCCC1)c1cccc1C(=O)c1cccc1                      |  |
| HOLZ_BP028   | 8.96   | O(CC(O)CN1CCCC1)c1cccc1-c1ncccc1                               |  |
| HOLZ_BP026   | 1.33   | O(CC(O)CN1CCCC1)c1cccc1NCc1cccc1                               |  |
| HOLZ_BP025   | 1.27   | O(CC(O)CN1CCCC1)c1c2c(ccc1)C(=O)c1c(ccc1)C2=O                  |  |
| HOLZ_BP024   | 1.36   | O(CC(O)CN1CCCC1)c1cc([N+](=O)[O-])ccc1C(=O)c1cccc1             |  |
| HOLZ_BP023   | 1.15   | O(CC(O)CN1CCCC1)c1cccc1Cc1cccc1                                |  |
| HOLZ_BP022   | 0.16   | Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2-                               |  |

## APPENDIX A3

|            |       |  |                                     |
|------------|-------|--|-------------------------------------|
|            |       | c2nc3c(cc2)cccc3)cc1   |                                     |
| HOLZ_BP021 | 0.05  | O(C[C@H](O)CN1CCN(CC1)c1cccc(C)c1C)c1cccc1-c1nc2c(cc1)cccc2      |                                     |
| HOLZ_BP020 | 0.04  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1-c1nc2c(cc1)cccc2              |                                     |
| HOLZ_BP019 | 5.44  | O1CCN(CC1)C[C@@H](O)COc1cccc1-c1nc2c(cc1)cccc2                   |                                     |
| HOLZ_BP018 | 1.04  | O(CC(O)CN1CCCC1)c1cccc1-c1nc2c(cc1)cccc2                         |                                     |
| HOLZ_BP017 | 2.75  | O(CC(O)CN1CCCC1)c1cccc1C(=NOC)c1cccc1                            |                                     |
| HOLZ_BP015 | 8.07  | O(CC(O)CN1CCCC1)c1cccc1C(=NO)c1cccc1                             |                                     |
| HOLZ_BP013 | 1.09  | Clc1cc(OCC(O)CN2CCCC2)ccc1C(=O)c1cccc1                           |                                     |
| HOLZ_BP011 | 0.55  | O(CC(O)CN1CCCC1)c1cc(OC)ccc1C(=O)c1cccc1                         |                                     |
| HOLZ_BP007 | 4.32  | O(CC(O)CN1CCCC1)c1ncnc1C(=O)c1cccc1                              |                                     |
| HOLZ_BP002 | 1.03  | s1ccc(OCC(O)CNCCCC)c1C(=O)c1cccc1                                |                                     |
| HOLZ_BP001 | 1.17  | s1ccc(OCC(O)CN2CCCC2)c1C(=O)c1cccc1                              | Submitted in J Med Chem (Chapter 6) |
| HOLZ_B59   | 1.20  | O(CC(O)CN1CCCC1)c1cccc1C(=O)c1cccc1                              |                                     |
| GPV0930    | 0.12  | O(CC(O)CN(C(C)C)C(C)C)c1ccc(cc1)C(=O)CCc1c2c(cc1)cccc2           |                                     |
| GPV0929    | 0.26  | O(CC(O)CN1CCCC1)c1ccc(cc1)C(=O)CCc1c2c(cc1)cccc2                 |                                     |
| GPV0927    | 0.44  | O(CC(O)CN(C(C)C)C(C)C)c1ccc(cc1)C(=O)c1cccc1                     |                                     |
| GPV0926    | 0.26  | O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cc(ccc1)C(=O)CCc1c2c(ccc1)cccc2 |                                     |
| GPV0921    | 0.41  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(cc1)C(=O)CCc1c2c(ccc1)cccc2     |                                     |
| GPV0920    | 0.56  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2)C(=O)CCc2c3c(ccc2)cccc3)cc1    |                                     |
| GPV0902    | 0.58  | Fc1ccc(N2CCN(CC2)CC(O)COc2cc(ccc2)C(=O)c2ccc(cc2)cc1             | Submitted in J Med Chem (Chapter 6) |
| GPV0901    | 7.84  | O(CC(O)CN(C(C)C)C(C)C)c1ccc(cc1)C(=O)CC                          | Submitted in J Med Chem (Chapter 6) |
| GPV0900    | 2.18  | O(CC(O)CN1CCCC1)c1ccc(cc1)C(=O)c1cccc1                           | Submitted in J Med Chem (Chapter 6) |
| GPV0899    | 5.32  | O1CCN(CC1)CC(O)COc1ccc(cc1)C(=O)c1cccc1                          | Submitted in J Med Chem (Chapter 6) |
| GPV0898    | 0.02  | Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(=O)CCc2c3c(ccc2)cccc3)cc1       | Submitted in J Med Chem (Chapter 6) |
| GPV0896    | 0.08  | O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cccc1C(=O)c1cccc1               | Submitted in J Med Chem (Chapter 6) |
| GPV0893    | 1.54  | O1CCN(CC1)CC(O)COc1ccc(cc1)C(=O)CCc1c2c(ccc1)cccc2               |                                     |
| GPV0865    | 32.91 | O(CC(O)CN1CCCC1)c1ccc(cc1)C(=O)CC                                |                                     |
| GPV0863    | 0.86  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(cc1)C(=O)CC                     |                                     |
| GPV0862    | 1.43  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2)C(=O)CC)cc1                    |                                     |
| GPV0861    | 1.42  | O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)CC                 |                                     |
| GPV0855    | 0.50  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(cc1)C(=O)c1cccc1                | Submitted in J Med Chem (Chapter 6) |
| GPV0845    | 3.18  | O(CC(O)CN1CCN(CC1)c1ccc(cc1)C)c1ccc(cc1C(=O)C)C                  | Submitted in J Med Chem (Chapter 6) |
| GPV0842    | 7.34  | O(CC(O)CN1CCN(CC1)C(=O)Nc1ccc(cc1)C)c1ccc(cc                     |                                     |



APPENDIX A3

|         |       |  |                                     |
|---------|-------|--|-------------------------------------|
|         |       | 1C(=O)C  |                                     |
| GPV0839 | 9.48  | O(CC(O)CN1CCN(CC1)C[C@H](O)COc1cccc1C(=O)c1cccc1)c1ccc(cc1C(=O)C)C | Submitted in J Med Chem (Chapter 6) |
| GPV0831 | 0.05  | O(CC(O)CN1CCN(CC1)CC(O)COc1cccc1C(=O)c1ccc(cc1)c1cccc1C(=O)c1cccc1 | Submitted in J Med Chem (Chapter 6) |
| GPV0826 | 0.48  | O(CC(O)CN1CCN(CC1)C(=O)Nc1ccc(cc1)C)c1cccc1C(=O)c1cccc1            | Submitted in J Med Chem (Chapter 6) |
| GPV0825 | 0.38  | S=C(Nc1ccc(cc1)C)N1CCN(CC1)CC(O)COc1cccc1C(=O)c1cccc1              | Submitted in J Med Chem (Chapter 6) |
| GPV0824 | 0.06  | O(CC(O)CN1CCN(CC1)c1ccc(C)c1C)c1ccc(cc1C(=O)CCc1cccc1)C            | Submitted in J Med Chem (Chapter 6) |
| GPV0823 | 0.24  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2C(=O)CCc2ccc2)C)cc1               |                                     |
| GPV0821 | 1.67  | O1CCN(CC1)CC(O)COc1ccc(cc1C(=O)CCc1cccc1)C                         |                                     |
| GPV0818 | 0.27  | O(CC(O)CN1CCCC1)c1ccc(cc1C(=O)CCc1cccc1)C                          |                                     |
| GPV0810 | 0.33  | O(CC(O)CN1CCN(CC1)c1ccc(C)c1C)c1ccc(cc1C(=O)C)C                    |                                     |
| GPV0809 | 3.60  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2C(=O)C)C)cc1                      |                                     |
| GPV0808 | 6.13  | O(CC(O)CN(C(C)C)C(C)C)c1ccc(cc1C(=O)C)C                            |                                     |
| GPV0807 | 12.65 | O1CCN(CC1)CC(O)COc1ccc(cc1C(=O)C)C                                 |                                     |
| GPV0806 | 10.77 | O(CC(O)CN1CCCC1)c1ccc(cc1C(=O)C)C                                  |                                     |
| GPV0798 | 0.12  | O(CC(O)CNCCC(c1cccc1)c1cccc1)c1cccc1C(=O)C                         | Kaiser et al 2005 <sup>2</sup>      |
| GPV0797 | 0.17  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1C(=O)CC                          |                                     |
| GPV0796 | 1.45  | O(CC(O)CN1CCN(CC1)Cc1cccc1)c1cccc1C(=O)CC                          | Kaiser et al 2005                   |
| GPV0795 | 4.29  | O(CC(O)CN(C(C)C)C(C)C)c1cccc1C(=O)CC                               | Kaiser et al 2005                   |
| GPV0793 | 0.08  | O(CC(O)CN1CCN(CC1)c1ccc(C)c1C)c1cccc1C(O)Cc1cccc1                  | Zdrzil 2007 <sup>3</sup>            |
| GPV0792 | 0.04  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1C(OC)Cc1cccc1                    | Zdrzil 2007                         |
| GPV0791 | 0.04  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1C(O)CCc1cccc1                    | Zdrzil 2007                         |
| GPV0790 | 0.08  | O(CC(O)CN1CCN(CC1)Cc1cccc1)c1cccc1C(OC)Cc1cccc1                    |                                     |
| GPV0789 | 0.19  | O(CC(O)CN1CCN(CC1)Cc1cccc1)c1cccc1C(O)CCc1cccc1                    |                                     |
| GPV0788 | 10.09 | O(CC(O)CN(C(C)C)C(C)C)c1cccc1C(=O)C                                |                                     |
| GPV0704 | 0.87  | O(CC(O)CN1CCN(CC1)C(OC(C)C)C)O)c1cccc1C(=O)c1cccc1                 |                                     |
| GPV0703 | 0.09  | O(CC(O)CN1CCC(O)(CC1)c1cccc1)c1ccc(OCc2cccc2)cc1C(=O)CCc1cccc1     | Submitted in J Med Chem (Chapter 6) |
| GPV0649 | 0.19  | O(CC(O)CNCCC(c1cccc1)c1cccc1)c1cc(ccc1)C(=O)CCc1cccc1              | Klein et al, 2002 <sup>4</sup>      |
| GPV0647 | 0.07  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(ccc1)C(=O)Cc1cccc1                 | Klein et al, 2002                   |
| GPV0645 | 0.18  | O(CC(O)CNCCC(c1cccc1)c1cccc1)c1ccc(cc1)C(=O)C                      | Klein et al, 2002                   |
| GPV0643 | 0.04  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(OC)ccc1C(=O)CCc1ccc(N(C)C)cc1      | Klein et al, 2002                   |
| GPV0636 | 0.12  | Clc1cc(C(=O)C)c(OCC(O)CN2CCN(CC2)c2cccc2)C)cc1Cl                   | Klein et al, 2002                   |
| GPV0626 | 0.25  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1C(=O)C                           | Klein et al, 2002                   |
| GPV0616 | 0.21  | O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(C)c(cc1C(=O)C)C(=O)C               | Klein et al, 2002                   |

## APPENDIX A3

|         |         |   |   |
|---------|---------|---|---|
| GPV0615 | 0.40    | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(cc1C(=O)C)C</chem>                           | Klein <i>et al</i> , 2002               |
| GPV0613 | 0.86    | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(OC)cc1C(=O)C</chem>                          | Klein <i>et al</i> , 2002               |
| GPV0610 | 0.01    | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1cccc1C(=O)CCc1ccc(N(C)C)cc1</chem>               | Klein <i>et al</i> , 2002               |
| GPV0608 | 0.30    | <chem>O(CC(O)CN1CCCCC1)c1cc(OC)ccc1C(=O)CCc1cccc1</chem>                            | Langer <i>et al</i> , 2004 <sup>5</sup> |
| GPV0600 | 0.23    | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)CCc1cccc1</chem>                | Langer <i>et al</i> , 2004              |
| GPV0595 | 48.92   | <chem>O1c2c(cccc2)C(C)C1CN1CCCCC1</chem>  | Klein <i>et al</i> , 2002               |
| GPV0594 | 1.62    | <chem>O1c2c(cccc2)C(CCC2cccc2)C1CN1CCCCC1</chem>                                    | Klein <i>et al</i> , 2002               |
| GPV0591 | 103.57  | <chem>O1c2c(cccc2)C(C)C1C(=O)N1CCCCC1</chem>  | Klein <i>et al</i> , 2002               |
| GPV0590 | 6.66    | <chem>O1c2c(cccc2)C(CCC2cccc2)C1C(=O)N1CCCCC1</chem>                                | Klein <i>et al</i> , 2002               |
| GPV0579 | 0.38    | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)C</chem>                        | Langer <i>et al</i> , 2004              |
| GPV0577 | 1.55    | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1ccc(cc1)C(=O)C</chem>                            | Langer <i>et al</i> , 2004              |
| GPV0576 | 0.01    | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cccc1C(=O)CCc1cccc1</chem>                   | Langer <i>et al</i> , 2004              |
| GPV0571 | 1587.49 | <chem>O(CC(O)=O)c1cccc1C(=O)CCc1cccc1</chem>  | Klein <i>et al</i> , 2002               |
| GPV0570 | 23.18   | <chem>O(CC(OCC)=O)c1cccc1C(=O)CCc1cccc1</chem>                                      | Klein <i>et al</i> , 2002               |
| GPV0568 | 1251.09 | <chem>O1c2c(cccc2)C(=O)CC1CN1CCOCC1</chem>  |   |
| GPV0566 | 8.86    | <chem>O1c2c(cccc2)C(=O)CC1CN1CCN(CC1)Cc1cccc1</chem>                                |   |
| GPV0563 | 2.06    | <chem>O1c2c(cccc2)C(=O)CC1CN1CCN(CC1)c1ccc(C)c1C</chem>                             |   |
| GPV0562 | 17.72   | <chem>Fc1ccc(N2CCN(CC2)CC2Oc3c(cccc3)C(=O)C2)cc1</chem>                             |   |
| GPV0559 | 1.73    | <chem>O1c2c(cccc2)C(CCC2cccc2)C1C(O)CN1CCCCC1</chem>                                | Langer <i>et al</i> , 2004              |
| GPV0543 | 176.30  | <chem>O1c2c(cccc2)C(=O)CC1CN1CCCCC1</chem>  | Langer <i>et al</i> , 2004              |
| GPV0525 | 0.29    | <chem>O1CCN(CC1)CC(O)COc1c2c(CC(Cc3cccc3)C2=O)ccc1</chem>                           |   |
| GPV0524 | 0.39    | <chem>O(CC(O)CN1CCN(CC1)c1ccc(OC)cc1)c1c2c(CC(Cc3cccc3)C2=O)ccc1</chem>             |   |
| GPV0523 | 0.17    | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1c2c(CC(Cc3cccc3)C2=O)ccc1</chem>             | Langer <i>et al</i> , 2004              |
| GPV0522 | 0.14    | <chem>O(CC(O)CN1CCCCC1)c1c2CC(Cc3cccc3)C(=O)c2ccc1</chem>                           | Langer <i>et al</i> , 2004              |
| GPV0491 | 0.29    | <chem>O(CC(O)CNC(C(C)C)C)c1cccc1)c1cccc1C(=O)CCc1cccc1</chem>                       | Langer <i>et al</i> , 2004              |
| GPV0485 | 9.68    | <chem>O(CC(O)CN1CCCCC1)c1c2c(CCC2=O)ccc1</chem>                                     | Langer <i>et al</i> , 2004              |
| GPV0476 | 0.02    | <chem>Clc1cc(C(=O)CCc2cccc2)c(OCC(O)CN2CCN(CC2)c2cccc2C)cc1Cl</chem>                | Langer <i>et al</i> , 2004              |
| GPV0472 | 0.15    | <chem>O(CC(O)CNC(C1CCCC1)c1cccc1)c1cccc1C(=O)CCc1cccc1</chem>                       | Langer <i>et al</i> , 2004              |
| GPV0470 | 0.21    | <chem>O(CC(O)CNC(C(C)C)c1cccc1)c1cccc1C(=O)CCc1cccc1</chem>                         | Langer <i>et al</i> , 2004              |
| GPV0442 | 1.21    | <chem>O(CC(O)CNCCC)c1ccc(cc1C(=O)c1cccc1)C</chem>                                   | Submitted in J Med Chem (Chapter 6)     |
| GPV0409 | 1.84    | <chem>S(=O)(=O)(N(CCN(CC(O)COc1cccc1C(=O)c1cccc1)C)C)c1c2c(ccc1)c(N(C)C)ccc2</chem> |   |
| GPV0391 | 302.05  | <chem>O1CCN(CC1)CC(O)COc1ccc(cc1)C(=O)C</chem>                                      |   |
| GPV0389 | 48.97   | <chem>O(CC(O)CN1CCCCC1)c1ccc(cc1)C(=O)C</chem>                                      | Langer <i>et al</i> , 2004              |
| GPV0388 | 0.13    | <chem>O(CC(O)CN1CCN(CC1)C(=O)c1cccc1)c1cccc1C(=O)CCc1cccc1</chem>                   | Langer <i>et al</i> , 2004              |
| GPV0386 | 10.07   | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2cc(ccc2)C(=O)C)cc1</chem>                           | Langer <i>et al</i> , 2004              |

