



## Discriminant and quantitative PLS analysis of competitive CYP2C9 inhibitors versus non-inhibitors using alignment independent GRIND descriptors

Lovisa Afzelius<sup>1,2,\*</sup>, Collen M. Masimirembwa<sup>2</sup>, Anders Karlén<sup>1</sup>, Tommy B. Andersson<sup>2</sup> & Ismael Zamora<sup>3</sup>

<sup>1</sup>Department of Organic Pharmaceutical Chemistry, Biomedical Center, Uppsala University, Uppsala, Sweden;

<sup>2</sup>Department of Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry, AstraZeneca R&D Mölndal, Sweden; <sup>3</sup>Lead Molecular Design, s.l. Francesc Cabanes i Alibau, 1-3, 2<sup>o</sup> I<sup>a</sup>, 08190 Sant Cugat del Valles, Barcelona, Spain

Received 16 January 2002; accepted in revised form 3 July 2002

**Key words:** cytochrome P450, computational modeling, CYP2C9 inhibitors, enzyme inhibition, 3D-QSAR, discriminant PLS, ALMOND, GRIND descriptors

### Summary

This study describes the use of alignment-independent descriptors for obtaining *qualitative* and *quantitative* predictions of the competitive inhibition of CYP2C9 on a serie of highly structurally diverse compounds. This was accomplished by calculating alignment independent descriptors in ALMOND. These GRid INdependent Descriptors (GRIND) represent the most important GRID-interactions as a function of the distance instead of the actual position of each grid-point. The experimental data was determined under uniform conditions. The inhibitor data set consists of 35 structurally diverse competitive stereospecific inhibitors of the cytochrome P450 2C9 and the non -inhibitor data set of 46 compounds. In a PLS *discriminant* analysis 21 inhibitors and 21 non-inhibitors (1 and 0 as activities) were analyzed using the ALMOND program obtaining a model with an  $r^2$  of 0.74 and a cross-validation value ( $q^2$ ) of 0.64. The model was externally validated with 39 compounds (14 inhibitors/25 non-inhibitors). 74% of the compounds were correctly predicted and an additional 13% was assigned to a borderline cluster. Thereafter, a model for *quantitative predictions* was generated by a PLS analysis of the GRIND descriptors using the experimental  $K_i$ -value for 21 of the competitive inhibitors ( $r^2 = 0.77$ ,  $q^2 = 0.60$ ). The model was externally validated using 12 compounds and predicted 11 out of 12 of the  $K_i$ -values within 0.5 log units. The discriminant model will be useful in screening for CYP2C9 inhibitors from large compound collections. The 3D-QSAR model will be used during lead optimization to avoid chemistry that result in inhibition of CYP2C9.

**Abbreviations:** P450 – cytochrome P450; 3D-QSAR – three-dimensional quantitative structure-activity relationship; CoMFA – comparative molecular field analysis; CoMSIA – comparative molecular similarity indices analysis; PCA – principal component analysis; PLS – partial least square analysis; GOLPE – generating optimal linear partial least-squares analysis estimations; SaSA – Synthesis and Structure Administration.

### Introduction

Three-dimensional quantitative structure activity relationship (3D-QSAR) is an established method to

predict the effect of compounds on biological systems. One of the major problems when performing 3D QSAR analysis is how to align the effector molecules. A proper alignment could reflect steric and electrostatic interactions with specific amino acids in the active site. In homologous series one can expect compounds

\*To whom correspondence should be sent: email: Lovisa.Afzelius@astrazeneca.com

to share a similar binding mode in the receptor, and alignment could be obtained by superimposing common molecular features and/or core-structures consistent throughout the series. However, even in homologous series, molecules can adopt different binding modes in the active site [1, 2]. In a series of diverse compounds where no common molecular features can be determined the alignment is more difficult. Experimentally, this problem can be solved if compounds are co-crystallised with their biological effector protein. This information is, however, not available for many ligand–receptor complexes because of technical difficulties, especially when dealing with membrane bound proteins.