APPENDIX A3

|         |      |  |  |
|---------|------|--|--|
| GPV0382 | 0.07 | <chem>O(CC(O)CN1CCC(O)(CC1)c1ccccc1)c1ccccc1C(=O)CCc1c2c(ccc1)cccc2</chem>   | Langer <i>et al</i> , 2004             |
| GPV0381 | 0.30 | <chem>O(CC(O)CN1CCC(O)(CC1)c1ccccc1)c1ccccc1C(=O)CC</chem>                   | Langer <i>et al</i> , 2004             |
| GPV0376 | 0.47 | <chem>O(CC(O)CN(C(C)C)C(C)C)c1ccccc1C(=O)CCc1c2c(ccc1)cccc2</chem>           | Langer <i>et al</i> , 2004             |
| GPV0366 | 1.69 | <chem>O(CC(O)CN(C(=O)c1ccccc1)CCC)c1ccccc1C(=O)CCc1ccccc1</chem>             | Langer <i>et al</i> , 2004             |
| GPV0363 | 0.19 | <chem>Fc1ccc(N2CCN(CC2)CCCCO)c2ccccc2C(=O)CCc2ccccc2)cc1</chem>              | Langer <i>et al</i> , 2004             |
| GPV0361 | 1.01 | <chem>O1CCN(CC1)CCCCOc1ccccc1C(=O)CCc1ccccc1</chem>                          | Langer <i>et al</i> , 2004             |
| GPV0359 | 2.57 | <chem>O(CC(O)CNc1ccc([N+](=O)[O-])cc1)c1ccccc1C(=O)CCc1ccccc1</chem>         | Langer <i>et al</i> , 2004             |
| GPV0358 | 8.49 | <chem>FC(F)(F)c1ccc(NCC(O)CO)c2ccccc2C(=O)CCc2ccccc2)cc1</chem>              | Langer <i>et al</i> , 2004             |
| GPV0356 | 0.18 | <chem>O(CC(O)CN1CCN(CC1)c1ccccc1C)c1ccc(OC)cc1C(=O)CCc1ccccc1</chem>         | Langer <i>et al</i> , 2004             |
| GPV0339 | 1.54 | <chem>O(CC(O)CNc1ccc(cc1)C(OC)=O)c1ccccc1C(=O)CCc1ccccc1</chem>              | Langer <i>et al</i> , 2004             |
| GPV0338 | 0.34 | <chem>O(CC(OC(=O)C)CN1CCCCC1)c1ccccc1C(=O)CCc1ccccc1</chem>                  | Langer <i>et al</i> , 2004             |
| GPV0336 | 0.01 | <chem>O(CC(O)CN1CCN(CC1)c1ccccc1C)c1ccccc1C(=O)CCc1ccc(OC)cc1</chem>         | Langer <i>et al</i> , 2004             |
| GPV0334 | 0.02 | <chem>Clc1ccc(cc1)CCC(=O)c1ccccc1OCC(O)CN1CCN(CC1)c1ccccc1C</chem>           | Langer <i>et al</i> , 2004             |
| GPV0323 | 3.20 | <chem>O(CC(O)CN1CCC(O)(CC1)c1ccccc1)c1ccccc1C(=O)CC</chem>                   | Langer <i>et al</i> , 2004             |
| GPV0321 | 0.36 | <chem>O(CC(O)CN1CCC(CC1)Cc1ccccc1)c1ccccc1C(=O)CC</chem>                     | Langer <i>et al</i> , 2004             |
| GPV0317 | 0.31 | <chem>O(CC(O)CN1CCC(O)(CC1)c1ccccc1)c1ccccc1C(=O)c1ccccc1</chem>             | Submitted in J Med Chem (Chapter 6)    |
| GPV0304 | 0.85 | <chem>O(CC(O)CN1CCCCC1)c1ccccc1C(O)CCc1ccccc1</chem>                         |  |
| GPV0265 | 0.40 | <chem>O(CC(O)CN1CCCCC1)c1ccccc1C(OC(C)C)CCc1ccccc1</chem>                    |  |
| GPV0264 | 0.46 | <chem>O(CC(O)CN1CCCCC1)c1ccccc1C(OC(C)C)CCc1ccccc1</chem>                    |  |
| GPV0254 | 0.53 | <chem>O(CC(O)CN1CCN(CC1)Cc1ccccc1)c1ccc(O)cc1C(=O)CCc1ccccc1</chem>          |  |
| GPV0253 | 0.12 | <chem>O(CC(O)CN1CCN(CC1)Cc1ccccc1)c1ccc(OCc2ccccc2)cc1C(=O)CCc1ccccc1</chem> |  |
| GPV0246 | 0.08 | <chem>O(CC(O)CN(C(C)C)C(C)C)c1ccc(OCc2ccccc2)cc1C(=O)CCc1ccccc1</chem>       | Chiba <i>et al</i> , 1997 <sup>6</sup> |
| GPV0245 | 1.04 | <chem>O(CC(O)CN(C(C)C)C(C)C)c1ccc(O)cc1C(=O)CCc1ccccc1</chem>                | Chiba <i>et al</i> , 1997              |
| GPV0242 | 0.34 | <chem>O(CC(O)CNC1CCCCC1)c1ccccc1C(=O)CCc1ccccc1</chem>                       |  |
| GPV0240 | 6.47 | <chem>O(CC(O)CNc1ccccc1)c1ccccc1C(=O)CCc1ccccc1</chem>                       |  |
| GPV0233 | 1.73 | <chem>O(CC(O)CN1CCCCC1)c1ccc(O)cc1C(=O)CCc1ccccc1</chem>                     | Langer <i>et al</i> , 2004             |
| GPV0232 | 0.17 | <chem>O(CC(O)CN1CCCCC1)c1ccc(OCc2ccccc2)cc1C(=O)CCc1ccccc1</chem>            | Chiba <i>et al</i> , 1997              |
| GPV0231 | 0.11 | <chem>O(CC(O)CNCCC)c1ccc(OCc2ccccc2)cc1C(=O)CCc1ccccc1</chem>                | Chiba <i>et al</i> , 1998              |
| GPV0227 | 1.81 | <chem>O1CCN(CC1)CC(O)COc1ccccc1C(OC)CCc1ccccc1</chem>                        | Langer <i>et al</i> , 2004             |
| GPV0226 | 9.54 | <chem>O1CCN(CC1)CC(O)COc1ccccc1C(O)CCc1ccccc1</chem>                         | Langer <i>et al</i> , 2004             |
| GPV0220 | 0.75 | <chem>O(CC(O)CNCCC(c1ccccc1)c1ccccc1)c1ccccc1C(OC)CCc1ccccc1</chem>          | Langer <i>et al</i> , 2004             |
| GPV0216 | 0.14 | <chem>O(CCCCCCN1CCCCC1)c1ccccc1C(=O)CCc1ccccc1</chem>                        |  |
| GPV0211 | 0.18 | <chem>O(CCCCCCN1CCCCC1)c1ccccc1C(=O)CCc1ccccc1</chem>                        | Langer <i>et al</i> , 2004             |

## APPENDIX A3

|         |        |   |                            |
|---------|--------|---|----------------------------|
| GPV0206 | 0.20   | O(CCCCCCN1CCCCC1)c1cccc1C(=O)CCc1cccc1                      | Langer <i>et al</i> , 2004 |
| GPV0201 | 0.24   | O(CCCCCN1CCCCC1)c1cccc1C(=O)CCc1cccc1                       | Langer <i>et al</i> , 2004 |
| GPV0195 | 0.53   | O(CCCCN1CCCCC1)c1cccc1C(=O)CCc1cccc1                        | Klein <i>et al</i> , 2002  |
| GPV0189 | 1.45   | O(CCN1CCCCC1)c1cccc1C(=O)CCc1cccc1                          | Langer <i>et al</i> , 2004 |
| GPV0186 | 0.65   | O(CCCN1CCCCC1)c1cccc1C(=O)CCc1cccc1                         | Langer <i>et al</i> , 2004 |
| GPV0184 | 0.72   | O(CC(O)CNCCC(c1cccc1)c1cccc1)c1cccc1C(=O)CCc1c2c(ccc1)cccc2 | Langer <i>et al</i> , 2004 |
| GPV0182 | 10.40  | O(CC(O)CNC(c1cccc1)c1cccc1)c1cccc1C(=O)CCc1cccc1            | Langer <i>et al</i> , 2004 |
| GPV0181 | 0.80   | O(CC(O)CNCC(c1cccc1)c1cccc1)c1cccc1C(=O)CCc1cccc1           | Langer <i>et al</i> , 2004 |
| GPV0164 | 0.66   | O(CC(O)CN(C(C)C)C(C)C)c1cccc1C(OC)CCc1cccc1                 | Langer <i>et al</i> , 2004 |
| GPV0163 | 1.74   | O(CC(O)CN(C(C)C)C(C)C)c1cccc1C(O)CCc1cccc1                  | Langer <i>et al</i> , 2004 |
| GPV0159 | 0.92   | O(CC(O)CN(C(C)C)C(C)C)c1ccc(cc1)C(=O)CCc1cccc1              | Langer <i>et al</i> , 2004 |
| GPV0156 | 0.23   | Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(OC)CCc2ccc2)cc1            | Langer <i>et al</i> , 2004 |
| GPV0155 | 0.67   | Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(O)CCc2ccc2)cc1             | Langer <i>et al</i> , 2004 |
| GPV0149 | 6.88   | O1CCN(CC1)CC(O)COc1ccc(cc1)C(=O)CCc1cccc1                   | Langer <i>et al</i> , 2004 |
| GPV0135 | 0.42   | O(CC(O)CN1CCCCC1)c1cc(ccc1)C(=O)CCc1cccc1                   | Langer <i>et al</i> , 2004 |
| GPV0134 | 2.54   | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2)C(=O)CCc2ccc2)cc1         | Langer <i>et al</i> , 2004 |
| GPV0129 | 3.02   | O(CC(O)CNCCC)c1ccc(O)cc1C(=O)CCc1cccc1                      | Langer <i>et al</i> , 2004 |
| GPV0128 | 0.26   | O(CC(O)CN1CCC(CC1)Cc1cccc1)c1cccc1C(=O)CCc1cccc1            | Langer <i>et al</i> , 2004 |
| GPV0095 | 0.18   | o1c2c(cccc2)c(CCc2cccc2)c1C(O)CN1CCN(CC1)c1cccc1OC          | Klein <i>et al</i> , 2002  |
| GPV0094 | 1.10   | o1c2c(cccc2)c(CCc2cccc2)c1C(O)CNCCC                         | Langer <i>et al</i> , 2004 |
| GPV0092 | 11.52  | o1c2c(cccc2)c(CC)c1C(O)CNCCC                                | Klein <i>et al</i> , 2002  |
| GPV0090 | 0.17   | O(CC(O)CN1CCCCC1)c1cccc1C(OC)CCc1cccc1                      | Langer <i>et al</i> , 2004 |
| GPV0062 | 0.06   | O(CC(O)CN1CCC(O)(CC1)c1cccc1)c1cccc1C(=O)CCc1cccc1          | Langer <i>et al</i> , 2004 |
| GPV0056 | 0.32   | O(CC(O)CN1CCCCC1)c1cccc1C(=O)C=Cc1cccc1                     |                            |
| GPV0051 | 2.32   | O(CC(O)CN(CC)CC)c1cccc1C(=O)c1cccc1                         | Chiba <i>et al</i> , 1997  |
| GPV0050 | 0.42   | O(CC(O)CN(CCC)C)c1cccc1C(=O)CCc1cccc1                       |                            |
| GPV0049 | 67.32  | O1CCN(CC1)CC(O)COc1cccc1C(=O)C                              |                            |
| GPV0048 | 2.48   | O(CC(O)CN(C)C)c1cccc1C(=O)CCc1cccc1                         |                            |
| GPV0046 | 207.20 | O1CCN(CC1)CC(O)COc1cccc1C(=O)CC                             |                            |
| GPV0031 | 0.07   | Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(=O)CCc2ccc2)cc1            | Langer <i>et al</i> , 2004 |
| GPV0029 | 0.66   | O(CC(O)CN1CCN(CC1)Cc1cccc1)c1cccc1C(=O)CCc1cccc1            | Langer <i>et al</i> , 2004 |
| GPV0023 | 0.68   | O(CC(O)CN1CCN(CC1)c1cc(OC)ccc1)c1cccc1C(=O)CCc1cccc1        | Langer <i>et al</i> , 2004 |
| GPV0019 | 0.61   | O(CC(O)CN1CCN(CC1)c1cccc1)c1cccc1C(=O)CCc1cccc1             | Langer <i>et al</i> , 2004 |
| GPV0017 | 31.96  | O(CC(O)CN1CCCCC1)c1cccc1C(=O)C                              | Langer <i>et al</i> , 2004 |
| GPV0012 | 14.31  | O(CC(O)CN1CCCCC1)c1cccc1C(=O)CC                             | Langer <i>et al</i> , 2004 |
| GPV0009 | 0.38   | O(CC(O)CN(C(C)C)C(C)C)c1cccc1C(=O)CCc1cccc1                 | Langer <i>et al</i> , 2004 |

APPENDIX A3

|                 |        |   |   |
|-----------------|--------|---|---|
| GPV0001R        | 0.28   | O(CC(O)CNCCC)c1ccccc1C(=O)CCc1ccccc1                        |   |
| ERK_Ti93        | 0.84   | S1CCCN(c2ccsc12)C(=O)CCNCCc1ccccc1                          |   |
| ERK_Pi98t       | 1.42   | S1c2sccc2N(CCN(C)C)C(=O)C(OC(=O)C)C1c1ccc(OC)cc1            |   |
| ERK_PA008       | 0.42   | S1c2sc(cc2N(CCN(C)C)C(=O)C(OC(=O)C)C1c1ccc(OC)cc1)Cc1ccccc1 |   |
| ERK_MSD057Ox    | 10.79  | s1c2SCCN(c2cc1CC)C(=O)NCCN1CCCC1                            |   |
| ERK_MSD028      | 3.39   | s1c2SCCN(c2cc1CC)C(=O)NCCCN(CC)CC                           |   |
| ERK_MSD025      | 0.58   | s1c2SCCN(c2cc1CC)C(=O)CCNCCc1cc(OC)c(OC)cc1                 |   |
| ERK_MSD024      | 0.22   | s1c2SCCN(c2cc1CC)C(=O)CCN(CCc1cc(OC)c(OC)cc1)C              |   |
| ERK_MSD023      | 38.20  | s1c2SCCN(c2cc1CC)C(=O)NCCN(C)C                              |   |
| ERK_MSD017      | 44.82  | s1c2SCCN(c2cc1CC)C(=O)N1CCN(CC1)C                           |   |
| ERK_MS098       | 26.07  | s1c2SCCN(c2cc1CC)C(=O)Nc1ccccc1                             |   |
| ERK_MS087HCl    | 72.13  | s1c2SCCN(c2cc1CC)C(=O)Nc1ccccc1                             |   |
| ERK_MS083       | 0.49   | O1CCN(c2c1cccc2)C(=O)NCCN(CCc1cc(OC)c(OC)cc1)C              |   |
| ERK_MS074HCl    | 20.23  | s1c2SCCN(c2cc1CC)C(=O)N1CCN(CC1)C                           |   |
| ERK_MS069HCl    | 119.37 | s1c2SCCN(c2cc1CC)C(=O)c1ccccc1                              |   |
| ERK_MS068HCl    | 0.28   | s1c2SCCN(c2cc1CC)C(=O)NCCN1CCc2cc(OC)c(OC)c2C1              |   |
| ERK_MS055Ox     | 28.89  | s1c2SCCN(c2cc1CC)C(=O)NCCN1CCOCC1                           |   |
| ERK_MS048Ox     | 5.12   | s1c2SCCN(c2cc1CC)C(=O)N1CCN(CC1)Cc1ccccc1                   |   |
| ERK_HO7Ox       | 0.36   | s1c2SCCN(c2cc1Cc1ccccc1)C(=O)NCCN(CCc1cc(OC)c(OC)cc1)C      |   |
| ERK_HO5Ox       | 4.77   | s1c2SCCN(c2cc1Cc1ccccc1)C(=O)N1CCN(CC1)C                    |   |
| CSOEL_UCHL14    | 0.08   | Fc1ccc(N2CCN(CC2)CC(O)COc2ccccc2NC(OCCC)=O)cc1              |   |
| CSOEL_UCHL13    | 0.20   | Fc1ccc(N2CCN(CC2)CC(O)COc2ccccc2NC(OCCC)=O)cc1              |   |
| CSOEL_UCHL12    | 0.33   | Fc1ccc(N2CCN(CC2)CC(O)COc2ccccc2NC(OCC)=O)c1                |   |
| AG-227/33912017 | 158.30 | [N+] <sub>1</sub> (=NC(C=C1C)C)c1nnc(nn1)NN=Cc1ccccc1       |   |
| AG-205/37047268 | 0.82   | Clc1ccc(N2C(Nc3c(cccc3)C2=O)c2ccc(cc2)C(CC)C)c1             |   |
| AG-205/33114008 | 102.20 | Oc1c(cccc1O)C=[NH+]NC(=O)C(=O)NC                            |   |
| AF-399/34011064 | 17.70  | Clc1cc(Cl)cc(Cl)c1NN=C(C(OCC)=O)C(OCC)=O                    |   |
| AF-399/13927090 | 311.80 | O(C(=O)c1ccc(N)cc1)CCOCC                                    |   |
|                 |        | Test Set  |   |
| UNT_REM 104     | 2.89   | s1c2c(cccc2)c(Cc2ccccc2)c1C(O)CNCCC                         |   |
| UNT_REM 014     | 1.03   | s1c2c(cccc2)c(Cc2ccccc2)c1C(O)CNC(C)C                       |   |
| UNT_REM 013     | 1.46   | s1c2c(cccc2)c(Cc2ccccc2)c1C(O)CNCCC                         |   |
| PCO_GP780       | 0.01   | O1c2c(cc(cc2)C#N)[C@H](N[C@@H](C(C)C)CO)[C@@H](O)C1(C)C     | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP768       | 0.06   | O1C2C(N(C)[C@H](C(C)C)C1=O)c1cc(ccc1OC2(C)C)C#N             | Submitted in Eur J Med Chem (Chapter 3) |

## APPENDIX A3

|             |       |  |   |
|-------------|-------|--|---|
| PCO_GP765   | 0.02  | <chem>O1c2c([C@H]3N(C)[C@@H](C(C)C)C(O[C@@H]3C1(C)C)=O)cc(cc2)C#N</chem>           | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP762   | 0.03  | <chem>O1C2C(N(C)[C@@H](C)C1=O)c1cc(ccc1OC2(C)C)C#N</chem>                          | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP738   | 0.99  | <chem>O1c2c(cc(cc2)C#N)[C@@H](N([C@H](C(C)C)C(OC(C)(C)C)=O)C)[C@H](O)C1(C)C</chem> | Submitted in Eur J Med Chem (Chapter 3) |
| PCO_GP734   | 1.01  | <chem>O1c2c(cc(cc2)C#N)[C@H](N([C@H](C(C)C)C(OC(C)(C)C)=O)C)[C@H](O)C1(C)C</chem>  | Submitted in Eur J Med Chem (Chapter 3) |
| HOLZ_BP031  | 0.41  | <chem>O(CC(O)CN1CCCC1)c1cccc1-c1nccc2c1cccc2</chem>                                |   |
| HOLZ_BP027  | 0.80  | <chem>O(CC(O)CN1CCCC1)c1cccc1CNc1cccc1</chem>                                      |   |
| HOLZ_BP017a | 0.73  | <chem>O(CC(O)CN1CCCC1)c1cccc1C(=NOC)c1cccc1</chem>                                 |   |
| HOLZ_BP016  | 0.16  | <chem>O(CC(O)CN1CCCC1)c1cccc1C(=NO)c1cccc1</chem>                                  |   |
| HOLZ_B047   | 0.11  | <chem>O(CC(O)CN1CCCC1)c1nccnc1C(=O)c1cccc1</chem>                                  |   |
| GPV0936     | 17.02 | <chem>O(CC(O)CN(C(C)C)C(C)C)c1cc(ccc1)C(=O)CCc1c2c(cc1)cccc2</chem>                |   |
| GPV0935     | 6.90  | <chem>O(CC(O)CN1CCCC1)c1cc(ccc1)C(=O)CCc1c2c(ccc1)cccc2</chem>                     |   |
| GPV0934     | 1.58  | <chem>O1CCN(CC1)CC(O)COc1cc(ccc1)C(=O)CCc1c2c(ccc1)cccc2</chem>                    |   |
| GPV0933     | 14.17 | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2cc(ccc2)C(=O)CCc2c3c(ccc2)cccc3)cc1</chem>         |   |
| GPV0932     | 6.35  | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2cc(ccc2)C(=O)CCc2c3c(ccc2)cccc3)cc1</chem>         |   |
| GPV0928     | 0.09  | <chem>O1CCN(CC1)CC(O)COc1cc(ccc1)C(=O)c1cccc1</chem>                               |   |
| GPV0919     | 2.08  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)CCc1c2c(ccc1)cccc2</chem>      |   |
| GPV0906     | 4.87  | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(ccc1)C(=O)c1cccc1</chem>                     | Submitted in J Med Chem (Chapter 6)     |
| GPV0904     | 0.28  | <chem>O(CC(O)CN1CCCC1)c1cc(ccc1)C(=O)c1cccc1</chem>                                | Submitted in J Med Chem (Chapter 6)     |
| GPV0903     | 5.76  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cc(ccc1)C(=O)c1cccc1</chem>                 | Submitted in J Med Chem (Chapter 6)     |
| GPV0897     | 0.07  | <chem>O1CCN(CC1)CC(O)COc1cccc1C(=O)c1cccc1</chem>                                  | Submitted in J Med Chem (Chapter 6)     |
| GPV0859     | 1.53  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)c1cccc1</chem>                 | Submitted in J Med Chem (Chapter 6)     |
| GPV0854     | 1.03  | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2)C(=O)c2ccc(cc2)cc1</chem>                  |   |
| GPV0840     | 0.29  | <chem>O(CC(O)CN1CCN(CC1)c1ccc(OC)cc1)c1ccc(cc1C(=O)C)C</chem>                      |   |
| GPV0794     | 1.20  | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(=O)CC)cc1</chem>                            | Zdrazil 2007                            |
| GPV0787     | 0.45  | <chem>Fc1ccc(N2CCN(CC2)CC(O)COc2cccc2C(=O)C)cc1</chem>                             |   |
| GPV0655     | 1.98  | <chem>O(CC(O)CNCCC(c1cccc1)c1cccc1)c1ccc(cc1)C(=O)CCc1cccc1</chem>                 | Klein <i>et al</i> , 2002               |
| GPV0653     | 0.29  | <chem>O1CCN(CC1)CC(O)COc1cc(ccc1)C(=O)CCc1cccc1</chem>                             | Klein <i>et al</i> , 2002               |
| GPV0651     | 7.30  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cc(ccc1)C(=O)CCc1cccc1</chem>               | Klein <i>et al</i> , 2002               |
| GPV0633     | 6.17  | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(ccc1C(=O)C)C</chem>                          |   |
| GPV0607     | 32.26 | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(OC)ccc1C(=O)CCc1cccc1</chem>                 |   |
| GPV0598     | 5.00  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1cccc1C(=O)C</chem>                          | Langer <i>et al</i> , 2004              |
| GPV0596     | 1.86  | <chem>O(CC(O)CN1CCN(CC1)c1cccc1C)c1cc(ccc1)C(=O)C</chem>                           | Langer <i>et al</i> , 2004              |
| GPV0574     | 0.74  | <chem>O(CC(O)CN1CCN(CC1)c1cccc(C)c1C)c1ccc(cc1)C(=O)C</chem>                       | Langer <i>et al</i> , 2004              |