There are currently a number of approaches to address the issue of superimposition which could be considered at two levels; (1) selection of the bioactive conformer and (2) alignment of the selected conformers. Most approaches available today try to solve these problems simultaneously by aligning hypothetical bioactive conformers, based on a pharmacophore hypothesis. In DISCO [3], a clique-detection method is used to find superpositions that contain at least one conformation of each molecule and user-specified number of point types and chirality. CATALYST [4] works by generating multiple conformers through a systematic search and defines pharmacophore-based alignment hypotheses comprising hydrogen bond acceptor/donor sites and/or hydrophobic regions. This program has been used to generate models for cytochrome P450 mediated metabolism [5, 6] and enzyme inhibition [6]. Another approach is manual alignment of conformationally constrained core structures based on a pharmacophore hypothesis generated from homology models and site directed mutagenesis [7]. Furthermore, as in the case presented here, docking into the active site can be used to constrain the conformational space. Similarity to template compounds with experimentally determined orientation in the active site can be used to select conformers of compounds where the orientation in the active site is unknown [8].

Once the molecules are aligned the molecular interaction fields for the selected conformers can be calculated using GRID [9] or CoMFA [10]. In these methods probes are used to determine the energies of interaction for each point in a grid lattice surrounding the aligned molecules. The molecular fields are then related to the biological variable to derive models using multivariate data analysis methods such as PLS.

In this study alignment independent GRIND descriptors (GRid-INdependent Descriptors) [11] calculated in ALMOND have been used. The concept is based on autocorrelation functions [12]. In the classical autocorrelation all correlations are summed, which makes it impossible to trace them back to the descriptor space. In the case of the GRIND autocorrelograms, the mathematical transformations have been simplified to include only maximum correlations. Therefore, the GRIND descriptors allow immediate backtracking of changes in activity to specific changes in structure.

The data sets analysed in this study are a set of competitive CYP2C9 inhibitors [8] and a set of compounds with no affinity to the enzyme (non-inhibitors). CYP2C is a subfamily of the cytochrome P450 (CYP450's) superfamily [13] which is a major enzyme system in the oxidative metabolism of xenobiotics. The subfamily contains at least four different isoforms (2C8, 2C9, 2C18 and 2C19) in the human liver and it is responsible for the metabolism of many drugs such as taxol, warfarin and omeprazole [14]. The isoform studied here, CYP2C9, is responsible for ~18% of all reactions catalysed by the superfamily [15]. There is a high structural diversity among its substrates and inhibitors [16], [6]. The structure-activity-relationship of CYP2C9 substrates have been described as amphipathic with a region of lipophilicity at the site of hydroxylation and an area of hydrophobicity around the hydrogen bond forming region [17]. The data set ranges from MW100 to 1000 and no obvious alignment rules can be found for these diverse compounds that are CYP2C9 inhibitors. Therefore, alignment independent GRIND descriptors [11] were used to address the problem of alignment and PLS was used to derive discriminative [18] and quantitative models [19] for CYP2C9 inhibitors.

Models derived in this study could successfully distinguish competitive CYP2C9 inhibitors and noninhibitors. The quantitative model could predict inhibition constants ( $K_i$ ) within 0.5 log units of the measured values.

## Methods and material

### *Software and molecular modeling*

Enzyme kinetic analysis was done using GraFit 4.0.12 (Erithacus Software Limited, Middlesex, U.K.) and SIMFIT 5.3 [20]. Protein homology and pharmacophore modeling were done on Silicon Graphics Oc-

tane and O<sub>2</sub> workstations respectively (Silicon Graphics Inc., Mountain View, CA, USA). Insight II 98.0 (Molecular Simulations Inc., San Diego, CA, USA) was used in the protein homology modeling. The software utilised in the molecular modelling were: GOLD 1.1 (Dr Gareth Jones, University of Sheffield, U.K.), GRID and VOLSURF (Molecular Discovery Ltd., University of Oxford, U.K.), GOLPE and ALMOND 2.0 (MIA, Perugia, Italy), SYBYL 6.5.3 and CONCORD (Tripos Associates Inc., St Louis, MO, USA) SaSA (AstraZeneca R&D Mölndal). Chemical structures were imported from ISIS-BASE database or drawn in ISIS-Draw (MDL information Systems Inc., San Leandro, CA, USA) and ChemDraw (Cambridge-Soft Corporation, MA, USA).

### Overview

The affinity for CYP2C9 was measured for a number of compounds to extract those with non-measurable affinity, the non-inhibitor data set. The inhibition constants,  $K_i$ , were determined for a set of competitive CYP2C9 inhibitors. The differences between the two data sets were explored in a discriminant model based on GRIND descriptors in ALMOND. In addition, the  $K_i$ -values of the competitive CYP2C9 inhibitors were used to generate a quantitatively predictive model also based on GRIND descriptors.

### Biological conditions

#### Discriminant / quantitative model

An assay for enzyme affinity (percentage inhibition) determination was set up using recombinant CYP2C9 enzyme. All incubations were done in a total volume of 200  $\mu$ M and with a final concentration of 0.1 M potassium phosphate buffer (pH 7.4), 1 mM NADPH and 8  $\mu$ M diclofenac (substrate of CYP2C9) dissolved in water. The substrate concentration of 8  $\mu$ M equals the  $K_m$  of this compound for CYP2C9. This allows for estimating  $K_i$  from  $IC_{50}$ ,  $K_i = IC_{50}/2$  for competitive inhibition [21]. The test compounds were dissolved in H<sub>2</sub>O, ethanol, methanol or acetonitrile at a final concentration of 100  $\mu$ M. The samples were analysed using an HPLC-UV method [21] and the percentage of metabolite (4-OH diclofenac) formed in comparison to a blank containing equal amount of solvent was measured. Inhibition assays were done in triplicate. All compounds inhibiting <10% of the CYP2C9 catalysed diclofenac 4-hydroxylation, were considered non-inhibitors.

Forty-six compounds were classified as non-inhibitors. In cases where racemates were tested (all chiral non-inhibitors except S-mephenytoin and S-propanolol) we assumed a negligible stereoselectivity for the on/off response and included all stereoisomers in the calculations. The competitive CYP2C9 inhibitor data set was generated by determining the full inhibition profile at four different inhibitor concentrations and twelve different substrate concentrations for each tested compound as described in ref [8]. This data set has now been expanded to include (R-) H 259/31, (S-) H 259/31, (R-) H 287/23 and (S-) H 287/23. Competitive inhibition was assigned from Michaelis-Menten saturation curves, Lineweaver-Burk and Eadie-Hoffstee plots calculated by regression methods using GraFit 4.0.12 and SIMFIT 5.3. Table 1 and 2 show the structures and activities of the compounds used to derive the discriminant and quantitative model.

### Selection and characterization of the data set

The position of the compounds in a 2D chemical space has been evaluated using ChemGPS (Chemical Global Positioning System). ChemGPS is a principal component analysis (PCA) [22] model built from a data set that aims to define the borders of a drug-like space [23]. The ChemGPS predictions were based on 72 descriptors calculated in either SaSA [24] or VolSurf [25]. The influence of each individual descriptor from SaSA was also examined for the classification of inhibitors versus non-inhibitors and for the quantitative predictions of the inhibitor data set.

### Conformer selection

Conformer selection was made based on docking template molecules using GOLD1.1 in a CYP2C9 homology model derived from the rabbit CYP2C5 crystal structure [8]. Three out of four template inhibitors were also substrates with an experimentally determined site of oxidation (S-warfarin (11), phenytoin (6) and progesterone (5)) [26]. The fourth template molecule, sulphaphenazole (1), hexa-coordinates to the heme with a specific nitrogen atom [27]. One solution was selected for each template molecule based on orientation and distance to the heme from the site of interaction. All other compounds, inhibitors and non-inhibitors, were docked under the same conditions. The GRID[9] interaction fields were calculated for all solutions using the DRY and OH probe and analysed

Table 1. Training set –  $K_i$  exp. ( $\mu$ M).