APPENDIX A3

|             |       |   |  |
|-------------|-------|---|--|
|             |       | O)C   |  |
| GPV0558     | 0.48  | O1c2c(ccccc2)C(CCC2ccccc2)C1C(O)CN1CCCCC1                   | Langer <i>et al</i> , 2004             |
| GPV0557     | 0.87  | O1c2c(ccccc2)C(CCC2ccccc2)C1C(O)CN1CCCCC1                   | Langer <i>et al</i> , 2004             |
| GPV0556     | 1.54  | O1c2c(ccccc2)C(CCC2ccccc2)C1C(O)CN1CCCCC1                   | Langer <i>et al</i> , 2004             |
| GPV0515     | 4.72  | O(CC(O)CN1CCCCC1)c1c2c(CC(Cc3ccccc3)C2=O)cc<br>c1           | Klein <i>et al</i> , 2002              |
| GPV0512     | 0.09  | O(CC(O)CN1CCCCC1)c1c2c(ccc1)C(=O)CC2                        | Klein <i>et al</i> , 2002              |
| GPV0479     | 0.18  | O(CC(O)CN1CCCCC1)c1ccc(cc1C(=O)C)C                          | Langer <i>et al</i> , 2004             |
| GPV0390     | 0.08  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccc(cc2)C(=O)C)cc1                | Langer <i>et al</i> , 2004             |
| GPV0385     | 0.11  | O(CC(O)CN1CCCCC1)c1cc(ccc1)C(=O)C                           | Langer <i>et al</i> , 2004             |
| GPV0384     | 0.01  | O1CCN(CC1)CC(O)COc1ccc(cc1)C(=O)C                           | Langer <i>et al</i> , 2004             |
| GPV0374     | 1.37  | O1CCN(CC1)CC(O)COc1ccccc1C(=O)CCc1c2c(ccc1)<br>cccc2        | Langer <i>et al</i> , 2004             |
| GPV0360     | 0.61  | O(CC(O)CN(Cc1ccccc1)C(=O)CC)c1ccccc1C(=O)CCc<br>1ccccc1     | Langer <i>et al</i> , 2004             |
| GPV0357     | 38.46 | O(CC(O)CN1CCN(CC1)c1ccccc1C)c1ccc(cc1C(=O)CC<br>c1ccccc1)C  | Langer <i>et al</i> , 2004             |
| GPV0354     | 16.95 | Clc1cc(C(=O)CCc2ccccc2)c(OCC(O)CN2CCN(CC2)c2<br>ccccc2C)cc1 | Langer <i>et al</i> , 2004             |
| GPV0337     | 26.32 | Clc1cc(ccc1Cl)CCC(=O)c1ccccc1OCC(O)CN1CCN(CC<br>1)c1ccccc1C |  |
| GPV0335     | 55.56 | O(CC(O)CN1CCN(CC1)c1ccccc1C)c1ccccc1C(=O)CC<br>c1ccc(cc1)C  | Langer <i>et al</i> , 2004             |
| GPV0319     | 6.80  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccccc2C(=O)c2ccccc<br>2)cc1       | Submitted in J Med Chem<br>(Chapter 6) |
| GPV0303     | 0.67  | O(CC(O)CN1CCCCC1)c1ccccc1C(O)CCc1ccccc1                     |  |
| GPV0302     | 1.14  | O(CC(O)CN1CCCCC1)c1ccccc1C(O)CCc1ccccc1                     | Chiba <i>et al</i> , 1998              |
| GPV0301     | 0.59  | O(CC(O)CN1CCCCC1)c1ccccc1C(O)CCc1ccccc1                     |  |
| GPV0238     | 1.39  | O(CC(O)CNCCC(c1ccccc1)c1ccccc1)c1ccccc1C(=O)C<br>Cc1ccccc1  |  |
| GPV0180     | 5.81  | O(CC(O)CN1CCCCC1)c1ccccc1C(=O)CCc1c2c(ccc1)c<br>ccc2        |  |
| GPV0157     | 0.89  | Fc1ccc(N2CCN(CC2)CC(O)COc2cc(ccc2)C(=O)CCc2c<br>cccc2)cc1   | Langer <i>et al</i> , 2004             |
| GPV0093     | 0.94  | o1c2c(ccccc2)c(Cc2ccccc2)c1C(O)CNCCC                        | Klein <i>et al</i> , 2002              |
| GPV0088     | 1.11  | O(CC(O)CN1CCCCC1)c1ccccc1C(O)CCc1ccccc1                     |  |
| GPV0073     | 1.03  | O(CC(O)CN1CCCCC1)c1ccc(cc1)C(=O)CCc1ccccc1                  | Klein <i>et al</i> , 2002              |
| GPV0057     | 0.27  | O1CCN(CC1)CC(O)COc1ccccc1C(=O)CCc1ccccc1                    | Langer <i>et al</i> , 2004             |
| GPV0045     | 0.48  | Fc1ccc(N2CCN(CC2)CC(O)COc2ccccc2C(=O)C)cc1                  |  |
| GPV0027     | 37.45 | O(CC(O)CN1CCN(CC1)c1ccccc1C)c1ccccc1C(=O)CC<br>c1ccccc1     | Langer <i>et al</i> , 2004             |
| GPV0025     | 3.90  | O(CC(O)CN1CCN(CC1)c1ccc(OC)cc1)c1ccccc1C(=O)<br>CCc1ccccc1  | Langer <i>et al</i> , 2004             |
| GPV0021     | 4.35  | O(CC(O)CN1CCN(CC1)c1ccccc1OC)c1ccccc1C(=O)C<br>Cc1ccccc1    | Langer <i>et al</i> , 2004             |
| GPV0005     | 1.67  | O(CC(O)CN1CCCCC1)c1ccccc1C(=O)CCc1ccccc1                    | Langer <i>et al</i> , 2004             |
| GPV0002     | 1.10  | O(CC(O)CN(CC)CC)c1ccccc1C(=O)CCc1ccccc1                     |  |
| GPV0001S    | 2.09  | O(CC(O)CNCCC)c1ccccc1C(=O)CCc1ccccc1                        |  |
| GPV0001     | 3.04  | O(CC(O)CNCCC)c1ccccc1C(=O)CCc1ccccc1                        |  |
| ERK_MS0840x | 0.74  | O1CCN(c2c1ccccc2)C(=O)NCCN(CCc1cc(OC)c(OC)cc<br>1)C         |  |

## APPENDIX A3

### SERT DATA SET

| #  | ChEMBL_ID | pIC50 | SMILES  |
|----|-----------|-------|---|
| 1  | 102223    | 9.09  | <chem>FC(F)(F)c1ccc(OC(CCNC)c2ccccc2)cc1</chem>                                 |
| 2  | 104584    | 2.96  | <chem>Oc1cc(ccc1O)CCN</chem>  |
| 3  | 106643    | 9.89  | <chem>Fc1ccc(cc1)[C@@H]1CCNC[C@H]1COc1cc2OCoc2cc1</chem>                        |
| 4  | 114193    | 10.10 | <chem>Br\C=C/c1ccc(cc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                          |
| 5  | 114239    | 6.41  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(Oc1ccccc1)=O)C</chem>                        |
| 6  | 114240    | 9.44  | <chem>lc1ccc(cc1)C1CC2NC(CC2)C1C(OC)=O</chem>                                   |
| 7  | 114241    | 8.38  | <chem>lc1ccc(cc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                                |
| 8  | 114255    | 8.87  | <chem>Br\C(\Br)=C\c1ccc(cc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                     |
| 9  | 114256    | 9.96  | <chem>\C=C/c1ccc(cc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                            |
| 10 | 114263    | 8.05  | <chem>O1CCNCC1C(Oc1ccccc1OCC)c1ccccc1</chem>                                    |
| 11 | 114309    | 10.30 | <chem>\C=C/c1ccc(cc1)C1CC2NC(CC2)C1C(OC)=O</chem>                               |
| 12 | 114353    | 10.10 | <chem>Br\C(\Br)=C\c1ccc(cc1)C1CC2NC(CC2)C1C(OC)=O</chem>                        |
| 13 | 114367    | 8.80  | <chem>Fc1ccc(cc1)C1(OCc2c1ccc(c2)C#N)CCCN(C)C</chem>                            |
| 14 | 114378    | 10.40 | <chem>Br\C=C/c1ccc(cc1)C1CC2NC(CC2)C1C(OC)=O</chem>                             |
| 15 | 126465    | 6.70  | <chem>Clc1cc(N2CCN(CC2)CCCN2N=C(N(CCOc3ccccc3)C2=O)CC)ccc1</chem>               |
| 16 | 127278    | 8.05  | <chem>O(C)c1ccc(cc1)C(CN(C)C)C1(O)CCCC1</chem>                                  |
| 17 | 127714    | 9.28  | <chem>Brc1cc(CO)c(Sc2ccccc2CN(C)C)cc1</chem>                                    |
| 18 | 128068    | 9.60  | <chem>S(c1ccccc1CN(C)C)c1ccc(cc1N)C</chem>                                      |
| 19 | 128220    | 6.08  | <chem>lc1cc(N)c(Sc2ccccc2CN(C)C)cc1</chem>                                      |
| 20 | 135586    | 5.68  | <chem>O(C(c1ccccc1)c1ccccc1)CCN1CCN(CC1)CCc1ccccc1</chem>                       |
| 21 | 146076    | 6.48  | <chem>O(C(=O)c1ccccc1)C1C[C@@H]2[N@](C(CC2)C1C(OC)=O)C</chem>                   |
| 22 | 146385    | 7.30  | <chem>Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1ccccc2</chem>                              |
| 23 | 149842    | 4.19  | <chem>O(C(=O)C(C1NCCCC1)c1ccccc1)C</chem>                                       |
| 24 | 153616    | 9.54  | <chem>Clc1cc(ccc1Cl)[C@@H]1CC[C@H](NC)c2c1ccccc2</chem>                         |
| 25 | 166246    | 6.69  | <chem>O(C(CCNC)c1ccccc1)c1ccccc1OC</chem>                                       |
| 26 | 166727    | 7.92  | <chem>Clc1cc(ccc1Cl)[C@H]1CC2N(C(CC2)[C@H]1C(OC)=O)C</chem>                     |
| 27 | 185619    | 7.13  | <chem>O(C)c1cc2CCN(Cc2cc1)C1CCC(CC1)c1c2cc(ccc2[nH]c1)C#N</chem>                |
| 28 | 185656    | 8.39  | <chem>O(C)c1cc2CCN(Cc2cc1)C1CCC(CC1)c1c2cc(ccc2[nH]c1)C#N</chem>                |
| 29 | 185717    | 8.35  | <chem>[nH]1cc(c2cc(ccc12)C#N)C1CCC(N2CCc3c(C2)ccccc3)CC1</chem>                 |
| 30 | 185740    | 7.04  | <chem>[nH]1cc(c2cc(ccc12)C#N)C1CCC(N2CCc3c(C2)ccccc3)CC1</chem>                 |
| 31 | 186089    | 8.05  | <chem>O(C(CCNC)c1ccccc1)c1ccccc1C</chem>  |
| 32 | 200496    | 5.62  | <chem>O(C(=O)[C@@H]1[C@@H]\2C[C@@H]3[N@@](C/C/2=C\c2ccccc2)C1CC3)C</chem>       |
| 33 | 200530    | 7.52  | <chem>Clc1cc(ccc1Cl)\C=C/1\ [C@H]2C[C@@H]3[N@@](C\1)C(CC3)[C@@H]2C(OC)=O</chem> |
| 34 | 200531    | 7.28  | <chem>O(C)c1ccc(cc1)\C=C/1\ [C@H]2C[C@@H]3[N@@](C\1)C(CC3)[C@@H]2C(OC)=O</chem> |
| 35 | 204260    | 7.41  | <chem>O(C(=O)C1C2N(C(CC1c1ccccc1)-c1ccccc1)CC2)C)C</chem>                       |
| 36 | 204525    | 8.91  | <chem>lc1cc(ccc1)C1CC2NC(CC2)C1C(OC)=O</chem>                                   |
| 37 | 204526    | 7.13  | <chem>s1cc(cc1)-c1cc(ccc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                       |
| 38 | 204735    | 6.05  | <chem>Fc1ccc(cc1)CCNCC1CCN(C1)CCOC(c1ccccc1)c1ccccc1</chem>                     |
| 39 | 204970    | 7.26  | <chem>o1ccccc1-c1cc(ccc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>                        |



APPENDIX A3

|    |        |       |   |
|----|--------|-------|---|
| 40 | 205064 | 6.68  | <chem>O(C(c1cccc1)c1cccc1)CCNCC1CCCN(C1)CCc1cccc1</chem>  |
| 41 | 205104 | 5.85  | <chem>O(C(c1cccc1)c1cccc1)CCN(C)C1CCN(CC1)CCc1cccc1</chem>                                      |
| 42 | 205443 | 8.02  | <chem>lc1cc(ccc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>  |
| 43 | 205523 | 6.28  | <chem>Fc1ccc(cc1)CCN(CC1CCCN(C1)CCOC(c1cccc1)c1cccc1)C</chem>                                   |
| 44 | 205555 | 6.46  | <chem>Fc1ccc(cc1)CCN1CC(CCC1)CNCCOC(c1cccc1)c1cccc1</chem>                                      |
| 45 | 205753 | 7.26  | <chem>O(C(c1cccc1)c1cccc1)CCN1CCC(NCCc2cccc2)CC1</chem>   |
| 46 | 205779 | 7.06  | <chem>O(C(c1cccc1)c1cccc1)CCN1CCC(N(CCCc2cccc2)C)CC1</chem>                                     |
| 47 | 205780 | 6.66  | <chem>Fc1ccc(cc1)CCNC1CCN(CC1)CCOC(c1cccc1)c1cccc1</chem>                                       |
| 48 | 205781 | 6.49  | <chem>O(C(c1cccc1)c1cccc1)CCN1CC(CCC1)CNCCc1cccc1</chem>  |
| 49 | 205788 | 5.67  | <chem>O(C(c1cccc1)c1cccc1)CCNC1CCN(CC1)Cc1cccc1</chem>  |
| 50 | 207150 | 6.99  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(=O)NCCCCCN<br/>C(=O)C1C2N(C(CC1c1ccc(Cl)cc1)CC2)C)C</chem>   |
| 51 | 207151 | 8.10  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(=O)NCCCCCCCN<br/>C(=O)C1C2N(C(CC1c1ccc(Cl)cc1)CC2)C)C</chem> |
| 52 | 207152 | 5.50  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(=O)NCCCCN(C=O)<br/>C1C2N(C(CC1c1ccc(Cl)cc1)CC2)C)C</chem>    |
| 53 | 207153 | 7.35  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(OC)=O)C</chem>   |
| 54 | 207666 | 6.45  | <chem>Clc1ccc(cc1)C1CC2N(C(CC2)C1C(=O)NCCCCN(C=O)<br/>C1C2N(C(CC1c1ccc(Cl)cc1)CC2)C)C</chem>    |
| 55 | 214406 | 4.76  | <chem>O1CCNCC1COc1cccc1OCC</chem>   |
| 56 | 233091 | 9.17  | <chem>lc1ccc(cc1)[C@H]1C[C@@H]2[N@@](C(CC2)[C@H]1C(OC)=O)C</chem>                               |
| 57 | 236836 | 4.00  | <chem>O(C)c1cc(ccc1OC)C1=NNC(=O)C2C1CC=CC2</chem>   |
| 58 | 260577 | 6.75  | <chem>O(C(=O)C1C2[N@](C(C[C@@H]1c1cccc1)CC2)C)C</chem>  |
| 59 | 261229 | 5.49  | <chem>Fc1ccc(cc1)C(OC1CC2N(C(C1)CC2)C)c1ccc(F)cc1</chem>  |
| 60 | 265576 | 3.21  | <chem>O(C(=O)c1cccc1-c1cc2c(cc1)cccc2)C</chem>  |
| 61 | 272541 | 9.48  | <chem>S(c1cccc1CN(C)C)c1ccc(cc1N)C(F)(F)F</chem>  |
| 62 | 272642 | 8.96  | <chem>S(c1cccc1CN(C)C)c1ccc(cc1N)C#N</chem>   |
| 63 | 272866 | 9.57  | <chem>Clc1cc(N)c(Sc2cccc2CN(C)C)cc1</chem>  |
| 64 | 272867 | 8.72  | <chem>S(c1cccc1CN(C)C)c1ccc(OC)cc1N</chem>  |
| 65 | 283149 | 7.92  | <chem>O(C)c1cc2c(cc(cc2)[C@H]2C3CC(CC3)[C@@H]2CNC)cc1</chem>                                    |
| 66 | 283190 | 8.44  | <chem>N(C[C@H]1C2CC(CC2)[C@@H]1c1cc2c(cc1)cccc2)C</chem>  |
| 67 | 283225 | 8.64  | <chem>N(C[C@H]1C2CC(CC2)[C@@H]1c1cc2c(cc1)cccc2)(C)C</chem>                                     |
| 68 | 283226 | 7.28  | <chem>N(CC1C2CC(CC2)C1c1c2c(ccc1)cccc2)(C)C</chem>  |
| 69 | 283269 | 8.55  | <chem>N(CC1C2CC(CC2)C1c1cc2c(cc1)cccc2)(C)C</chem>  |
| 70 | 283270 | 8.24  | <chem>N(C[C@@H]1C2CC(CC2)[C@H]1c1cc2c(cc1)cccc2)(C)C</chem>                                     |
| 71 | 283772 | 7.24  | <chem>Clc1cc(ccc1Cl)[C@H]1C2CCC(CC2)[C@H]1CN(C)C</chem>   |
| 72 | 283813 | 8.31  | <chem>N(CC1C2CC(CC2)C1c1cc2c(cc1)cccc2)C</chem>   |
| 73 | 283814 | 7.60  | <chem>O(C)c1cc2c(cc(cc2)[C@H]2C3CC(CC3)[C@@H]2CN(C)C)cc1</chem>                                 |
| 74 | 283891 | 8.59  | <chem>Fc1cc2c(cc(cc2)[C@H]2C3CC(CC3)[C@@H]2CN(C)C)cc1</chem>                                    |
| 75 | 283892 | 7.60  | <chem>Clc1cc(ccc1)[C@H]1C2CCC(CC2)[C@@H]1CN(C)C</chem>  |
| 76 | 283893 | 7.33  | <chem>O(C)c1cc2c(cc(cc2)[C@H]2C3CC(CC3)[C@@H]2CNC)cc1</chem>                                    |
| 77 | 283894 | 8.36  | <chem>Fc1cc2c(cc(cc2)[C@H]2C3CC(CC3)[C@@H]2CNC)cc1</chem>                                       |
| 78 | 286129 | 11.00 | <chem>S(C)c1ccc(cc1)C1c2c(C3N(C1)CCC3)cccc2</chem>  |
| 79 | 286845 | 10.01 | <chem>lc1cc(CO)c(Sc2cccc2CN(C)C)cc1</chem>  |
| 80 | 292607 | 2.51  | <chem>s1cccc1[C@H](Oc1c2c(ccc1)c(OS(=O)(=O)[O-])c(OC)cc2)CCNC.[Na+]</chem>                      |