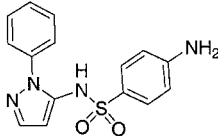
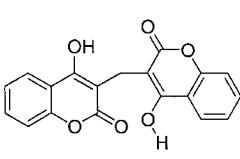
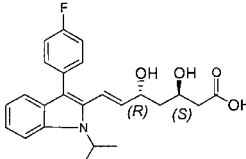
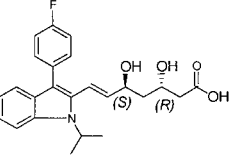
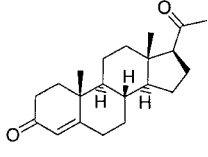
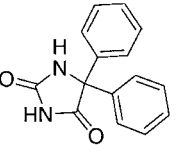
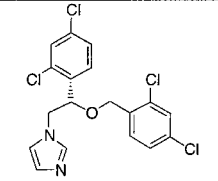
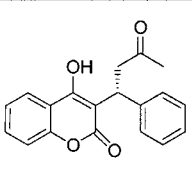
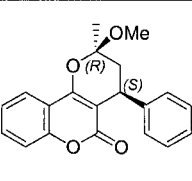
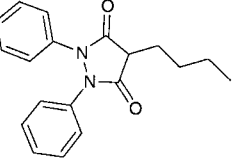
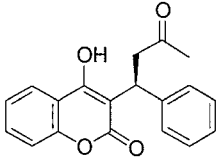
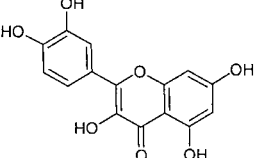
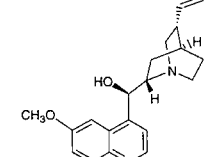
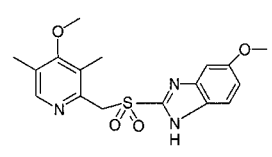
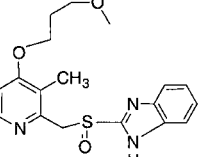
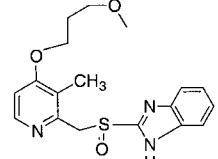
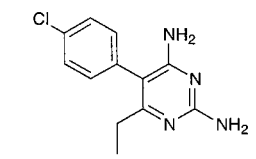
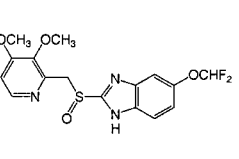
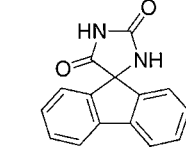
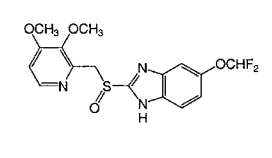
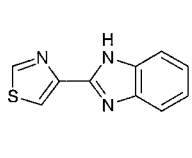
Table 1 Training set -- $K_i$ exp. ( $\mu$ M)		
		
sulphaphenazole (1) $K_i = 0.5$	dicoumarol (2) $K_i = 1.9$	( <i>R,S</i> )-fluvastatin (3) $K_i = 2.2$
		
( <i>S,R</i> )-fluvastatin (4) $K_i = 3.3$	progesterone (5) $K_i = 5.5$	phenytoin (6) $K_i = 6$
		
( <i>S</i> )-miconazole (7) $K_i = 6$	( <i>R</i> )-warfarin (8) $K_i = 13.6$	( <i>R,S</i> )-pyranocoumarin (9) $K_i = 16$
		
phenylbutazone (10) $K_i = 19.3$	( <i>S</i> )-warfarin (11) $K_i = 20$	quercetin (12) $K_i = 27$
		
quinine (13) $K_i = 32$	omeprazole sulfone (14) $K_i = 35$	( <i>R</i> )-rabeprazole (15) $K_i = 36$
		
( <i>S</i> )-rabeprazole (16) $K_i = 37$	pyrimethamine (17) $K_i = 51.5$	( <i>S</i> )-pantoprazole (18) $K_i = 64$
		
SFID (19) $K_i = 125$	( <i>R</i> )-pantoprazole (20) $K_i = 145$	thiabendazole (21) $K_i = 245$

Table 1. Continued.