## APPENDIX A3

|     |        |      |  |
|-----|--------|------|--|
| 81  | 292608 | 2.84 | s1cccc1[C@@H](Oc1c2c(cc(OC3OC(C(O)=O)C(O)C(O)C3O)cc2)ccc1)CCNC                           |
| 82  | 292634 | 4.19 | s1cccc1[C@@H](Oc1c2c(cccc2)c(O)cc1)CCNC  |
| 83  | 292635 | 5.97 | s1cccc1[C@@H](Oc1c2c(cc(O)cc2)ccc1)CCNC  |
| 84  | 292643 | 2.00 | s1cccc1[C@@H](Oc1c2c(cccc2)c(OC2OC(C(O)=O)C(O)C(O)C2O)cc1)CCNC                           |
| 85  | 292644 | 4.50 | s1cccc1[C@@H](Oc1c2c(ccc1)c(O)c(O)cc2)CCNC   |
| 86  | 292867 | 5.44 | S(c1cccc1CN(C)C)c1ccc(OC)cc1N  |
| 87  | 292899 | 3.58 | s1cccc1[C@@H](Oc1c2c(ccc1)c(O)c(OC)cc2)CCNC  |
| 88  | 292900 | 5.02 | s1cccc1[C@@H](Oc1c2c(ccc1)c(O)ccc2)CCNC  |
| 89  | 292908 | 9.10 | s1cccc1[C@@H](Oc1c2c(ccc1)ccc2)CCNC  |
| 90  | 293317 | 2.00 | s1cccc1[C@@H](Oc1c2c(ccc1)c(OC1OC(C(O)=O)C(O)C(O)C1O)c(OC1OC(C(O)=O)C(O)C(O)C1O)cc2)CCNC |
| 91  | 293318 | 2.00 | s1cccc1[C@@H](Oc1c2c(ccc1)c(OC1OC(C(O)=O)C(O)C(O)C1O)c(OC)cc2)CCNC                       |
| 92  | 299712 | 5.82 | O(C(c1cccc1)c1cccc1)CCN(CCCN(CCCc1cccc1)C)C  |
| 93  | 299713 | 6.94 | Fc1ccc(cc1)C(OCCN1CCN(CC1=O)CCc1cccc1)c1ccc(F)cc1  |
| 94  | 299920 | 6.90 | Fc1ccc(cc1)C(OCCNCCCNCc1cccc1)c1ccc(F)cc1  |
| 95  | 299921 | 5.94 | Fc1ccc(cc1)C(OCCNCCCNC(=O)Cc1cccc1)c1ccc(F)cc1   |
| 96  | 299924 | 6.28 | Fc1ccc(cc1)C(OCCNCCCNC(=O)Cc1cccc1)c1ccc(F)cc1   |
| 97  | 299926 | 6.11 | Fc1ccc(cc1)C(OCCN(CCN(CCOc1ccc(F)cc1)c1ccc(F)cc1)C)C)c1ccc(F)cc1                         |
| 98  | 299953 | 5.23 | O=C1N(CCN(C1)CCOC(c1cccc1)c1cccc1)CCc1cccc1  |
| 99  | 300006 | 6.26 | Fc1ccc(cc1)CCN(CCN(CCOc1cccc1)c1cccc1)C)C  |
| 100 | 300010 | 6.36 | O(C(c1cccc1)c1cccc1)CCN(CCCN(CCCc1cccc1)C)C  |
| 101 | 300046 | 6.25 | Fc1ccc(cc1)C(OCCNCCNC(=O)Cc1cccc1)c1ccc(F)cc1  |
| 102 | 300049 | 6.45 | Fc1ccc(cc1)C(OCCNCCNC(=O)Cc1ccc(F)cc1)c1ccc(F)cc1  |
| 103 | 300050 | 6.23 | BrC1ccc(cc1)CC(=O)N\C=C\NCCOC(c1ccc(F)cc1)c1ccc(F)cc1                                    |
| 104 | 318085 | 8.64 | OC1(N2C(=NCC2)c2c1cccc2)c1cc2c(cc1)cccc2   |
| 105 | 318244 | 7.96 | OC1(N2C(=NCC2)c2c1ccc1c2cccc1)c1cc2c(cc1)cccc2   |
| 106 | 318269 | 7.59 | Clc1ccc(cc1)C1N2C(=NCC2)c2c1cccc2  |
| 107 | 318272 | 8.26 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1ccc1c2cccc1   |
| 108 | 318273 | 6.65 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1cccc2   |
| 109 | 318282 | 7.41 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1cccc2   |
| 110 | 318306 | 7.42 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1cccc2F  |
| 111 | 318401 | 5.75 | OC1(N2C(=NCC2)c2c1cccc2)c1cccc1  |
| 112 | 318763 | 6.22 | Clc1ccc(cc1)C1(O)N2[C@@H]3[C@@H](N=C2c2c1cccc2)CCCC3                                     |
| 113 | 318824 | 5.70 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1c(OC)ccc2   |
| 114 | 318825 | 5.87 | Clc1cc2c(cc1Cl)C=1N(CCN=1)C2(O)c1ccc(Cl)cc1  |
| 115 | 318833 | 7.25 | O(C)c1ccc(cc1)C1(O)N2C(=NCC2)c2c1cccc2   |
| 116 | 318853 | 7.11 | OC1(N2C(=NCC2)c2c1cccc2)c1ccc(O)cc1  |
| 117 | 318935 | 6.06 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1cccc2OC   |
| 118 | 318936 | 5.99 | Clc1ccc(cc1)C1(O)N2C(=NCC2)c2c1cc(OC)cc2   |
| 119 | 319035 | 5.00 | O(C(=O)C=1C2N(C(CC=1c1cc(C(=O)C)c(cc1)C(=O)C)CC2)C)C                                     |
| 120 | 319036 | 6.77 | O(C(=O)C1C2N(C(CC1c1cc(C(=O)C)c(cc1)C(=O)C)CC2)C)C                                       |
| 121 | 319219 | 5.00 | Oc1cc(ccc1O)C1CC2N(C(CC2)C1C(OC)=O)C   |
| 122 | 319255 | 5.00 | Oc1cc(ccc1O)C=1CC2N(C(CC2)C=1C(OC)=O)C   |

APPENDIX A3

|     |        |       |   |
|-----|--------|-------|---|
| 123 | 319532 | 7.34  | O(C)c1cc(ccc1OC)C1CC2N(C(CC2)C1C(OC)=O)C                        |
| 124 | 337830 | 8.98  | S(c1ccccc1CN(C)C)c1ccc(cc1N)CF                                  |
| 125 | 343756 | 6.13  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](NC(C)(C)C)c2c1cccc2                |
| 126 | 343759 | 5.33  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(C(C)(C)C)C)c2c1cccc2             |
| 127 | 343761 | 7.36  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(C(C)C)C)c2c1cccc2                |
| 128 | 343778 | 7.89  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(CC)C)c2c1cccc2                    |
| 129 | 343779 | 6.70  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(CCC)C)c2c1cccc2                   |
| 130 | 343780 | 7.14  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(C(C)C)C)c2c1cccc2                 |
| 131 | 343789 | 6.74  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(Cc2cccc2)C)c2c1cccc2             |
| 132 | 343790 | 8.62  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(C)C)c2c1cccc2                     |
| 133 | 343910 | 10.22 | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(C)C)c2c1cccc2                    |
| 134 | 343943 | 5.00  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(C(C)(C)C)C)c2c1cccc2              |
| 135 | 343975 | 7.03  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](NC(C)C)c2c1cccc2                   |
| 136 | 344059 | 7.42  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NC)c2c1cccc2                        |
| 137 | 344092 | 6.42  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](NCc2cccc2)c2c1cccc2                |
| 138 | 344099 | 8.30  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](NC)c2c1cccc2                       |
| 139 | 344195 | 7.04  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(CCC)C)c2c1cccc2                  |
| 140 | 344220 | 6.89  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NC(C)C)c2c1cccc2                    |
| 141 | 344221 | 6.30  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NCCC)c2c1cccc2                      |
| 142 | 344222 | 8.33  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NCC)c2c1cccc2                       |
| 143 | 344393 | 5.68  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NCc2cccc2)c2c1cccc2                 |
| 144 | 344418 | 5.77  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](NC(C)(C)C)c2c1cccc2                 |
| 145 | 344452 | 7.80  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](N(CC)C)c2c1cccc2                   |
| 146 | 344453 | 8.70  | Clc1cc(ccc1Cl)[C@@H]1C[C@@H](NCC)c2c1cccc2                      |
| 147 | 344483 | 7.77  | lc1cc2c(cc1)[C@@H](C[C@H]2N(C)C)c1cc(Cl)c(Cl)cc1                |
| 148 | 344484 | 6.52  | Clc1cc(ccc1Cl)[C@@H]1C[C@H](N(Cc2cccc2)C)c2c1cccc2              |
| 149 | 347074 | 6.04  | n1ccc(cc1)C(Cc1cccc1)c1cccc1                                    |
| 150 | 347479 | 5.94  | n1ccc(cc1)C(c1cccc1)c1cccc1                                     |
| 151 | 347480 | 5.00  | n1ccc(cc1)Cc1cccc1  |
| 152 | 347503 | 5.00  | n1cc(ccc1)-c1c2c(cc3c1cccc3)cccc2                               |
| 153 | 347506 | 5.00  | n1ccc(cc1)-c1c2c(cc3c1cccc3)cccc2                               |
| 154 | 347584 | 4.46  | n1cccc1C(c1cccc1)c1cccc1  |
| 155 | 347585 | 4.90  | n1cc(ccc1)C(c1cccc1)c1cccc1                                     |
| 156 | 388042 | 10.30 | l\C=C/c1ccc(cc1)[C@H]1CC2NC(CC2)[C@H]1C(OC)=O                   |
| 157 | 388504 | 8.94  | l\C=C\c1ccc(cc1)[C@H]1CC2NC(CC2)[C@H]1C(OC)=O                   |
| 158 | 397259 | 5.53  | Fc1ccc(cc1)CC1CCCN(C1)C[C@@H]1CCCC[C@H]1NC(=O)Nc1cc(ccc1)C(=O)C |
| 159 | 401194 | 5.58  | Clc1cc2c(CCNC[C@@H]2C)cc1                                       |
| 160 | 394158 | 8.82  | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1OC                          |
| 161 | 402750 | 7.83  | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1C                           |
| 162 | 402799 | 6.18  | O1CCNC[C@H]1[C@@H](Oc1cccc1OCC)c1cccc1                          |
| 163 | 403062 | 7.44  | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1                            |
| 164 | 409927 | 8.09  | s1c2c(cccc2O[C@@H](CCNC)c2cccc2)cc1                             |
| 165 | 409928 | 8.72  | s1c2c(cccc2O[C@H](CCNC)c2cccc2)cc1                              |

## APPENDIX A3

|     |        |       |   |
|-----|--------|-------|---|
| 166 | 409946 | 8.92  | s1c2c(cc(O[C@H](CCNC)c3cccc3)cc2)cc1                                    |
| 167 | 409970 | 9.00  | s1c2c(cccc2OC(CCNC)c2ccc(F)cc2)cc1                                      |
| 168 | 409971 | 8.82  | s1cccc1[C@@H](Oc1c2sccc2ccc1)CCNC                                       |
| 169 | 409992 | 8.27  | s1c2c(cc1)c(O[C@H](CCNC)c1cccc1)ccc2                                    |
| 170 | 409993 | 9.30  | s1c2c(cc1)c(O[C@@H](CCNC)c1cccc1)ccc2                                   |
| 171 | 410036 | 9.30  | s1c2cc(O[C@@H](CCNC)c3cccc3)ccc2cc1                                     |
| 172 | 410070 | 9.30  | s1c2c(cccc2OC(CCNC)c2ccc(F)cc2)cc1                                      |
| 173 | 410071 | 8.77  | s1c2c(cc(O[C@@H](CCNC)c3cccc3)cc2)cc1                                   |
| 174 | 410104 | 8.62  | s1c2c(cccc2OC(CCNC)c2cc(ccc2)C)cc1                                      |
| 175 | 410120 | 8.96  | s1c2c(cccc2OC(CCNC)c2cc(F)ccc2)cc1                                      |
| 176 | 410137 | 8.66  | s1c2c(cccc2OC(CCNC)c2cc(OC)ccc2)cc1                                     |
| 177 | 410201 | 9.15  | s1c2c(cc1)c(O[C@H](CCNC)c1cccc1)ccc2F                                   |
| 178 | 410207 | 9.05  | s1c2c(cc1)c(O[C@@H](CCNC)c1cccc1)ccc2F                                  |
| 179 | 410219 | 8.49  | s1c2c(cccc2OC(CCNC)c2cc(ccc2)C(F)(F)F)cc1                               |
| 180 | 413195 | 11.05 | S(C)c1ccc(cc1)[C@@H]1c2c([C@H]3N(C1)CCC3)cccc2                          |
| 181 | 413405 | 9.92  | lc1cc(CO)c(Oc2ccccc2CN(C)C)cc1  |
| 182 | 413592 | 7.31  | lc1cc(N)c(cc1)Cc1cccc1CN(C)C  |
| 183 | 413759 | 9.43  | lc1cc(N)c(Oc2ccccc2CN(C)C)cc1   |
| 184 | 416660 | 8.69  | S(c1cccc1CNC)c1ccc(cc1N)CC  |
| 185 | 416692 | 8.80  | S(c1cccc1CN(C)C)c1cccc1N  |
| 186 | 416693 | 8.95  | S(c1cccc1CN(C)C)c1ccc(cc1N)C=C  |
| 187 | 416848 | 7.30  | S(c1cccc1CN(C)C)c1ccc(cc1N)CCO  |
| 188 | 416849 | 8.07  | S(c1cccc1CN(C)C)c1ccc(cc1N)CCF  |
| 189 | 416956 | 7.08  | S(c1cccc1CNC)c1cccc1N   |
| 190 | 417023 | 9.31  | S(c1cccc1CN(C)C)c1ccc(cc1N)CC   |
| 191 | 190937 | 9.24  | S(c1cccc1CN(C)C)c1ccc(cc1N)CO   |
| 192 | 417171 | 7.03  | S(c1cccc1CN(C)C)c1ccc(cc1N)CCCO   |
| 193 | 417172 | 8.07  | S(c1cccc1CN(C)C)c1ccc(cc1N)CCCF   |
| 194 | 417571 | 7.82  | S(c1cccc1CNC)c1ccc(cc1N)CO  |
| 195 | 420276 | 8.84  | S(c1cccc1CN(C)C)c1ccc(F)cc1N  |
| 196 | 420351 | 7.15  | S(c1cccc1CNCC)c1ccc(cc1N)C  |
| 197 | 420383 | 7.15  | S(c1cccc1CNCCF)c1ccc(cc1N)C   |
| 198 | 420472 | 7.09  | S(c1cccc1CNCC=C)c1ccc(cc1N)C  |
| 199 | 420566 | 7.52  | S(c1cccc1CN)c1ccc(cc1N)C  |
| 200 | 420593 | 6.75  | S(c1cccc1CN(C)C)c1cc(F)c(cc1N)C   |
| 201 | 420665 | 5.66  | S(c1cccc1CNCCC)c1ccc(cc1N)C   |
| 202 | 420692 | 6.29  | S(c1cccc1CNCCCF)c1ccc(cc1N)C  |
| 203 | 423713 | 5.41  | Fc1ccc(cc1)C[C@@H]1CCCN(C1)C[C@@H]1CCCC[C@H]1NC(=O)Nc1ccc(cc1)-c1nnnn1C |
| 204 | 424590 | 5.63  | s1c(C(=O)c(nc1NC(=O)N[C@@H]1CCCC[C@H]1CN1C[C@@H](CCC1)Cc1ccc(F)cc1)C    |
| 205 | 196468 | 9.74  | [nH]1cc(c2cc(ccc12)C#N)[C@H]1C[C@@H]1CN(C)C                             |
| 206 | 429090 | 8.08  | FC(F)(F)c1ccc(OC(CCN(C)C)c2cccc2)cc1                                    |
| 207 | 429873 | 7.00  | O=C1N(c2c(CC1CCNC)cccc2)c1cccc1   |

APPENDIX A3

|     |        |      |   |
|-----|--------|------|---|
| 208 | 429874 | 7.00 | O=C1N(c2c(CC1CCCNC)cccc2)c1cccc1  |
| 209 | 429890 | 7.00 | Fc1cc(N2c3c(CC(CCCNC)C2=O)cccc3)ccc1  |
| 210 | 429891 | 7.00 | Clc1cc(N2c3c(CC(CCCNC)C2=O)cccc3)ccc1   |
| 211 | 429913 | 7.00 | O=C1N(c2c(CC1CCNC)cccc2)c1cccc1   |
| 212 | 430009 | 6.80 | O=C1N(c2c(CC1(CCCNC)CCC)cccc2)c1ccc(cc1)C   |
| 213 | 430015 | 7.00 | Fc1cc2CC(CCCNC)C(=O)N(c2cc1)c1ccc(cc1)C   |
| 214 | 430023 | 7.00 | Clc1cc2CC(CCCNC)C(=O)N(c2cc1)c1ccc(cc1)C  |
| 215 | 430031 | 6.89 | Clc1cc2CC(CCCNC)(CC)C(=O)N(c2cc1)c1ccc(cc1)C  |
| 216 | 430070 | 7.00 | O=C1N(c2c(CC1(CCCNC)C)cccc2)c1cccc1   |
| 217 | 430076 | 7.00 | O=C1N(c2c(CC1(CCCNC)C)cccc2)c1cc(ccc1)C   |
| 218 | 430077 | 7.00 | Fc1ccc(N2c3c(CC(CCCNC)C2=O)cccc3)cc1  |
| 219 | 430083 | 7.00 | Clc1ccc(N2c3c(CC(CCCNC)C2=O)cccc3)cc1   |
| 220 | 430084 | 7.00 | O=C1N(c2c(CC1CCCNC)cccc2)c1ccc(cc1)C  |
| 221 | 430089 | 7.00 | O=C1N(c2c(CC1CCCNC)cccc2)c1ccc(cc1)CC   |
| 222 | 430090 | 7.00 | FC(F)(F)c1ccc(N2c3c(CC(CCCNC)C2=O)cccc3)cc1   |
| 223 | 430096 | 7.00 | O=C1N(c2c(CC1(CCCNC)C)cccc2)c1ccc(cc1)C   |
| 224 | 430097 | 7.00 | O=C1N(c2c(CC1(CCCNC)C)cccc2)c1ccc(cc1)C(C)C   |
| 225 | 430102 | 7.00 | Clc1cc(N2c3c(CC(CCCNC)C2=O)cccc3)ccc1Cl   |
| 226 | 430108 | 7.00 | Fc1cc(N2c3c(CC(CCCNC)(C)C2=O)cccc3)ccc1F  |
| 227 | 430109 | 7.00 | Fc1cc(N2c3c(CC(CCCNC)(C)C2=O)cccc3)cc(F)c1  |
| 228 | 430115 | 7.00 | Clc1cc(N2c3c(CC(CCCNC)(C)C2=O)cccc3)cc(Cl)c1  |
| 229 | 430116 | 7.00 | O=C1N(c2c(CC1(CCCNC)CC)cccc2)c1cccc1  |
| 230 | 430123 | 7.00 | O=C1N(c2c(CC1(CCCNC)CCC)cccc2)c1cccc1   |
| 231 | 430124 | 7.00 | O=C1N(c2c(CC1(CCCNC)CC)cccc2)c1ccc(cc1)C  |
| 232 | 431966 | 6.15 | Clc1cc(ccc1Cl)[C@H]1CC2N(C(CC2)[C@H]1C(OC)=O)CCNC(=O)C1CCC(CC1)CN<br>S(=O)(=O)c1cc(S(=O)(=O)[O-]<br>])c(cc1)C=1c2c(OC=3C=1C=C/C(=[N+](/CC)\CC)/C=3)cc(N(CC)CC)cc2 |
| 233 | 432003 | 6.64 | Clc1cc(ccc1Cl)[C@H]1CC2N(C(CC2)[C@H]1C(OC)=O)CCNC(=O)CCCN1CCCc2c1<br>cc1c(C=C3C(=C\C(=[N+](/C)\C)\C=C3)C1(C)C)c2  |
| 234 | 432046 | 5.83 | Clc1cc(ccc1Cl)[C@H]1CC2N(C(CC2)[C@H]1C(OCCNC(=O)CCCCNS(=O)(=O)c1<br>cc(S(=O)(=O)[O-]<br>])c(cc1)C=1c2c(OC=3C=1C=C/C(=[N+](/CC)\CC)/C=3)cc(N(CC)CC)cc2)=O)C        |
| 235 | 432087 | 6.41 | Clc1cc(ccc1Cl)[C@H]1CC2N(C(CC2)[C@H]1C(OC)=O)CCNC(=O)CCCCNS(=O)(<br>=O)c1cc(S(=O)(=O)[O])c(cc1)C=1c2c(OC=3C=1C=C/C(=[N+](/CC)\CC)/C=3)cc(<br>N(CC)CC)cc2          |
| 236 | 435202 | 5.21 | O(C)c1cc2c(cc1)cccc2N1CCN(CC1)CCCCN1N=CC(=O)N(C)C1=O  |
| 237 | 438679 | 5.48 | o1cccc1-c1ccc(cc1)C(=O)C(N1CCCC1)CCC  |
| 238 | 438680 | 6.20 | Clc1cc(ccc1Cl)C(=O)C(CCC)CN1CCCC1   |
| 239 | 438681 | 6.09 | Clc1cc(ccc1Cl)C(=O)C(N1CCCC1)CC   |
| 240 | 438753 | 5.54 | O=C(C(N1CCCC1)CC(C)C)c1ccc(cc1)C  |
| 241 | 438754 | 5.00 | O=C(C(N1CCCC1)CC=C)c1ccc(cc1)C  |
| 242 | 438762 | 5.97 | Clc1cc(ccc1Cl)C(=O)C(N1CCCC1)CC=C   |
| 243 | 438769 | 5.00 | OC(C(N1CCCC1)CCC)c1ccc(cc1)C  |
| 244 | 438770 | 6.45 | O(C(=O)c1cccc1)[C@H]1CC2N(C(CC2)[C@H]1OC(OC)=O)C  |
| 245 | 438973 | 5.48 | s1cccc1-c1ccc(cc1)C(=O)C(N1CCCC1)CCC  |
| 246 | 438988 | 6.08 | Brc1ccc(cc1)C(=O)C(N1CCCC1)CCC  |