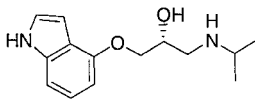
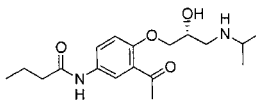
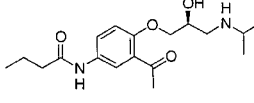
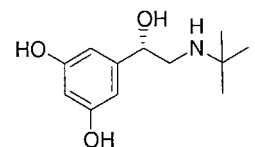
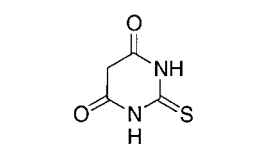
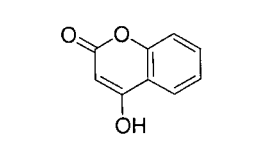
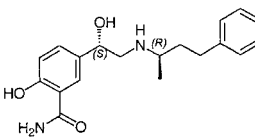
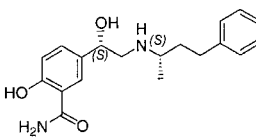
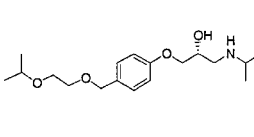
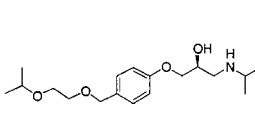
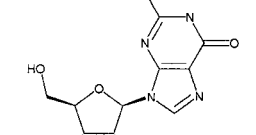
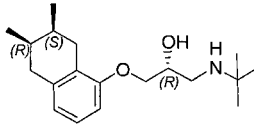
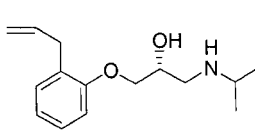
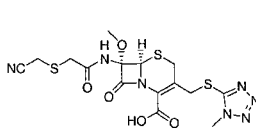
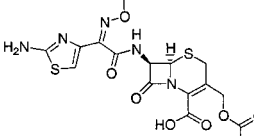
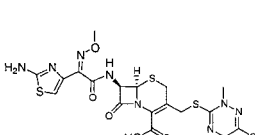
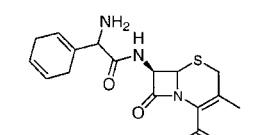
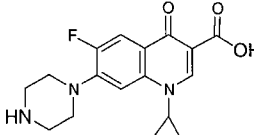
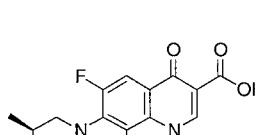
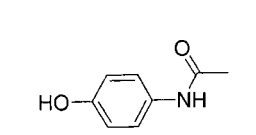
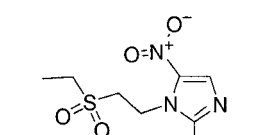
		
<i>(R)</i> -pindolol ( <b>22</b> ) <i>non-inhib</i>	<i>(R)</i> -acebutolol ( <b>23</b> ) <i>non-inhib</i>	<i>(S)</i> -acebutolol ( <b>24</b> ) <i>non-inhib</i>
		
<i>(S)</i> -terbutaline ( <b>25</b> ) <i>non-inhib</i>	2-thiobarbituric acid ( <b>26</b> ) <i>non-inhib</i>	4-OH-coumarin ( <b>27</b> ) <i>non-inhib</i>
		
<i>(4S,15R)</i> -labetalol ( <b>28</b> ) <i>non-inhib</i>	<i>(4S,15S)</i> -labetalol ( <b>29</b> ) <i>non-inhib</i>	<i>(R)</i> -bisoprolol ( <b>30</b> ) <i>non-inhib</i>
		
<i>(S)</i> -bisoprolol ( <b>31</b> ) <i>non-inhib</i>	guanosine ( <b>32</b> ) <i>non-inhib</i>	<i>(R)</i> -nadolol ( <b>33</b> ) <i>non-inhib</i>
		
<i>(R)</i> -alprenolol ( <b>34</b> ) <i>non-inhib</i>	cefmetazole ( <b>35</b> ) <i>non-inhib</i>	cefotaxime ( <b>36</b> ) <i>non-inhib</i>
		
ceftriaxone ( <b>37</b> ) <i>non-inhib</i>	cephradine ( <b>38</b> ) <i>non-inhib</i>	ciprofloxacin ( <b>39</b> ) <i>non-inhib</i>
		
<i>(S)</i> -lomefloxacin ( <b>40</b> ) <i>non-inhib</i>	paracetamol ( <b>41</b> ) <i>non-inhib</i>	tinidazole ( <b>42</b> ) <i>non-inhib</i>

Table 2. Test set –  $K_i$  exp. /  $K_i$  pred ( $\mu$ M).