## APPENDIX A3

|     |        |      |  |
|-----|--------|------|--|
| 247 | 438989 | 6.52 | <chem>lc1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                               |
| 248 | 438990 | 5.85 | <chem>lc1cc(ccc1)C(=O)C(N1CCCC1)CCC</chem>                               |
| 249 | 439000 | 5.39 | <chem>O=C(C(CCC)CN1CCCC1)c1ccc(cc1)C</chem>                              |
| 250 | 439012 | 5.00 | <chem>O=C(C(N1CCCC1)CCC)c1ccc(cc1)C#N</chem>                             |
| 251 | 439024 | 5.42 | <chem>O=C(C(N1CCCC1)CCC)c1ccc(cc1)C</chem>                               |
| 252 | 439025 | 5.65 | <chem>O=C([C@@H](N1CCCC1)CCC)c1ccc(cc1)C</chem>                          |
| 253 | 439032 | 5.00 | <chem>O=C([C@H](N1CCCC1)CCC)c1ccc(cc1)C</chem>                           |
| 254 | 439033 | 5.00 | <chem>O=C(C(N1CCCC1)CCC)c1cccc1</chem>                                   |
| 255 | 439034 | 5.00 | <chem>Fc1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                               |
| 256 | 439040 | 5.00 | <chem>O=C(C(N1CCCC1)CCC)c1ccc(NC(=O)C)cc1</chem>                         |
| 257 | 439041 | 6.02 | <chem>FC(F)(F)c1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                        |
| 258 | 439042 | 5.17 | <chem>O=C(C(N1CCCC1)CCC)c1ccc(cc1)C#CC</chem>                            |
| 259 | 439047 | 5.43 | <chem>O=C(C(N1CCCC1)CCC)c1cccc1C</chem>                                  |
| 260 | 439048 | 5.00 | <chem>O=C(C(N1CCCC1)CCC)c1ccc(cc1)-c1n(ccc1)C</chem>                     |
| 261 | 439054 | 5.00 | <chem>O=C(CC(N1CCCC1)CC)c1ccc(cc1)C</chem>                               |
| 262 | 439055 | 5.41 | <chem>Clc1cc(ccc1Cl)C(=O)CC(N1CCCC1)CC</chem>                            |
| 263 | 439057 | 5.00 | <chem>O=C(C(N1CCCC1)C#C)c1ccc(cc1)C</chem>                               |
| 264 | 439144 | 5.29 | <chem>Clc1cc(ccc1Cl)C(=O)C(NCCCC)CCC</chem>                              |
| 265 | 439151 | 5.61 | <chem>Clc1cc(ccc1Cl)C(=O)C(N1CCCC1)CCC</chem>                            |
| 266 | 439317 | 5.00 | <chem>OCc1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                              |
| 267 | 439318 | 5.00 | <chem>Oc1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                               |
| 268 | 439319 | 5.61 | <chem>O=C(C(N1CCCC1)CCC)c1ccc([N+](=O)[O-])cc1</chem>                    |
| 269 | 439331 | 7.48 | <chem>O=C(C(N1CCCC1)CCC)c1cc2c(cc1)cccc2</chem>                          |
| 270 | 439339 | 5.23 | <chem>O=C(C(N1CCCC1)CCC)c1cc(ccc1)C</chem>                               |
| 271 | 439502 | 5.39 | <chem>O(C)c1ccc(cc1)C(=O)C(N1CCCC1)CCC</chem>                            |
| 272 | 439503 | 6.70 | <chem>Clc1cc(ccc1Cl)C(=O)C(N1CCCC1)CCC</chem>                            |
| 273 | 439511 | 5.00 | <chem>Oc1cc(ccc1O)C(=O)C(N1CCCC1)CCC</chem>                              |
| 274 | 439512 | 5.13 | <chem>O(C)c1cc(ccc1OC)C(=O)C(N1CCCC1)CCC</chem>                          |
| 275 | 440274 | 6.64 | <chem>O(C(c1cccc1)c1cccc1)CCC1CCN(CC1)CC[C@H](O)c1cccc1</chem>           |
| 276 | 440293 | 7.37 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)CC[C@H](O)c1cccc1)c1ccc(F)cc1</chem>     |
| 277 | 440437 | 6.35 | <chem>O(C(c1cccc1)c1cccc1)CCC1CCN(CC1)C[C@H](O)Cc1cccc1</chem>           |
| 278 | 440444 | 7.39 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](O)Cc1cccc1)c1ccc(F)cc1</chem>     |
| 279 | 440448 | 6.54 | <chem>O(C(c1cccc1)c1cccc1)CCC1CCN(CC1)C[C@H](O)Cc1cccc1</chem>           |
| 280 | 440451 | 7.09 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](O)Cc1cccc1)c1ccc(F)cc1</chem>     |
| 281 | 440543 | 7.66 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)CC[C@H](O)c1cccc1)c1ccc(F)cc1</chem>     |
| 282 | 440552 | 7.24 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)CC[C@H](C)c1cccc1)c1ccc(F)cc1</chem>     |
| 283 | 440558 | 7.12 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)CC[C@H](C)c1cccc1)c1ccc(F)cc1</chem>     |
| 284 | 440593 | 6.90 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)CC=1CCc2c(C=1)cccc2)c1ccc(F)cc1</chem>   |
| 285 | 440601 | 7.80 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H]1C[C@H]1c1cccc1)c1ccc(F)cc1</chem> |
| 286 | 440657 | 6.64 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](N)Cc1cccc1)c1ccc(F)cc1</chem>     |
| 287 | 440716 | 6.64 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](N)Cc1cccc1)c1ccc(F)cc1</chem>     |
| 288 | 440756 | 7.72 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](O)CCc1cccc1)c1ccc(F)cc1</chem>    |
| 289 | 440757 | 7.42 | <chem>Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](O)CCc1cccc1)c1ccc(F)cc1</chem>    |

APPENDIX A3

|     |         |       |  |
|-----|---------|-------|--|
| 290 | 440909  | 6.87  | Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H]1C[C@@H]1c1ccccc1)c1ccc(F)cc1  |
| 291 | 441010  | 7.28  | Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@@H](Cc1ccccc1)C)c1ccc(F)cc1      |
| 292 | 441017  | 6.91  | Fc1ccc(cc1)C(OCCC1CCN(CC1)C[C@H](Cc1ccccc1)C)c1ccc(F)cc1       |
| 293 | 441025  | 6.55  | O(C(c1ccccc1)c1ccccc1)CCC1CCN(CC1)CC[C@@H](O)c1ccccc1          |
| 294 | 442514  | 6.00  | O(CC1CC1)Cc1nc2N3[C@H](Cc2cc1)CNC[C@H]3C                       |
| 295 | 444371  | 5.00  | O(C)c1cc(N2CCN(CC2)CCCCN2N=CC(=O)N(C)C2=O)ccc1                 |
| 296 | 1140708 | 9.32  | N1CCC(N(Cc2cc(ccc2)C#N)CC(C)C)CC1                              |
| 297 | 447015  | 8.82  | N1CCC(N(Cc2ccccc2)CC(C)C)CC1                                   |
| 298 | 447022  | 8.66  | N1CCC(N(Cc2cc(ccc2)C)CC(C)C)CC1                                |
| 299 | 447063  | 8.72  | Fc1ccc(cc1)CN(CC(C)C)C1CCNCC1                                  |
| 300 | 447116  | 9.02  | Clc1cc(Cl)ccc1CN(CC(C)C)C1CCNCC1                               |
| 301 | 447118  | 9.52  | Fc1cc(C(F)(F)F)c(cc1)CN(CC(C)C)C1CCNCC1                        |
| 302 | 447203  | 8.37  | FC(F)(F)c1ccccc1CN(CC(C)C)C1CCNCC1                             |
| 303 | 447342  | 8.80  | Clc1ccccc1CN(CC(C)C)C1CCNCC1                                   |
| 304 | 447415  | 8.85  | FC(F)(F)c1ccccc1CN(CC(C)C)C1CCNCC1                             |
| 305 | 447459  | 8.51  | Fc1cc(F)ccc1CN(CC(C)C)C1CCNCC1                                 |
| 306 | 447469  | 8.24  | Fc1cc(ccc1CN(CC(C)C)C1CCNCC1)C(F)(F)F                          |
| 307 | 449376  | 5.84  | o1c(ccc1Oc1cc2c(cc1)CC2(C)C)C(=O)Nc1c(OC)nc(nc1OC)NCCCN1CCOCC1 |
| 308 | 455079  | 5.85  | O1CCNC[C@@H]1[C@H](Oc1ccccc1OCC)c1ccccc1                       |
| 309 | 457974  | 8.11  | N12CC(C3C1CCC2CC3c1ccc(cc1)C)=C                                |
| 310 | 458003  | 9.55  | Brc1ccc(cc1)C1C2C3N(CC2=C)C(C1)CC3                             |
| 311 | 458004  | 9.09  | Clc1ccc(cc1)C1C2C3N(CC2=C)C(C1)CC3                             |
| 312 | 458091  | 10.05 | lc1ccc(cc1)C1C2C3N(CC2=C)C(C1)CC3                              |
| 313 | 458120  | 7.01  | O(C(=O)C=C\1/C2C3N(C/1)C(CC2c1ccc(cc1)C)CC3)C                  |
| 314 | 458121  | 8.25  | Brc1ccc(cc1)C1C\2C3N(C/C/2=C/C(OC)=O)C(C1)CC3                  |
| 315 | 458122  | 7.73  | Clc1ccc(cc1)C1C\2C3N(C/C/2=C/C(OC)=O)C(C1)CC3                  |
| 316 | 459546  | 6.54  | lc1ccc(cc1)[C@H]1CCN(C[C@H]1C(OC)=O)C                          |
| 317 | 461477  | 9.70  | Br\C=C\c1cc(ccc1)[C@H]1CC2NC(CC2)[C@H]1C(OC)=O                 |
| 318 | 465711  | 5.41  | Clc1ccc(cc1)C(CC(C)C)C1NCCCC1                                  |
| 319 | 465712  | 5.49  | Clc1ccc(cc1)C(CC1CCCC1)C1NCCCC1                                |
| 320 | 465713  | 6.05  | Clc1ccc(cc1)C(CC1CCCC1)C1NCCCC1                                |
| 321 | 465721  | 5.38  | O(C)c1cc(ccc1)C(CC(C)C)C1NCCCC1                                |
| 322 | 465722  | 5.48  | N1CCCCC1C(CC(C)C)c1ccc(cc1)C(C)C                               |
| 323 | 465723  | 5.11  | O=C(C(C1NCCCC1)c1ccccc1)C                                      |
| 324 | 465731  | 5.22  | Clc1ccc(cc1)C(C(OC)=O)C1NCCCC1                                 |
| 325 | 465732  | 5.04  | Clc1ccc(cc1)C(C)C1NCCCC1                                       |
| 326 | 465740  | 5.37  | Clc1ccc(cc1)C(Cc1ccccc1)C1NCCCC1                               |
| 327 | 465741  | 6.19  | Clc1ccc(cc1)C(CCCc1ccccc1)C1NCCCC1                             |
| 328 | 465742  | 5.82  | Clc1ccc(cc1)C(CCCc1ccccc1)C1NCCCC1                             |
| 329 | 465749  | 5.32  | Clc1ccc(cc1)C(C(CC)CC)C1NCCCC1                                 |
| 330 | 465750  | 5.34  | Clc1ccc(cc1)C(C1CCCC1)C1NCCCC1                                 |
| 331 | 465751  | 5.49  | Clc1cc(ccc1)C(CC(C)C)C1NCCCC1                                  |
| 332 | 465758  | 5.11  | Clc1ccc(cc1)C(CC)C1NCCCC1                                      |

## APPENDIX A3

|     |        |      |  |
|-----|--------|------|--|
| 333 | 465759 | 5.49 | Clc1ccc(cc1)C(CCC)C1NCCCC1   |
| 334 | 465760 | 2.00 | Clc1ccc(cc1)C(C(C)C)C1NCCCC1   |
| 335 | 465761 | 5.32 | Clc1ccc(cc1)C(CCCC)C1NCCCC1  |
| 336 | 465768 | 6.35 | Clc1cc(ccc1Cl)C(CC(C)C)C1NCCCC1  |
| 337 | 465769 | 5.89 | Clc1cc(ccc1Cl)C(CC(C)C)C1N(CCCC1)C   |
| 338 | 465770 | 6.02 | O(C)c1ccc(cc1)C(CC(C)C)C1NCCCC1  |
| 339 | 465777 | 5.60 | Clc1cc(ccc1Cl)C(C(OC)=O)C1NCCCC1   |
| 340 | 465778 | 5.74 | Clc1cc(ccc1Cl)C(CCC)C1NCCCC1   |
| 341 | 465779 | 5.96 | Clc1cc(ccc1Cl)C(CCCC)C1NCCCC1  |
| 342 | 465787 | 5.37 | Clc1ccc(cc1)C(CC(C)C)C1NCCCC1  |
| 343 | 465788 | 5.66 | Clc1ccc(cc1)C(CCCCC)C1NCCCC1   |
| 344 | 465789 | 5.30 | Clc1ccc(cc1)C(CCC(C)C)C1NCCCC1   |
| 345 | 467469 | 7.40 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCO)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C     |
| 346 | 467470 | 6.67 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCCO)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C    |
| 347 | 467474 | 7.96 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC(=O)C)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C |
| 348 | 467475 | 6.15 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC(=O)N)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C |
| 349 | 467480 | 7.31 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCF)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C     |
| 350 | 467481 | 6.52 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC#C)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C    |
| 351 | 467486 | 6.48 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC(F)F)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C  |
| 352 | 467491 | 5.26 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)C(=O)C)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C  |
| 353 | 467496 | 6.65 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)C(=O)CC)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C |
| 354 | 467497 | 7.72 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)C(N)=N)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C  |
| 355 | 467508 | 7.27 | O=C(Nc1cc(ccc1)C#N)NCCC[C@@H]1NCC[C@H](C1)Cc1ccccc1                        |
| 356 | 467530 | 7.38 | Fc1ccc(cc1)C[C@H]1C[C@@H](NCC1)CCCNC(=O)Nc1cc(ccc1)C(=O)C                  |
| 357 | 467531 | 7.74 | Fc1ccc(cc1)C[C@H]1C[C@@H](NCC1)CCCNC(=O)Nc1cc(ccc1)C(=O)C                  |
| 358 | 467536 | 8.46 | Fc1ccc(cc1)C[C@H]1C[C@@H](NCC1)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C          |
| 359 | 467537 | 6.24 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)C)CCCNC(=O)Nc1cc(ccc1)C(=O)C               |
| 360 | 467542 | 6.53 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCC)CCCNC(=O)Nc1cc(ccc1)C(=O)C             |
| 361 | 467543 | 6.93 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC1CC1)CCCNC(=O)Nc1cc(ccc1)C(=O)C          |
| 362 | 467553 | 6.66 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CC=C)CCCNC(=O)Nc1cc(ccc1)C(=O)C            |
| 363 | 467554 | 7.30 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCO)CCCNC(=O)Nc1cc(ccc1)C(=O)C             |
| 364 | 467559 | 6.78 | Fc1ccc(cc1)C[C@H]1C[C@@H](N(CC1)CCC)CCCNC(=O)Nc1cc(cc(c1)C(=O)C)C(=O)C     |
| 365 | 476460 | 6.73 | O1CCNC[C@H]1[C@@H](Oc1ccccc1OC)c1ccccc1                                    |
| 366 | 482415 | 5.00 | s1cc(nc1CF)C#Cc1cc(cc(F)c1)C#N   |
| 367 | 482978 | 5.57 | Clc1ccc(cc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1c1sc(cn1)-c1ccccc1)C            |
| 368 | 482979 | 6.19 | Clc1ccc(cc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1c1sc(cn1)-c1ccc(F)cc1)C         |
| 369 | 482980 | 5.62 | Clc1ccc(cc1)-c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(Cl)cc1)CC2)C     |



APPENDIX A3

|     |        |       |   |
|-----|--------|-------|---|
| 370 | 482981 | 5.24  | Brc1ccc(cc1)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(Cl)cc1)CC2)C          |
| 371 | 482982 | 5.82  | Brc1cc(ccc1)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(Cl)cc1)CC2)C          |
| 372 | 482983 | 5.74  | Clc1ccc(cc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1c1sc(nc1)-<br>c1ccc([N+](=O)[O-])cc1)C   |
| 373 | 482984 | 6.13  | Clc1ccc(cc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1c1sc(nc1)-<br>c1cc([N+](=O)[O-])ccc1)C   |
| 374 | 482985 | 5.28  | Clc1ccc(cc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1c1sc(nc1)-c1ccc(OC)cc1)C                 |
| 375 | 482988 | 5.57  | Clc1ccc(cc1)C(CCCCC)C1NCCCC1  |
| 376 | 482989 | 5.44  | Clc1ccc(cc1)C(CCC(C)C)C1NCCCC1  |
| 377 | 482990 | 5.76  | s1c(cnc1[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C)-<br>c1ccc(F)cc1            |
| 378 | 482991 | 5.88  | Clc1ccc(cc1)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C           |
| 379 | 482992 | 5.73  | Brc1ccc(cc1)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C           |
| 380 | 482993 | 6.06  | Brc1cc(ccc1)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C           |
| 381 | 482995 | 5.79  | s1c(cnc1[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C)-<br>c1ccc([N+](=O)[O-])cc1 |
| 382 | 482996 | 6.24  | s1c(cnc1[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C)-<br>c1cc([N+](=O)[O-])ccc1 |
| 383 | 483010 | 5.44  | s1c(cnc1[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C)-<br>c1ccc(OC)cc1           |
| 384 | 483011 | 5.53  | Clc1cc(ccc1Cl)-<br>c1sc(nc1)[C@@H]1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C         |
| 385 | 483012 | 5.11  | s1c(cnc1C1[C@@H]2N([C@H](C[C@@H]1c1ccc(cc1)C)CC2)C)-<br>c1cc(OC)c(OC)cc1            |
| 386 | 488386 | 10.10 | I\C=C/c1ccc(cc1)C1CC2NC(CC2)C1C(OCCF)=O   |
| 387 | 488452 | 10.10 | I\C=C/c1ccc(cc1)C1CC2NC(CC2)C1C(OCCF)=O   |
| 388 | 492076 | 8.22  | lc1ccc(cc1)[C@H]1C[C@@H]2N([C@@H](CC2)[C@H]1C(OC)=O)C                               |
| 389 | 492077 | 5.73  | Clc1cc(ccc1Cl)[C@H]1CCN(C[C@H]1C(OC)=O)C  |
| 390 | 492083 | 4.57  | O(C(=O)[C@@H]1CN(CC[C@@H]1c1ccccc1)C)C  |
| 391 | 492087 | 4.33  | O(C(=O)[C@@H]1CN(CC[C@@H]1c1ccccc1)C)C  |
| 392 | 492091 | 5.55  | lc1ccc(cc1)[C@H]1CCN(C[C@H]1C(OC)=O)C   |
| 393 | 492092 | 6.36  | O(C(=O)C=1CNCCC=1c1cc2c(cc1)cccc2)CC  |
| 394 | 492093 | 6.66  | Clc1cc(ccc1Cl)C=1CCNCC=1C(OCC)=O  |
| 395 | 492095 | 5.52  | O(C(=O)C=1CNCCC=1c1cc([N+](=O)[O-])ccc1)CC  |
| 396 | 492097 | 3.10  | Clc1cc(ccc1Cl)C=1CCCC=1C(OC)=O  |
| 397 | 492099 | 2.55  | Clc1ccc(cc1)C=1CCCC=1C(OC)=O  |
| 398 | 492101 | 2.65  | Clc1cc(ccc1)C=1CCCC=1C(OC)=O  |
| 399 | 492102 | 3.38  | O(C(=O)C=1CCCC=1c1cc2c(cc1)cccc2)C  |
| 400 | 492103 | 2.57  | O(C)c1ccc(cc1)C=1CCCC=1C(OC)=O  |
| 401 | 492104 | 2.84  | O(C(=O)C=1CCCC=1c1ccc(cc1)C(C)(C)C)C  |
| 402 | 492105 | 2.42  | FC(F)(F)Oc1ccc(cc1)C=1CCCC=1C(OC)=O   |
| 403 | 492107 | 3.53  | O(C(=O)C=1CCCC=1c1ccc(cc1)C(=O)C)C  |
| 404 | 492109 | 2.71  | O(C(=O)C=1CCCC=1c1cc(N)ccc1)C   |
| 405 | 492111 | 2.94  | O(C(=O)C=1CCCC=1c1ccc([N+](=O)[O-])cc1)C  |
| 406 | 492113 | 3.64  | Clc1cc(ccc1Cl)C=1CCCC=1C(OC)=O  |

## APPENDIX A3

|     |         |       |   |
|-----|---------|-------|---|
| 407 | 492115  | 2.73  | Clc1ccc(cc1)C=1CCCC=1C(OC)=O                            |
| 408 | 492117  | 2.93  | Clc1cc(ccc1)C=1CCCC=1C(OC)=O                            |
| 409 | 492119  | 3.96  | O(C(=O)C=1CCCC=1c1cc2c(cc1)cccc2)C                      |
| 410 | 492120  | 3.08  | O(C(=O)C=1CCCC=1c1ccc(cc1)-c1cccc1)C                    |
| 411 | 492122  | 2.65  | O(C(=O)C=1CCCC=1c1ccc(cc1)C(C)(C)C                      |
| 412 | 492125  | 2.33  | FC(F)(F)Oc1ccc(cc1)C=1CCCC=1C(OC)=O                     |
| 413 | 492126  | 3.18  | O(C(=O)C=1CCCC=1c1cccc1)C                               |
| 414 | 492128  | 2.30  | FC(F)(F)c1cc(ccc1)C=1CCCC=1C(OC)=O                      |
| 415 | 492130  | 2.33  | FC(F)(F)c1ccc(cc1)C=1CCCC=1C(OC)=O                      |
| 416 | 492132  | 3.24  | Oc1ccc(cc1)C=1CCCC=1C(OC)=O                             |
| 417 | 492134  | 4.54  | O(C(=O)C=1CCCC=1c1cc([N+](=O)[O-])ccc1)C                |
| 418 | 492139  | 3.83  | Clc1cc(ccc1Cl)C=1CCCC=1C(OC)=O                          |
| 419 | 492141  | 3.15  | Clc1ccc(cc1)C=1CCCC=1C(OC)=O                            |
| 420 | 492143  | 3.33  | Clc1cc(ccc1)C=1CCCC=1C(OC)=O                            |
| 421 | 492144  | 3.27  | O(C(=O)C=1CCCC=1c1cc2c(cc1)cccc2)C                      |
| 422 | 492149  | 4.49  | O(C(=O)C=1CCCC=1c1ccc(cc1)-c1cccc1)C                    |
| 423 | 492151  | 3.42  | Clc1cc(ccc1Cl)C=1CCCC=1\C=C\C                           |
| 424 | 492375  | 3.41  | O(C(=O)[C@@H]1CCCC[C@@H]1c1ccc(cc1)-c1cccc1)C           |
| 425 | 492376  | 2.04  | O(C(=O)[C@@H]1CCCC[C@@H]1c1ccc(cc1)C(C)(C)C             |
| 426 | 492379  | 3.22  | FC(F)(F)Oc1ccc(cc1)[C@H]1CCCC[C@H]1C(OC)=O              |
| 427 | 492380  | 2.82  | O(C(=O)[C@@H]1CCCC[C@@H]1c1cccc1)C                      |
| 428 | 492382  | 2.72  | FC(F)(F)c1cc(ccc1)[C@H]1CCCC[C@H]1C(OC)=O               |
| 429 | 492385  | 3.73  | FC(F)(F)c1ccc(cc1)[C@H]1CCCC[C@H]1C(OC)=O               |
| 430 | 492386  | 4.14  | O(C(=O)[C@@H]1CCCC[C@@H]1c1ccc(cc1)-c1cccc1)C           |
| 431 | 492389  | 2.74  | O(C(=O)[C@@H]1CCCC[C@@H]1c1ccc(cc1)C(C)(C)C             |
| 432 | 492394  | 2.98  | O(C(=O)C=1CCCC=1c1ccc(cc1)C(C)(C)C                      |
| 433 | 496256  | 7.62  | Fc1cc(ccc1C)[C@@H]1C[C@@H]2N[C@H](CC2)[C@H]1C(OC)=O     |
| 434 | 496260  | 7.55  | O(C(=O)[C@@H]1[C@@H]2N[C@@H](C[C@H]1c1ccc(cc1)C)CC2)C   |
| 435 | 496262  | 8.68  | lc1ccc(cc1)[C@@H]1C[C@@H]2N[C@H](CC2)[C@H]1C(OC)=O      |
| 436 | 496264  | 8.72  | O(C(=O)[C@@H]1[C@@H]2N[C@@H](C[C@H]1c1ccc(cc1)C=C)CC2)C |
| 437 | 1153074 | 8.74  | O(C(=O)[C@@H]1[C@@H]2N[C@@H](C[C@H]1c1ccc(cc1)C#C)CC2)C |
| 438 | 499913  | 9.10  | S(C)c1ccc(Oc2ccc(cc2CNC)C#CCN2CCOCC2)cc1                |
| 439 | 499923  | 8.72  | S(C)c1ccc(Oc2ccc(cc2CNC)C#CCN2CCOCC2)cc1C               |
| 440 | 499942  | 9.05  | S(C)c1ccc(Oc2ccc(cc2CNC)C#CCN2CCC(F)CC2)cc1             |
| 441 | 499962  | 8.96  | S(C)c1ccc(Oc2ccc(cc2CNC)C#CCN(C)C2CC2)cc1               |
| 442 | 505795  | 9.60  | S(c1ccc(OCCF)cc1CN(C)c1cccc1N                           |
| 443 | 505799  | 10.00 | S(c1ccc(OCCF)cc1CN(C)c1ccc(F)cc1N                       |
| 444 | 1144290 | 10.52 | Clc1cc(N)c(Sc2ccc(OCCF)cc2CN(C)c1                       |
| 445 | 505805  | 10.30 | Brc1cc(N)c(Sc2ccc(OCCF)cc2CN(C)c1                       |
| 446 | 505806  | 8.85  | S(c1ccc(OCCF)cc1CN(C)c1cccc1N                           |
| 447 | 505809  | 9.02  | S(c1ccc(OCCF)cc1CN(C)c1ccc(F)cc1N                       |
| 448 | 505811  | 9.82  | Brc1cc(N)c(Sc2ccc(OCCF)cc2CN(C)c1                       |
| 449 | 253184  | 8.96  | S(c1ccc(OCCO)cc1CN(C)c1cccc1N                           |

APPENDIX A3

|     |         |      |  |
|-----|---------|------|--|
| 450 | 138354  | 8.92 | S(c1ccc(OCCO)cc1CN(C)C)c1ccc(F)cc1N                                  |
| 451 | 1147706 | 8.89 | S(c1ccc(OCCCO)cc1CN(C)C)c1ccccc1N                                    |
| 452 | 505816  | 9.68 | Brc1cc(N)c(Sc2ccc(OCCCO)cc2CN(C)C)cc1                                |
| 453 | 505817  | 8.30 | S(c1ccc(OC)cc1CN(C)C)c1ccccc1N                                       |
| 454 | 505819  | 7.77 | S(c1ccc(O)cc1CN(C)C)c1ccccc1N  |
| 455 | 510296  | 8.38 | S(C)c1ccc(cc1)C1c2c(cc(OCCCN3CCC(F)CC3)cc2)CN(C1)C                   |
| 456 | 510302  | 8.42 | S(C)c1ccc(cc1)C1c2c(cc(OCCCN3CCOCC3)cc2)CN(C1)C                      |
| 457 | 510304  | 7.89 | O1C[C@@H](N(CC1)CCCOc1cc2c(cc1)[C@@H](CN(C2)C)c1ccc(OC)cc1)C         |
| 458 | 512300  | 8.08 | O(CCCN1CCCC1)c1cc2[C@@H]3N(C[C@H](c2cc1)c1ccnc1)CCC3                 |
| 459 | 516686  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)N1CCOCC1                                 |
| 460 | 516688  | 3.28 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)N1CCCC1                                  |
| 461 | 516696  | 3.52 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)N(CC)CC                                  |
| 462 | 516697  | 3.00 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)N1CCCC1                                  |
| 463 | 516708  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)NCCF                                     |
| 464 | 516709  | 3.27 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)N(OC)C                                   |
| 465 | 516710  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)NOC                                      |
| 466 | 516711  | 3.96 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)NCC1sccc1                                |
| 467 | 516712  | 4.62 | Clc1cc(ccc1Cl)C=1CCCC=1C(=O)NCC1occc1                                |
| 468 | 516713  | 3.00 | Clc1cc(ccc1Cl)C=1CCCC=1C(OC1CCCC1)=O                                 |
| 469 | 516714  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(OC(C)C)=O                                   |
| 470 | 516715  | 4.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(OCCO)=O                                     |
| 471 | 516716  | 3.41 | Clc1cc(ccc1Cl)C=1CCCC=1C(OCCF)=O                                     |
| 472 | 516717  | 3.00 | Clc1cc(ccc1Cl)C=1CCCC=1C(OCC(F)(F)F)=O                               |
| 473 | 516718  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(Oc1ccccc1)=O                                |
| 474 | 516720  | 3.30 | Clc1cc(ccc1Cl)C=1CCCC=1C(Oc1ccccc1C)=O                               |
| 475 | 516721  | 3.00 | Clc1cc(ccc1Cl)C=1CCCC=1C(Oc1cc(ccc1)C)=O                             |
| 476 | 516722  | 3.04 | Clc1cc(ccc1Cl)C=1CCCC=1C(Oc1sccc1)=O                                 |
| 477 | 516723  | 4.62 | Clc1cc(ccc1Cl)C=1CCCC=1C(Oc1occc1)=O                                 |
| 478 | 1138354 | 8.90 | S(c1ccc(F)cc1CN(C)C)c1ccc(cc1N)CO                                    |
| 479 | 241646  | 9.54 | Brc1cc(CN(C)C)c(Sc2ccc(cc2N)CO)cc1                                   |
| 480 | 519030  | 9.51 | Ic1cc(CN(C)C)c(Sc2ccc(cc2N)CO)cc1                                    |
| 481 | 521450  | 5.54 | Fc1ccc(cc1)C[C@@H]1CCCN(C1)C[C@@H]1CCCC[C@H]1NC(=O)Nc1cc(ccc1)C(=O)C |
| 482 | 521513  | 4.54 | Clc1cc(ccc1Cl)C=1CCCC=1C(OC1CCCC1)=O                                 |
| 483 | 527083  | 7.74 | Clc1cc(CC)c(Oc2cc(ncc2CNC)C)cc1                                      |
| 484 | 527889  | 6.13 | Fc1cc(OC)c(O[C@H]([C@H]2OCCNC2)c2ccccc2)cc1                          |
| 485 | 527891  | 6.96 | Clc1cc(OC)c(O[C@H]([C@H]2OCCNC2)c2ccccc2)cc1                         |
| 486 | 527892  | 7.96 | Clc1cc(OC)c(O[C@@H]([C@H]2OCCNC2)c2ccccc2)cc1                        |
| 487 | 527976  | 7.55 | Clc1cc(OC)c(OC(C2OCCNC2)c2ccccc2)cc1                                 |
| 488 | 528065  | 7.22 | Clc1cc(OCC)c(OC(C2OCCNC2)c2ccccc2)cc1                                |
| 489 | 528112  | 6.62 | Clc1ccccc(OC(C2OCCNC2)c2ccccc2)c1OCC                                 |
| 490 | 528327  | 7.06 | Clc1cc(OC)c(O[C@H]([C@H]2OCCNC2)c2ccc(cc2)C)cc1                      |
| 491 | 528376  | 7.24 | Clc1cc(OC)c(O[C@H]([C@H]2OCCNC2)c2cc(F)ccc2)cc1                      |

## APPENDIX A3

|     |        |      |   |
|-----|--------|------|---|
| 492 | 528378 | 7.02 | Clc1cc(OC)ccc1O[C@H]([C@H]1OCCNC1)c1cccc1             |
| 493 | 528429 | 7.70 | O1CCNC[C@@H]1[C@H](Oc1c2OCCc2ccc1)c1cccc1             |
| 494 | 528430 | 7.00 | Clc1c(O[C@H]([C@H]2OCCNC2)c2cccc2)cccc1Cl             |
| 495 | 528471 | 7.92 | Fc1c(O[C@@H]([C@@H]2OCCNC2)c2cccc2)cccc1F             |
| 496 | 528474 | 7.47 | Fc1cc(C)c(O[C@@H]([C@@H]2OCCNC2)c2cccc2)cc1           |
| 497 | 528475 | 5.47 | Fc1c(O[C@H]([C@H]2OCCNC2)c2cccc2)cccc1F               |
| 498 | 528520 | 6.59 | Clc1cc(F)ccc1O[C@H]([C@H]1OCCNC1)c1cccc1              |
| 499 | 528521 | 7.70 | Clc1cc(F)ccc1O[C@@H]([C@@H]1OCCNC1)c1cccc1            |
| 500 | 528522 | 6.73 | Fc1cc(C)c(O[C@H]([C@H]2OCCNC2)c2cccc2)cc1             |
| 501 | 533477 | 5.85 | O(c1cc(ncc1CN(C)C)C)c1cccc1OC                         |
| 502 | 533479 | 6.66 | O(c1cc(ncc1CNC)C)c1cccc1OCC                           |
| 503 | 533510 | 6.58 | O(c1cc(ncc1CNC)C)c1cccc1OC                            |
| 504 | 533512 | 6.36 | O(c1cc(ncc1CNC)C)c1cccc1CC                            |
| 505 | 533513 | 6.55 | O(c1cc(ncc1CN(C)C)C)c1cccc1CC                         |
| 506 | 533726 | 5.39 | O(c1cc(ncc1CN(C)C)C)c1cccc1Oc1cccc1                   |
| 507 | 533759 | 6.08 | O(c1cc(ncc1CNC)C)c1cccc1Oc1cccc1                      |
| 508 | 533799 | 5.55 | O(c1cc(ncc1CN(C)C)C)c1cccc1CCC                        |
| 509 | 533802 | 6.09 | O(c1cc(ncc1CNC)C)c1cccc1CCC                           |
| 510 | 533804 | 5.83 | O(c1cc(ncc1CN(C)C)C)c1cccc1C(C)C                      |
| 511 | 533834 | 6.62 | Clc1cccc1Oc1cc(ncc1CN(C)C)C                           |
| 512 | 533836 | 5.70 | O(c1cc(ncc1CNC)C)c1cccc1C(C)C                         |
| 513 | 533867 | 7.06 | Clc1cccc1Oc1cc(ncc1CNC)C                              |
| 514 | 533868 | 6.42 | S(C)c1cccc1Oc1cc(ncc1CN(C)C)C                         |
| 515 | 533895 | 6.88 | S(C)c1cccc1Oc1cc(ncc1CNC)C                            |
| 516 | 533897 | 6.63 | O(c1cc(ncc1CN(C)C)C)c1cccc1C                          |
| 517 | 533927 | 6.87 | O(c1cc(ncc1CNC)C)c1cccc1C                             |
| 518 | 533955 | 5.32 | O(c1cc(ncc1CN(C)C)C)c1cccc1OCC                        |
| 519 | 537760 | 8.05 | Clc1cc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)ccc1Cl     |
| 520 | 537761 | 8.00 | Clc1ccc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)cc1       |
| 521 | 537804 | 7.92 | S(C)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)cc1C    |
| 522 | 538179 | 7.77 | FC(F)(F)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)cc1 |
| 523 | 538226 | 6.63 | Fc1cc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)ccc1        |
| 524 | 538272 | 7.33 | Clc1cc(Oc2ccc(cc2CNC)C(=O)N2CCCN(CC2)C2CC2)ccc1       |
| 525 | 538273 | 6.64 | O(c1ccc(cc1CNC)C(=O)N1CCCN(CC1)C1CC1)c1cccc1          |
| 526 | 538399 | 8.09 | S(C)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C(C)C)cc1C     |
| 527 | 538451 | 8.30 | S(C)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C2CC2)cc1      |
| 528 | 538503 | 8.23 | S(C)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C(C)C)cc1      |
| 529 | 538546 | 6.81 | Clc1cc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C2CC2)ccc1        |
| 530 | 538630 | 7.55 | O(c1ccc(cc1CNC)C(=O)N1CCN(CC1)C(C)C)c1cc(OC)ccc1      |
| 531 | 538777 | 7.64 | Clc1ccc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C2CC2)cc1        |
| 532 | 538879 | 8.17 | Clc1cc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C2CC2)ccc1Cl      |
| 533 | 538924 | 8.20 | Clc1cc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C(C)C)ccc1Cl      |
| 534 | 539103 | 7.96 | S(C)c1ccc(Oc2ccc(cc2CNC)C(=O)N2CCN(CC2)C2CC2)cc1C     |