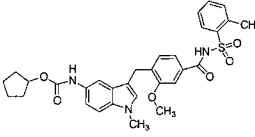
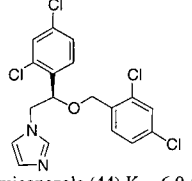
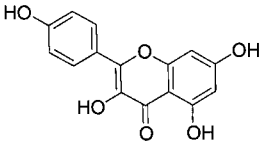
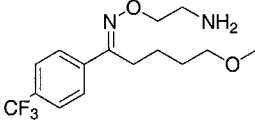
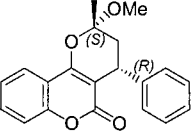
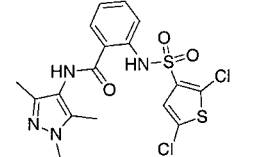
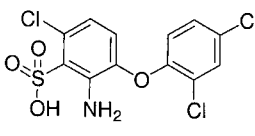
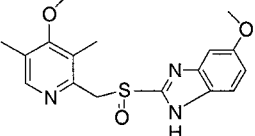
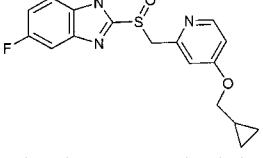
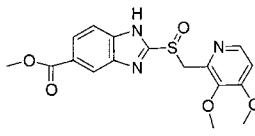
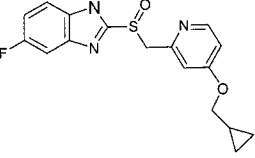
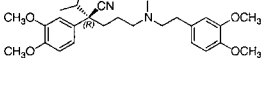
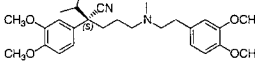
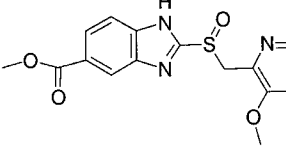
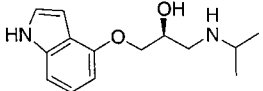
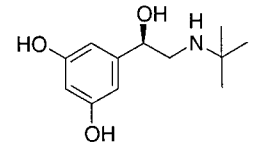
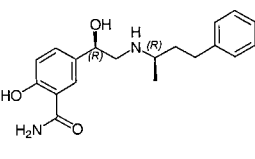
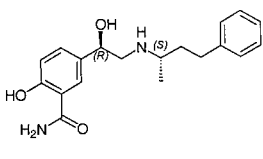
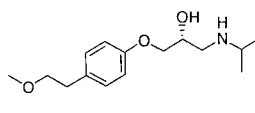
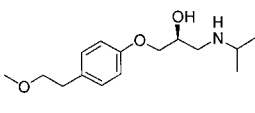
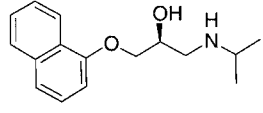
Table 2. Test set -- $K_i$ exp. / $K_i$ pred ( $\mu$ M)		
		
zafirlukast (43) $K_i$ = 2.5 / 2.1	( <i>R</i> )-miconazole (44) $K_i$ = 6.0 / 4.7	kaempferol (45) $K_i$ = 6.0 / 7.1
		
fluvoxamine (46) $K_i$ = 8.5 / 18.8	( <i>S,R</i> )-pyranocoumarin (47) $K_i$ = 16 / 14.8	D-62126 (48) $K_i$ = 17.0 / 5.3
		
A-4438 (49) $K_i$ = 20 / 11.2	( <i>R</i> )-omeprazole (50) $K_i$ = 45.0 / 44.2	<i>S</i> -H259/31 (51) $K_i$ = 60 / 57.4
		
( <i>S</i> )-H287/23 (52) $K_i$ = 64 / 24.8	( <i>R</i> )-H259/31 (53) $K_i$ = 93 / 30.6	( <i>R</i> )-verapamil (54) $K_i$ = 90 – 127*
		
( <i>S</i> )-verapamil (55) $K_i$ = 90 – 127*	( <i>R</i> )-H287/23 (56) $K_i$ = 140 / 17.6	( <i>S</i> )-pindolol (57) noninh/noninh
		
( <i>R</i> )-terbutaline (58) noninh/ noninh	( <i>R,R</i> )-labetalol (59) noninh/ noninh	( <i>R,S</i> )-labetalol (60) noninh/ noninh
		
( <i>R</i> )-metoprolol (61) noninh/ noninh	( <i>S</i> )-metoprolol (62) noninh/ noninh	( <i>S</i> )-propanolol (63) noninh/ border

Table 2. Continued.

(S)-nadolol (64) <i>noninh/ noninh</i>	caffeine (65) <i>noninh/ border</i>	cefaclor (66) <i>noninh/ noninh</i>
cefoperazone (67) <i>noninh/ inhib</i>	cefuroxime (68) <i>noninh/ noninh</i>	cephalothin (69) <i>noninh/ noninh</i>
cephapirin (70) <i>noninh/ border</i>	(R)-lomefloxacin (71) <i>noninh/ noninh</i>	metronidazole (72) <i>noninh/ noninh</i>
primidone (73) <i>noninh/ noninh</i>	theophylline (74) <i>noninh/ border</i>	timolol (75) <i>noninh/ noninh</i>
violuric acid (76) <i>noninh/ noninh</i>	yohimbine (77) <i>noninh/ border</i>	(S)-alprenolol (78) <i>noninh/ inhib</i>
(S)-mephénytoin (79) <i>noninh/ nhib</i>	(R)-levobunolol (80) <i>noninh/ noninh</i>	(S)-levobunolol (81) <i>noninh/ noninh</i>

\* absolute configuration not determined

in GOLPE[28]. The euclidian distances were computed in a 4D PCA score space and one solution for each compound closest in distance to the four template molecules were selected as described previously [8].

#### Alignment independent descriptors

ALMOND uses GRIND-descriptors; (GRId INdependent Descriptors) [11] that increase the signal to noise ratio and compresses the information compared to GRID interaction fields. These descriptors were computed in a three-step process : (1) the interac-

tion fields of the compounds were calculated using chemical probes, DRY (steric and hydrophobic), N1 (amide nitrogen-hydrogen bond donor) and O (carbonyl oxygen-hydrogen bond acceptor) probes, in GRID. (2) The second step involved a filtering process on the grid interaction fields, selecting the points with the lowest energy (most favourable interactions) and the greatest distance between them, using a Federov-like algorithm [29, 30] procedure in the grid interaction field variables. Since only negative energy values are considered in the generation of these new descriptors a compression of the information is obtained and consequently some of the information is lost. The energy values for each block of variables were scaled by a pre-computed maximum value of interaction defined for each probe. For the full set of variables the default raw scaling was used. (3) In the final step the maximum cross and auto correlograms were calculated. In this process the distances obtained from the interaction fields are binned into different distance categories. The bin sizes are defined as smoothing windows. If the smoothing window is assigned a value slightly lower than the grid box size, the desired 'smoothing effect' can be achieved. The maximum interaction for each smoothing window is collected. These interaction energies at specified distances correspond to the GRIND descriptors. The advantage of the filtering process and the subsequent distance grouping and translation to correlograms is that the variables after this treatment are not dependent on the alignment of the molecules. They do not relate back to grid points with a defined spatial position but instead they correlate to distances between points. By converting the GRID points into ALMOND distances, 6 degrees of freedom, three rotational and three translational, are solved. These new grid independent variables can be back-tracked into the space of the original molecules.

#### *Discriminant model*

The discriminant model was built to be able to distinguish new compounds as being similar to competitive CYP2C9 inhibitors or not. According to the Pattern Recognition (PARC) rules [31] the data describes an asymmetric case, where one class (non-inhibitors) has no detectable structure and observations are spread randomly and where the other class is supposed to have a tight structure with distinct structural features in common. The discriminant PLS model rotates the projection of the principal components to latent variables that focus on class separation [18] by introducing

binary y-variables (0 and 1). The goal is to find a restricted volume in x-space, a discriminant plane that separates the compounds.

The DRY, N1 and O probes were used to calculate the GRIND descriptors for all molecules, (inhibitors = 35 and non-inhibitors = 46). The grid space was set to 0.5 Å, the number of filtered points to 150 and the smoothing window to 0.8 grid units. In the filtering process the weighting between distance and energy was set to 50% and raw scaling was used. The number of points was also increased to examine its influence on the model. Data sets containing large compounds sometimes require a larger number of points to be extracted in order to retain adequate information in the compression process [11]. The structural characteristics of the data set explained by the GRIND descriptors were evaluated using ALMOND. A PCA was done on the descriptor matrix, and the score plot was examined to look for structural characteristics of the data set independent of the y classification variable. The y-value, a binary-variable, was then set to 0 (non-inhibitor) or 1 (competitive inhibitor) and a PLS-DA (partial least squares discriminant analysis) model was built from 21 inhibitors and 21 non-inhibitors (Table 1). The model was cross-validated and thereafter externally validated using 14 competitive inhibitors and 25 non-inhibitors (Table 2).