APPENDIX A3

|     |         |      |  |
|-----|---------|------|--|
| 535 | 540306  | 6.92 | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1OCCF   |
| 536 | 540307  | 6.30 | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1C(OC)=O  |
| 537 | 540308  | 6.13 | S([C@H]([C@H]1OCCNC1)c1cccc1)c1cccc1OCCF   |
| 538 | 540332  | 6.21 | O1CCNC[C@H]1[C@@H](Oc1cccc1C(OC)=O)c1cccc1   |
| 539 | 540334  | 7.04 | S(C)c1cccc1O[C@H]([C@H]1OCCNC1)c1cccc1   |
| 540 | 546745  | 6.56 | O(C(=O)c1cccc1)[C@H]1C[C@H]2N([C@H](CC2)[C@H]1C(OC)=O)C                            |
| 541 | 549148  | 7.20 | Clc1cc(ccc1Cl)C12C(C1)CNC2   |
| 542 | 554328  | 9.37 | I\C=C/c1cc(ccc1)C1CC2NC(CC2)C1C(OCCF)=O  |
| 543 | 554329  | 9.48 | Br\C=C/c1cc(ccc1)C1CC2NC(CC2)C1C(OCCF)=O   |
| 544 | 554400  | 9.59 | I\C=C/c1cc(ccc1)C1CC2NC(CC2)C1C(OCCCF)=O   |
| 545 | 554401  | 9.48 | Br\C=C/c1cc(ccc1)C1CC2NC(CC2)C1C(OCCCF)=O  |
| 546 | 554402  | 7.82 | I\C=C\c1cc(ccc1)C1CC2NC(CC2)C1C(OCCF)=O  |
| 547 | 1152751 | 8.68 | Fc1cc(CCC2OC(CC2)CCN)c(OC)cc1  |
| 548 | 558469  | 8.15 | Fc1cc(CCCC2OC(CC2)CCN)c(OC)cc1   |
| 549 | 558470  | 4.00 | BrCCCCCN1N=C([C@@H]2[C@@H](CC=CC2)C1=O)c1cc(OC)c(OC)cc1                            |
| 550 | 558471  | 6.81 | Fc1cc(CCC2OC(CC2)CCNCCCCCN2N=C([C@@H]3[C@@H](CC=CC3)C2=O)c2cc(OC)c(OC)cc2)c(OC)cc1 |
| 551 | 558472  | 6.71 | Fc1ccc(OC)cc1CCCC1OC(CC1)CCNCCCCCN1N=C([C@@H]2[C@@H](CC=CC2)C1=O)c1cc(OC)c(OC)cc1  |
| 552 | 558542  | 7.67 | Fc1cc(CCC2OC(CC2)CCNC)c(OC)cc1   |
| 553 | 572103  | 7.03 | O1CCNC[C@H]1[C@@H](Oc1cccc1C)c1cccc1   |
| 554 | 572104  | 6.37 | FCc1cccc1O[C@H]([C@H]1OCCNC1)c1cccc1   |
| 555 | 572105  | 6.33 | FCc1cccc1O[C@H]([C@H]1OCCNC1)c1cccc1   |
| 556 | 573109  | 7.30 | FC(F)(F)c1ccc(O[C@@H](CCN)c2cccc2)cc1  |
| 557 | 573110  | 7.59 | FC(F)(F)c1ccc(O[C@H](CCN)c2cccc2)cc1   |
| 558 | 573174  | 4.00 | BrCCCCCN1N=C(C2C(CC=CC2)C1=O)c1cc(OC)c(OC)cc1                                      |
| 559 | 573175  | 6.52 | FC(F)(F)c1ccc(O[C@@H](CCNCCCCCN2N=C(C3C(CC=CC3)C2=O)c2cc(OC)c(OC)cc2)c2cccc2)cc1   |
| 560 | 573176  | 7.19 | FC(F)(F)c1ccc(O[C@H](CCNCCCCCN2N=C(C3C(CC=CC3)C2=O)c2cc(OC)c(OC)cc2)c2cccc2)cc1    |
| 561 | 576087  | 5.98 | Clc1cc(F)ccc1[C@H]1C[C@@H]1CN  |
| 562 | 577102  | 5.42 | BrC1CCCC1[C@H]1C[C@@H]1CN  |
| 563 | 577104  | 5.57 | BrC1CCCC1C1CC1CN   |
| 564 | 577180  | 5.26 | NC[C@H]1C[C@@H]1c1cc(ccc1)C  |
| 565 | 577184  | 5.57 | Fc1ccc(cc1)[C@H]1C[C@@H]1CN  |
| 566 | 578777  | 5.22 | Fc1cc(ccc1)-c1ccc(cc1C#N)C(=O)Nc1nc(ccc1)C   |
| 567 | 579324  | 5.22 | s1cc(nc1C)C#Cc1cc(cc(F)c1)-c1cccnc1  |
| 568 | 579326  | 5.22 | s1cc(nc1C)C#Cc1cc(F)c(cc1)-c1cccnc1  |
| 569 | 579418  | 5.22 | O=C(Nc1nc(ccc1)C)c1cc(C#N)c(cc1)-c1cccc1   |
| 570 | 586074  | 4.28 | O(C)c1ccc(cc1)CCN(CCN1CCCC1)C  |
| 571 | 586075  | 5.15 | FC(F)(F)Oc1ccc(cc1)CCN(CCN1CCCC1)C   |
| 572 | 586930  | 5.42 | FC(F)(F)Oc1ccc(cc1)CCN(CCN1CCCC1)CC  |
| 573 | 586931  | 5.02 | FC(F)(F)Oc1ccc(cc1)CCN(CCN1CCCC1)C   |
| 574 | 586932  | 5.48 | FC(F)(F)Oc1ccc(cc1)CCN(CCN1CCCC1)CC  |
| 575 | 586933  | 4.00 | FC(F)(F)Oc1cccc1CCN(CCN1CCCC1)C  |

## APPENDIX A3

|     |        |      |  |
|-----|--------|------|--|
| 576 | 586934 | 4.00 | FC(F)(F)Oc1cccc1CCN(CCN1CCCC1)CC                     |
| 577 | 586935 | 4.00 | FC(F)(F)Oc1cccc1CCN(CCN1CCCC1)C                      |
| 578 | 587001 | 4.00 | FC(F)(F)Oc1cccc1CCN(CCN1CCCC1)CC                     |
| 579 | 587002 | 5.00 | FC(F)(F)Oc1cc(ccc1)CCN(CCN1CCCC1)C                   |
| 580 | 587003 | 5.00 | FC(F)(F)Oc1cc(ccc1)CCN(CCN1CCCC1)CC                  |
| 581 | 587004 | 5.00 | FC(F)(F)Oc1cc(ccc1)CCN(CCN1CCCC1)C                   |
| 582 | 587005 | 5.00 | FC(F)(F)Oc1cc(ccc1)CCN(CCN1CCCC1)CC                  |
| 583 | 589211 | 6.02 | O=C(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)C(C)C             |
| 584 | 589345 | 8.21 | FCCCOc1cc2CCN(Cc2cc1)C1CCC(CC1)c1c2cc(ccc2[nH]c1)C#N |
| 585 | 589346 | 6.49 | FCCCOc1cc2CCN(Cc2cc1)C1CCC(CC1)c1c2cc(ccc2[nH]c1)C#N |
| 586 | 589527 | 6.14 | O=C(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)CCC               |
| 587 | 589530 | 6.26 | O=C(N(Cc1cccc1C(C)C)[C@H]1CCNC1)C(C)C                |
| 588 | 589531 | 6.91 | O=C(N(Cc1cccc1C1CC1)[C@H]1CCNC1)C(C)C                |
| 589 | 589534 | 5.39 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)C           |
| 590 | 589535 | 5.25 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)CC          |
| 591 | 589536 | 5.25 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)CCC         |
| 592 | 589540 | 6.39 | O(CC)C(=O)N(Cc1cccc1-c1cccc1)[C@@H]1CCNC1            |
| 593 | 589541 | 5.27 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@@H]1CCNC1)C          |
| 594 | 589606 | 7.30 | S(C)c1cccc1CN(C(=O)C(C)C)[C@H]1CCNC1                 |
| 595 | 589607 | 6.64 | O(c1cccc1CN(C(=O)C(C)C)[C@H]1CCNC1)c1cccc1           |
| 596 | 589608 | 5.35 | O(C(=O)N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)C              |
| 597 | 589609 | 5.95 | O(CC)C(=O)N(Cc1cccc1-c1cccc1)[C@H]1CCNC1             |
| 598 | 589610 | 5.33 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)C(C)C       |
| 599 | 589611 | 5.28 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)C(F)(F)F    |
| 600 | 589612 | 5.54 | S(=O)(=O)(N(Cc1cccc1C(F)(F)F)[C@H]1CCNC1)C           |
| 601 | 589613 | 5.90 | S(C)c1cccc1CN(S(=O)(=O)C)[C@H]1CCNC1                 |
| 602 | 589614 | 5.57 | S(=O)(=O)(N(Cc1cccc1Oc1cccc1)[C@H]1CCNC1)C           |
| 603 | 589615 | 5.48 | S(=O)(=O)(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)N(C)C       |
| 604 | 589682 | 5.97 | N1C[C@@H](NCc2cccc2-c2cccc2)CC1                      |
| 605 | 589683 | 6.06 | O=CN(Cc1cccc1-c1cccc1)[C@H]1CCNC1                    |
| 606 | 589686 | 6.10 | O(C(C)C)C(=O)N(Cc1cccc1-c1cccc1)[C@H]1CCNC1          |
| 607 | 589687 | 6.12 | O(CC)C(=O)N(Cc1cccc1C(C)C)[C@H]1CCNC1                |
| 608 | 589688 | 6.37 | FC(F)(F)c1cccc1CN(C(OCC)=O)[C@H]1CCNC1               |
| 609 | 589690 | 6.35 | O(c1cccc1CN(C(OCC)=O)[C@H]1CCNC1)c1cccc1             |
| 610 | 589691 | 5.30 | O=C(N(Cc1cccc1-c1cccc1)[C@H]1CCNC1)N(C)C             |
| 611 | 609149 | 8.39 | Fc1cc2c([nH]cc2CCCN2CCc3cc(OC)ccc3C2)cc1             |
| 612 | 609150 | 8.19 | Fc1cc2c([nH]cc2CCCN2CCc3c(C2)ccc3OC)cc1              |
| 613 | 609205 | 7.94 | Brc1cc2c([nH]cc2CCCN2CCc3cc(OC)ccc3C2)cc1            |
| 614 | 609206 | 5.72 | Brc1cc2c([nH]cc2CCCN2CCc3c(C2)ccc3OC)cc1             |
| 615 | 609207 | 6.43 | lc1cc2c([nH]cc2CCCN2CCc3cc(OC)ccc3C2)cc1             |
| 616 | 609208 | 6.67 | lc1cc2c([nH]cc2CCCN2CCc3c(C2)ccc3OC)cc1              |
| 617 | 615167 | 7.89 | [nH]1c2cc(ccc2cc1)C1(CCN1)Cc1cccc1                   |
| 618 | 615170 | 7.64 | [nH]1cc(c2c1cccc2)C1(CCN1)Cc1cccc1                   |

## APPENDIX A3

|     |        |      |   |
|-----|--------|------|---|
| 619 | 615171 | 7.28 | [nH]1c2c(cc(cc2)C2(CCNC2)Cc2ncccc2)cc1          |
| 620 | 615172 | 7.82 | Fc1ccc(cc1)CC1(CCNC1)c1cc2c([nH]cc2)cc1         |
| 621 | 615242 | 7.47 | O(C)c1cc(ccc1)CC1(CCNC1)c1cc2c([nH]cc2)cc1      |
| 622 | 615243 | 7.89 | [nH]1c2c(cc(cc2)C2(CCNC2)CCc2cccc2)cc1          |
| 623 | 615244 | 8.40 | n1(c2c(cc(cc2)C2(CCNC2)Cc2cccc2)cc1)C           |
| 624 | 615245 | 8.40 | O(C(=O)c1[nH]c2c(cc(cc2)C2(CCNC2)Cc2cccc2)c1)CC |
| 625 | 615246 | 9.40 | [nH]1c2c(cc(cc2)C2(CCNC2)Cc2cccc2)cc1C#N        |
| 626 | 615247 | 8.70 | O=C(N)c1[nH]c2c(cc(cc2)C2(CCNC2)Cc2cccc2)c1     |
| 627 | 615307 | 7.89 | O=C(NC)c1[nH]c2c(cc(cc2)C2(CCNC2)Cc2cccc2)c1    |
| 628 | 615308 | 7.66 | O=C(N(C)C)c1[nH]c2c(cc(cc2)C2(CCNC2)Cc2cccc2)c1 |
| 629 | 615309 | 7.18 | [nH]1c2ncc(cc2cc1)C1(CCNC1)Cc1cccc1             |
| 630 | 615310 | 6.06 | [nH]1c2c(cc(cc2)C2(CCNC2)C)cc1                  |
| 631 | 615311 | 7.18 | [nH]1c2c(cc(cc2)C2(CCNC2)CCC)cc1                |
| 632 | 615312 | 7.80 | [nH]1c2c(cc(cc2)C2(CCNC2)CCCC)cc1               |
| 633 | 616227 | 8.05 | [nH]1cc(c2c1cccc2)C(CCNC)c1cccc1                |
| 634 | 616304 | 8.00 | [nH]1c2c(cc(cc2)C2(CCNC2)Cc2cccc2)cc1           |
| 635 | 616307 | 9.00 | [nH]1ncc2cc(ccc12)C1(CCNC1)Cc1cccc1             |
| 636 | 616308 | 8.15 | O=C(N(C)C)c1[nH]c2c(cc(cc2)C2(CCNC2)Cc2cccc2)c1 |
| 637 | 616309 | 7.39 | [nH]1cc(c2cc(ccc12)C1(CCNC1)Cc1cccc1)C#N        |
| 638 | 618133 | 7.37 | O1C(CCC1CN)c1cccc1                              |
| 639 | 618134 | 5.96 | Clc1ccc(cc1)C1OC(CC1)CN                         |
| 640 | 618199 | 5.95 | Brc1ccc(cc1)C1OC(CC1)CN                         |
| 641 | 618200 | 5.16 | O1C(CCC1CN)c1ccc(OC)cc1                         |
| 642 | 618201 | 5.00 | O1C(CCC1CN)c1ccc(cc1)C(C)(C)C                   |
| 643 | 618202 | 7.35 | O1C(CCC1CN)c1cc2c(cc1)cccc2                     |
| 644 | 618203 | 6.76 | O1C(CCC1CN)c1c2c(ccc1)cccc2                     |
| 645 | 618204 | 7.17 | O1C(CCC1CN)CCc1cccc1                            |
| 646 | 618205 | 7.74 | O1C(CCC1CN)CCc1c2c(ccc1)cccc2                   |
| 647 | 618206 | 5.67 | Fc1c(CCC2OC(CC2)CN)c(F)c(F)c(F)c1F              |
| 648 | 618279 | 5.34 | O1C(CCC1CN)CCc1cccnc1                           |
| 649 | 618280 | 5.42 | O1C(CCC1CCN)Cc1cccc1                            |
| 650 | 618281 | 6.16 | Fc1ccc(cc1C)CCC1OC(CC1)CCN                      |
| 651 | 618282 | 5.13 | O1C(CCC1CCN)CCc1ccc(O)cc1                       |
| 652 | 618283 | 5.00 | O1C(CCC1CCN)c1occc1                             |
| 653 | 618284 | 5.20 | Fc1cc([C@@H]2O[C@H](CC2)CN)c(OC)cc1             |
| 654 | 618285 | 5.16 | Fc1cc(C2OC(CC2)CN)c(OC)cc1                      |
| 655 | 618286 | 5.74 | Fc1cc([C@@H]2O[C@H](CC2)CCN)c(OC)cc1            |
| 656 | 618385 | 5.32 | Fc1cc([C@H]2O[C@H](CC2)CCN)c(OC)cc1             |
| 657 | 618386 | 6.72 | Fc1cc(C[C@@H]2O[C@H](CC2)CN)c(OC)cc1            |
| 658 | 618387 | 6.82 | Fc1cc(C[C@H]2O[C@H](CC2)CN)c(OC)cc1             |
| 659 | 618388 | 7.70 | Fc1cc(C[C@@H]2O[C@H](CC2)CCN)c(OC)cc1           |
| 660 | 618389 | 7.68 | Fc1cc(C[C@H]2O[C@H](CC2)CCN)c(OC)cc1            |
| 661 | 511500 | 8.49 | Fc1cc(CC[C@@H]2O[C@H](CC2)CN)c(OC)cc1           |

## APPENDIX A3

|     |         |      |   |
|-----|---------|------|---|
| 662 | 618391  | 7.66 | Fc1cc(CC[C@H]2O[C@H](CC2)CN)c(OC)cc1          |
| 663 | 1152805 | 8.68 | Fc1cc(CC[C@@H]2O[C@H](CC2)CCN)c(OC)cc1        |
| 664 | 618493  | 7.71 | Fc1cc(CC[C@H]2O[C@H](CC2)CCN)c(OC)cc1         |
| 665 | 618494  | 9.10 | Fc1cc(CCC[C@@H]2O[C@H](CC2)CN)c(OC)cc1        |
| 666 | 618495  | 8.55 | Fc1cc(CCC[C@H]2O[C@H](CC2)CN)c(OC)cc1         |
| 667 | 618498  | 8.40 | Fc1cc(CCCC[C@@H]2O[C@H](CC2)CN)c(OC)cc1       |
| 668 | 618499  | 8.70 | Fc1cc(CCCCC2OC(CC2)CCN)c(OC)cc1               |
| 669 | 618664  | 7.04 | O1C(CCC1CCN)CCc1c2c(cccc2)c(OC)cc1            |
| 670 | 618665  | 7.07 | O1C(CCC1CCN)CCc1c2c(ccc1OC)cccc2              |
| 671 | 647314  | 7.59 | O1c2cc(ccc2OC1)C1(O)N2C(=NCCC2)c2c1cccc2      |
| 672 | 647748  | 7.32 | O[C@H]([C@H](CN(C)C)c1cccc1)c1cccc1           |
| 673 | 647750  | 8.21 | O[C@H]([C@H](CN)c1cc2c(cc1)cccc2)C(C)(C)C     |
| 674 | 647768  | 7.40 | O[C@H]([C@H](CN(C)C)c1cccc1)C(C)(C)C          |
| 675 | 647773  | 5.52 | O[C@H]([C@H](CN)c1cccc1)c1cccc1               |
| 676 | 647796  | 8.92 | O[C@H]([C@H](CN(C)C)c1cc2c(cc1)cccc2)C(C)(C)C |
| 677 | 647886  | 8.07 | O[C@H]([C@H](CN(C)C)c1cccc1)c1c(cc(cc1C)C)C   |
| 678 | 652260  | 8.25 | O[C@H]([C@H](CN(C)C)c1cc2c(cc1)cccc2)c1cccc1  |
| 679 | 652273  | 8.21 | O[C@H]([C@H](CN)c1cc2c(cc1)cccc2)c1cccc1      |
| 680 | 652284  | 7.12 | O(C)c1ccc(cc1)[C@H]([C@H](O)C1CCCC1)CN        |
| 681 | 652309  | 5.96 | O[C@H]([C@H](CN)c1cccc1)C(C)(C)C              |
| 682 | 652326  | 7.52 | O(C)c1ccc(cc1)[C@H]([C@H](O)C1CCCC1)CN(C)C    |
| 683 | 652343  | 6.93 | O[C@H]([C@H](CN)c1cccc1)c1c(cc(cc1C)C)C       |
| 684 | 655551  | 6.68 | FC(F)(F)c1cccc1C(=O)N([C@H]1CCNC1)C1CCC1      |
| 685 | 655577  | 6.18 | O(c1cccc1C(=O)N(C)[C@H]1CCNC1)c1cccc1         |
| 686 | 655578  | 6.65 | O(c1cccc1C(=O)N([C@H]1CCNC1)C1CCC1)c1cccc1    |
| 687 | 655590  | 6.00 | [nH]1c2c(cc(N(C3CCNCC3)c3cccc3)cc2)cc1        |
| 688 | 655610  | 8.20 | [nH]1c2c(cc(N(Cc3cccc3)C3CCNCC3)cc2)cc1       |
| 689 | 657263  | 6.63 | O=C(N(CC1CCC1)[C@H]1CCNC1)c1cccc1CC           |
| 690 | 657280  | 7.80 | FC(F)(F)c1cccc1C(=O)N(CC1CC1)[C@H]1CCNC1      |
| 691 | 657281  | 6.95 | FC(F)(F)c1cccc1C(=O)N([C@H]1CCNC1)c1cccc1     |
| 692 | 657309  | 9.10 | [nH]1c2c(cc(N(Cc3cccc3)C3CCNCC3)cc2)cc1       |
| 693 | 657310  | 8.20 | [nH]1c2c(cc(N(Cc3cccc3)C3CCNCC3)cc2)cc1       |
| 694 | 658956  | 7.41 | Clc1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1           |
| 695 | 658965  | 7.21 | S(C)c1cccc1C(=O)N([C@H]1CCNC1)C1CCC1          |
| 696 | 658980  | 6.24 | FC(F)(F)c1cccc1C(=O)N(CC1CC1)[C@H]1CCNC1      |
| 697 | 659013  | 6.73 | FC(F)(F)c1cccc1CN(C(=O)C1CCC1)[C@H]1CCNC1     |
| 698 | 659028  | 8.90 | [nH]1ncc2cc(N(Cc3cccc3)C3CCNCC3)ccc12         |
| 699 | 659029  | 7.40 | [nH]1c2c(cc(N(C)C)c3cccc3)C3CCNCC3)cc2)cc1    |
| 700 | 659059  | 8.60 | [nH]1c2c(cc(N(Cc3ccc(cc3)C#N)C3CCNCC3)cc2)cc1 |
| 701 | 660749  | 6.18 | O=C(N(CC(C)C)[C@H]1CCNC1)c1cccc1-c1cccc1      |
| 702 | 660762  | 6.07 | O=C(N(CC1CCC1)[C@H]1CCNC1)c1cccc1C(C)C        |
| 703 | 660763  | 6.82 | O(c1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1)c1cccc1   |
| 704 | 660770  | 6.34 | FC(F)(F)c1cccc1C(=O)N(C(C)C)[C@H]1CCNC1       |



APPENDIX A3

|     |         |      |  |
|-----|---------|------|--|
| 705 | 660782  | 5.54 | O=C(N(CC1CC1)[C@H]1CCNC1)c1cccc1C(C)C                |
| 706 | 660783  | 6.18 | O=C(N([C@H]1CCNC1)C1CCCC1)c1cccc1C(C)C               |
| 707 | 660784  | 6.48 | O=C(N(CC1CCCC1)[C@H]1CCNC1)c1cccc1C(C)C              |
| 708 | 660798  | 5.78 | FC(F)(F)c1cccc1C(=O)N(CC1CCC1)[C@@H]1CCNC1           |
| 709 | 660828  | 8.70 | Fc1ccc(cc1)CN(C1CCNCC1)c1cc2c([nH]cc2)cc1            |
| 710 | 660836  | 9.60 | [nH]1c2c(cc(N(Cc3cc(ccc3)C#N)C3CCNCC3)cc2)cc1        |
| 711 | 662579  | 6.75 | S(C)c1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1                |
| 712 | 662580  | 6.46 | O(CC)c1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1               |
| 713 | 662588  | 6.10 | O=C(N(CC1CCC1)[C@H]1CCNC1)c1cccc1C1CCCC1             |
| 714 | 662602  | 6.11 | O=C(N([C@H]1CCNC1)C1CCC1)c1cccc1C(C)C                |
| 715 | 662631  | 7.60 | O=C(N(C1CCNCC1)c1cc2c([nH]cc2)cc1)c1cccc1            |
| 716 | 662642  | 7.40 | O1CCC(CC1)CN(C1CCNCC1)c1cc2c([nH]cc2)cc1             |
| 717 | 662643  | 8.40 | [nH]1c2c(cc(N(Cc3ccccc3C#N)C3CCNCC3)cc2)cc1          |
| 718 | 662650  | 8.80 | O(C)c1cc(ccc1)CN(C1CCNCC1)c1cc2c([nH]cc2)cc1         |
| 719 | 662670  | 8.68 | s1cccc1[C@]1(Oc2c3c(ccc2)cccc3)C[C@H]1CNC            |
| 720 | 664268  | 6.93 | O=C(N(CC1CCC1)[C@H]1CCNC1)c1cccc1C                   |
| 721 | 664274  | 6.65 | FC(F)(F)c1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1            |
| 722 | 664286  | 5.96 | S(C)c1cccc1C(=O)N(CCC)[C@H]1CCNC1                    |
| 723 | 664287  | 7.44 | S(C)c1cccc1C(=O)N([C@H]1CCNC1)C1CCCC1                |
| 724 | 664302  | 5.28 | O=C(N(CCC)[C@H]1CCNC1)c1cccc1C(C)C                   |
| 725 | 664321  | 6.40 | [nH]1c2c(cc(N(CCNC)c3ccccc3)cc2)cc1                  |
| 726 | 664335  | 7.70 | [nH]1c2c(cc(N(Cc3ccccc3)C3CCNCC3)cc2)cc1             |
| 727 | 1138354 | 8.20 | S(=O)(=O)(N)c1cc(ccc1)CN(C1CCNCC1)c1cc2c([nH]cc2)cc1 |
| 728 | 666074  | 6.58 | S(CC)c1cccc1C(=O)N(CC1CCC1)[C@H]1CCNC1               |
| 729 | 666090  | 7.19 | S(C)c1cccc1C(=O)N([C@H]1CCNC1)c1cccc1                |
| 730 | 666111  | 5.79 | O(c1cccc1C(=O)N(CC)[C@H]1CCNC1)c1cccc1               |
| 731 | 666112  | 6.80 | O(c1cccc1C(=O)N(CC(C)C)[C@H]1CCNC1)c1cccc1           |
| 732 | 666119  | 6.23 | O(c1cccc1C(=O)N(CC1CCC1)[C@@H]1CCNC1)c1cccc1         |
| 733 | 666134  | 8.40 | [NH2+]1CCC(N(Cc2ccccc2)c2cc3c(n(cc3)C)cc2)CC1        |
| 734 | 666174  | 7.48 | s1cccc1[C@@]1(Oc2c3c(ccc2)cccc3)C[C@H]1CNC           |
| 735 | 667945  | 6.96 | FC(F)(F)c1cccc1C(=O)N(CC1CCCC1)[C@H]1CCNC1           |
| 736 | 667949  | 7.17 | O(c1cccc1C(=O)N([C@H]1CCNC1)c1cccc1)c1cccc1          |
| 737 | 667950  | 5.97 | S(C)c1cccc1C(=O)N(CC1CCC1)[C@@H]1CCNC1               |
| 738 | 667961  | 5.70 | O=C(N(CC1CCC1)[C@@H]1CCNC1)c1cccc1C(C)C              |
| 739 | 667967  | 7.50 | [nH]1c2c(cc(cc2)C(CCNC)c2cccc2)cc1                   |
| 740 | 667979  | 7.90 | s1c2c(cc(N(Cc3ccccc3)C3CCNCC3)cc2)cc1                |
| 741 | 667993  | 8.40 | Fc1cccc1CN(C1CCNCC1)c1cc2c([nH]cc2)cc1               |
| 742 | 667994  | 8.90 | Fc1cc(ccc1)CN(C1CCNCC1)c1cc2c([nH]cc2)cc1            |
| 743 | 668021  | 7.75 | s1cccc1[C@]1(Oc2c3c(ccc2)cccc3)C[C@@H]1CNC           |
| 744 | 668022  | 7.52 | s1cccc1[C@@]1(Oc2c3c(ccc2)cccc3)C[C@@H]1CNC          |

## APPENDIX A3

### References

1. Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **2007**, 6, 881-90.
2. Kaiser, D.; Zdrazil, B.; Ecker, G. F. Similarity-based descriptors (SIBAR)--a tool for safe exchange of chemical information? *J Comput Aided Mol Des* **2005**, 19, 687-92.
3. Zdrazil, B.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity-Based Descriptors (SIBAR) as Tool for QSAR Studies on P-Glycoprotein Inhibitors: Influence of the Reference Set. *QSAR & Combinatorial Science* **2007**, 26, 669-678.
4. Klein, C.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity based SAR (SIBAR) as tool for early ADME profiling. *J Comput Aided Mol Des* **2002**, 16, 785-93.
5. Langer, T.; Eder, M.; Hoffmann, R. D.; Chiba, P.; Ecker, G. F. Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model. *Arch Pharm (Weinheim)* **2004**, 337, 317-27.
6. Chiba, P.; Tell, B.; Jager, W.; Richter, E.; Hitzler, M.; Ecker, G. Studies on propafenone-type modulators of multidrug-resistance (IV1): synthesis and pharmacological activity of 5-hydroxy and 5-benzyloxy derivatives. *Arch Pharm (Weinheim)* **1997**, 330, 343-7.

## CURRICULUM VITAE

|                |                         |
|----------------|-------------------------|
| Full Name      | Ishrat Jabeen           |
| Sex            | Female                  |
| Date of Birth  | 17-04-1982              |
| Place of Birth | Dist. Chakwal, Pakistan |
| Marital Status | Single                  |
| Nationality    | Pakistani               |
| E-mail Address | ishrat_chem@yahoo.com   |

### **Educational Background**

#### **2008-2011**

University of Vienna, Faculty of Life-Sciences, Department of Medicinal Chemistry, Pharmacoinformatics Research Group, Vienna, Austria

#### **Ph.D in Natural Sciences in Chemistry**

Advisor: Univ.-Prof. Dr. Gerhard F. Ecker                      E-mail: gerhard.f.ecker@univie.ac.at

**Thesis Title:** Combined Ligand- and Structure-Based Studies on inhibitors of P-glycoprotein

#### **2005-2007**

University of Agriculture Faisalabad, Department of Chemistry, Faisalabad, Pakistan

#### **M.Phil in Organic Chemistry**

Advisor: Prof. Dr. Rakshanda Nawaz                      E-mail: nawazn@fsd.paknet.com.pk

**Thesis Title:** Vivo Chelation of Iron and Zinc after Defereprone administration to male human beings

#### **2003-2005**

University of Agriculture Faisalabad, Department of Chemistry, Faisalabad, Pakistan

#### **M.Sc in Organic Chemistry**

Advisor: Prof. Dr. Rakshanda Nawaz                      E-mail: nawazn@fsd.paknet.com.pk

**Thesis Title:** In vivo distribution of sulfadiazine in whole and skimmed milk of buffalos

## **CURRICULUM VITAE**

### **2001-2003**

Punjab University Lahore, Pakistan

**B.Sc in Chemistry**

### **1999-2001**

Board of Intermediate and Secondary Education Rawalpindi, Pakistan

**F.Sc in Chemistry, Physics, Biology**

### **1999**

Board of Intermediate and Secondary Education Rawalpindi, Pakistan

**Matriculation** in Chemistry, Physics, Biology, Mathematics

## **Internship**

### **April -July, 2011**

Universitat Pompeu Fabra, Barcelona, Computer-Assisted Drug Design Laboratory (CADD), Research Unit on Biomedical Informatics (GRIB), Barcelona, Spain.

Advisor: Prof. Manuel Pastor

E-mail: manuel.pastor@upf.edu

Project: Study of hERG drug blockade mechanisms using structure-based and ligand-based computational methods.

## **Scholarships**

Merit Scholarship of Board of Intermediate and Secondary Education Rawalpindi, Pakistan 1999-2001

Merit Scholarship of University of Agriculture Faisalabad, Pakistan, 2005-2007.

Scholarship award from HEC (Higher Education Commission Pakistan, "Overseas scholarship for Ms/M.Phil Leading to PhD in selected fields' PHASE-II Batch-I, Austria" 2008-2011

## Publications

**Jabeen Ishrat**, Wetwitayaklung Penpun, Chiba Peter, Pastor Manuel and Ecker Gerhard F, Synthesis, Biological Activity and Quantitative Structure-Activity Relationship Studies of a Series of Benzopyranes and Benzopyrano[3,4-b][1,4]oxazines as Inhibitors of the Multidrug Transporter P-glycoprotein, *European Journal of Medicinal Chemistry*, 2011, Submitted article.

**Jabeen Ishrat**, Pleban Karin, Chiba Peter and Ecker Gerhard F. Structure-Activity Relationships, Ligand Efficiency and Lipophilic Efficiency Profiles of Benzophenone-Type Inhibitors of the Multidrug Transporter P-glycoprotein, *Journal of Medicinal Chemistry*, 2011, Submitted article.

**Jabeen I**, Wetwitayaklung P, Klepsch F, Parveen Z, Chiba P, Ecker GF. Probing the stereoselectivity of P-glycoprotein – synthesis, biological activity and preliminary docking studies of a set of enantiopure benzopyrano[3,4b][1,4]oxazines. *Chem Comm* **47**, 2586-2588 (2011)

Klepsch F, **Jabeen I**, Chiba P, Ecker GF. Pharmacoinformatic approaches to design natural product type ligands of ABC-transporters. *Curr Pharm Design*, **16**, 1742-1752 (2010)

## Contributions to Scientific Conferences

**I. Jabeen**, B. Plagens, W. Holzer, G. F. Ecker, QSAR, HQSAR and GRIND Studies on a Set of Heterocyclic Propafenone-Type Inhibitors of P-Glycoprotein, 21<sup>st</sup> Scientific Congress of the Austrian Pharmaceutical Society-Vienna, April 16-18, 2009. (Poster Presentation)

**Ishrat Jabeen**, Penpun Wetwitayaklung, Peter Chiba, Gerhard Ecker, Combine 2D and 3D- QSAR analysis of benzophenone type inhibitors of P-glycoprotein, Joint meeting of Medicinal Chemistry Budapest, Jun 24-27, 2009. (Poster Presentation)

**Ishrat Jabeen**, Penpun Wetwitayaklung, Peter Chiba, Gerhard Ecker, GRIND, CoMFA and CoMSIA studies of benzopyran type inhibitors of P-glycoprotein” in, Frontiers in Medicinal Chemistry (FMC) Barcelona, October 4-6, 2009. (Poster Presentation), *Drugs of the future*, Volume: 34, Pages: 168, Published: OCT 2009.

Gerda Brunhofer, **Ishrat Jabeen**, Zahida Parveen, Peter Chiba; Gerhard Ecker; Thomas Erker, QSAR-guided synthesis of chalcone-like P-glycoprotein inhibitors, 238<sup>th</sup> ACS national meeting, Washington, DC, United States, August 16-09, 2009. MEDI, 378. Publisher: American Chemical society, Washington, DC. CODEN 69: LVCL. AN: 2009: 984572.

**Ishrat Jabeen**, Zahida Parveen, Uwe Rinner, Peter Chiba, Gerhard F. Ecker, SAR, ligand efficiency (LE) and lipophilic efficiency (LipE) studies of a series of benzophenone-type inhibitors of the multidrug transporter P-glycoprotein, 240th ACS National Meeting, Boston, MA, United States, August 22-26, 2010, COMP-244. Publisher: (American Chemical Society, Washington, D. C) CODEN: 69, NAQG AN: 2010, 1009871. (Poster Presentation)

## **CURRICULUM VITAE**

**I. Jabeen**, P. Wetwitayaklung, F. Klepsch, P. Chiba, G. F. Ecker, Stereoselective interactions of Benzopyrano [3,4-b][1,4]Oxazine with P-glycoprotein, talk in 18th European Symposium on Quantitative Structure- Activity Relationships in Rhodes Greece September 19-24, 2010.

**Ishrat Jabeen**, Karin Pleban, Peter Chiba, Gerhard F. Ecker, SAR, ligand efficiency (LE) and lipophilic efficiency (LipE) studies of a series of benzophenone-type inhibitors of the multidrug transporter P-glycoprotein, Joint Meeting of the Austrian and German Pharmaceutical Societies, Innsbruck, Austria, September 20 - 23, 2011. (Poster Presentation)



Abstract der Dissertation

angestrebter akademischer Grad: Doktor der Naturwissenschaften Chemie (Dr. rer.nat)

A 091419 Chemie

Ishrat Jabeen, M.Phil Chemistry  
Pharmacoinformatics Research Group  
Department für Medizinische Chemie  
Althanstraße 14, 1090, Wien  
Universität Wien, Österreich

Der aktive Effluxtransporter P-Glykoprotein (P-gp) ist verantwortlich für Multidrug Resistenz (MDR) in Tumoren und beeinflusst außerdem die ADME Eigenschaften von Arzneistoffkandidaten. P-gp zeigt eine sehr breite Substratspezifität und transportiert daher eine hohe Anzahl von strukturell und funktionell diversen Substanzen aus Tumorzellen hinaus und über physiologische Barrieren hinweg. Obwohl in den letzten zwei Jahrzehnten einige Inhibitoren von P-gp identifiziert wurden, scheiterten alle von ihnen in klinischen Studien, entweder wegen schwerwiegenden Nebenwirkungen, oder wegen fehlender Wirksamkeit. Dies betont die Notwendigkeit von verlässlichen *in silico* Modellen für die Vorhersage von P-gp Substraten und Inhibitoren bereits in frühen Phasen der Wirkstoffentwicklung. In dieser Arbeit wurden daher unterschiedliche *in silico* Methoden verwendet um Einblicke in die dreidimensionalen strukturellen Voraussetzungen der Liganden, ihren Bindungsmodus und ihre Stereoselektivität gegenüber P-gp zu erhalten.

Verschiedene 2D- und 3D-QSAR Modelle wurden mit einfachen physicochemischen sowie komplexen 3D-Deskriptoren (GRIND) für unterschiedliche chemische Grundkörper erstellt, um globale strukturelle Merkmale von P-gp Inhibitoren zu untersuchen. Um die vielversprechendsten P-gp Liganden mit dem besten Wirksamkeits/Lipophilie- oder Größenverhältnis zu identifizieren, verwendeten wir zum bisher ersten Mal ligandeneffizienz- und lipophilieeffizienz-basierte Ansätze. Interessanterweise überschritt keine der vielversprechendsten Substanzen den LipE Grenzwert von 5. Dies könnte mit dem einzigartigen Zugangsweg der Substanzen

## **ABSTRACT (German)**

zusammenhängen, der anders als bei anderen Transportern oder Ionenkanälen direkt aus der Zellmembran erfolgt. Unsere Dockingstudien bieten einen ersten Nachweis über unterschiedliche Bindungsareale für zwei diastereomere Substanzserien und zeigen eine stereoselektive Ligandenerkennung von P-gp. Zusätzlich war es uns möglich zu zeigen, dass sich ein Benzophenon-Dimer so platzieren lässt, dass diese beiden Areale verbunden werden, was die Hypothese von mehreren, teilweise überlappenden, Bindungsarealen von P-gp verstärkt. Die in dieser Dissertation beschriebene Arbeit wird den Weg für die Entwicklung von zukünftlichen neuen und vielversprechenderen Inhibitoren von P-gp bereiten, die bessere ADME Eigenschaften und verringerte Toxizität besitzen.

**Keywords:** P-glycoprotein, ABCB1, Multidrug Resistenz (MDR), Lipophilieeffizienz-basierte, Ligandeneffizienz, QSAR, GRIND, Docking.





Abstract of the Dissertation Submitted to the Examination Department (Prüfungsreferat Naturwissenschaften Chemie) of University of Vienna, Austria, in Partial Fulfilment of the Requirements for the Degree of Doctor of Natural Sciences (Dr. rer.nat)  
A 091419 Chemistry

Ishrat Jabeen, M.Phil Chemistry  
Pharmacoinformatics Research Group  
Department of Medicinal Chemistry  
Althanstraße 14, 1090, Vienna  
University of Vienna, Austria

The drug efflux pump P-glycoprotein (P-gp) has been shown to cause multidrug resistance (MDR) in tumors as well as to influence ADME properties of drug candidates. P-gp is highly promiscuous in its ligand recognition profiles and thus transports numerous structurally and functionally diverse compounds out of tumor cells and across physiological barriers. Several inhibitors of P-gp mediated drug efflux have been identified in the past two decades, but all of them failed in clinical trials due to severe side effects and lack of efficacy. This further emphasizes the necessity of reliable *in-silico* tools for prediction of P-gp substrates and inhibitors during the early phases of drug discovery. Therefore, in this thesis, various *in silico* tools have been utilized to get insights into 3D structural requirements of ligands, their binding modes, as well as their stereoselectivity towards P-gp.

Different 2D- and 3D-QSAR models using simple physicochemical and GRID independent molecular descriptors have been constructed across different chemical scaffolds to investigate global structural attributes of P-gp inhibitors. In order to identify most promising P-gp ligands with best potency/lipophilicity or size ratio, we, for the first time, also used ligand efficiency and lipophilic efficiency based approaches. Interestingly, none of the P-gp inhibitors/substrates cross the LipE threshold of 5 for highly promising compounds. This might be linked to the unique entry pathway directly from the membrane bilayer, which is rather unique for transporters and ion

## **ABSTRACT (English)**

channels. Our docking studies provide the first evidence for different binding areas of two diastereomeric compound series and provides evidence for stereoselective ligand recognition of by P-gp. In addition we could show that a benzophenone dimer is well docked in a pose bridging these two distinct binding sites, which further strengthens the hypothesis of multiple, partly overlapping binding sites at P-gp. The work described in this thesis will pave the way for the design of new and more promising inhibitors of P-gp in the future with better ADME properties and reduced toxicity.

**Keywords:** P-glycoprotein, ABCB1, Multidrug Resistance (MDR), Lipophilic Efficiency, Ligand Efficiency, Stereoselectivity, QSAR, GRIND, Docking.

**List of Abbreviations**

|          |  |
|----------|--|
| P-gp     | P-glycoprotein   |
| ABC      | ATP Binding Cassette   |
| MDR      | Multidrug Resistance   |
| TMD      | Transmembrane Domain   |
| NBD      | Nucleotide Binding Domain                                    |
| ADME-Tox | Absorption, Distribution, Metabolism, Excretion and Toxicity |
| BBB      | Blood–Brain Barrier  |
| B-CSF    | Blood Cerebrospinal Fluid                                    |
| SAR      | Structure Activity Relationship                              |
| QSAR     | Quantitative Structure Activity Relationship                 |
| HQSAR    | Hologram Quantitative Structure Activity Relationship        |
| CoMFA    | Comparative Molecular Field Analysis                         |
| CoMSIA   | Comparative Molecular Similarity Index Analysis              |
| MIF      | Molecular Interaction Field                                  |
| GRIND    | GRID Independent Descriptors                                 |
| TPSA     | Topological Polar Surface Area                               |
| LE       | Ligand Efficiency  |
| LipE     | Lipophilic Efficiency  |
| MOE      | Molecular Operating Environment                              |
| GOLD     | Genetic Optimization for Ligand Docking                      |
| RMSD     | Root Mean Square Deviation                                   |
| PLS      | Partial Least Square   |
| PC       | Principal Component  |
| PCA      | Principal Component Analysis                                 |
| PDB      | Protein Data Bank  |
| hERG     | human <i>Ether-à-go-go</i> Related Gene                      |
| SERT     | Serotonin transporter  |
| CYP3A4   | Cytochrome P <sub>450</sub> 3A4                              |