To explore the effect of the different correlograms in the model a two-factor full factorial design was applied to the six different correlograms; the auto-correlograms (DRY-DRY, O-O and N1-N1) and the cross-correlograms (DRY-O, DRY-N1 and O-N1). This yielded  $2^6 = 64$  different combinations for which the predictive power, SDEP, (standard deviation of error of predictions) [32] was calculated using two components in the model. In order to calculate the two-factor interaction effects, the Yate's algorithm [33] was applied to the observations after they had been written in a standard order where the  $k$ th column consists of  $2^{k-1}$  minus signs followed by  $2^{k-1}$  plus signs. The predictive power (SDEP) and estimates of effects were assessed using normal probability plots. A normal distribution would plot as a straight line on normal probability scale and hence used to assess combinations of correlograms that yield the most predictive model and evaluate if the increase in predictivity is statistically significant.



### Quantitative model

The training set of 21 competitive inhibitors was used to obtain a quantitative model. The grid space was set to 0.5 Å, the number of points to 150 and the smoothing window to 0.8 grid units. The weighting in the filtering step between distance and energy was set to 50% and the DRY, O and N1 probes were used in the calculations. The number of points and the weighting were also varied to explore the effects of this parameter in the predictive power of this model. The PCA score plots were examined to distinguish object differences and how much of the variance that was explained. Thereafter, the inhibitory constant ( $K_i$ ) was introduced as the y variable and a PLS [19] model was built, cross validated and externally tested using 12 competitive inhibitors with defined stereochemistry. A full factorial design was applied to the data set in the same way as for the discriminant model to evaluate the effect of the different correlograms on the predictivity and the interactions of the correlograms. The results were evaluated using normal probability plots based on the SDEP values for a two-component model. Since no model was significantly better than the initial model generated from all correlograms, this model was further explored using GOLPE variable selection [34] and externally validated.

## Results

### Selection and characterization of the data set

The data set of non-inhibitors was generated from a set of 100 drug compounds distributed evenly in a PCA score plot based on VolSurf descriptors. The first two components explained mainly hydrophobicity and polarity in ChemGPS based on VolSurf descriptors [35]. Out of the 100 compounds 46 non-inhibitors were determined. The rest were discarded due to affinity to the enzyme or problems with solvation/compound availability. The data set of competitive CYP2C9 inhibitors was collected by experimentally testing compounds described in literature as inhibitors, substrates or mentioned in the context of CYP2C9 to generate such a diverse set of structures as possible (Tables 1 and 2).

Inspection of the PCA score plots showed no clustering of non-inhibitors vs inhibitors in any of the six components that ChemGPS calculated from the SaSA or the VolSurf descriptors. Examination of the individual SaSA descriptors showed a strong correlation between clogP, calculated using Daylight software,

and the discrimination between inhibitors and non-inhibitors. Compounds with a clogP below 1 were non-inhibitors, between 1 and 3 both inhibitors and non-inhibitors and above 3 inhibitors. Since there were no outliers, 65% of all compounds were correctly classified in the non-inhibitor or in the inhibitor class; the remaining 35% were classified as an intermediate class. Relationships between lipophilic character and metabolism of organic compounds by microsome P450's has been shown previously [36] and reflects probably only a compound's possibility to be presented to the active site. None of the other 71 descriptors showed such a discriminant behaviour. For the competitive inhibitor data set no correlation between the activity and clogP or any other descriptor was seen.

### Discriminant analysis

The PCA score plot generated from the GRIND descriptors of the training set of non-inhibitors and inhibitors were examined in ALMOND. The two first components explained 58% of the variance and there was a clear discrimination between inhibitors and non-inhibitors (Fig. 1). The existence of structure in the x-space validated the use of the discriminant PLS-methodology based on binary variables. The initial model was predictive in two components ( $r^2 = 0.65$ ,  $q^2_{(LOO)} = 0.46$   $q^2_{(5 \text{ random groups})} = 0.42$ ) without any variable selection. When this model was used to predict the test set of 39 compounds (14 competitive inhibitors and 25 non-inhibitors) 67% were correctly classified, 18% were in a border region and 15% were misclassified. Compounds predicted between 0.4 and 0.6 were considered borderline cases to account for the experimental error included in the binary variables (compounds resulting in <10% inhibition are considered non-inhibitors). The scores of the first two components of the initial PCA and PLS were compared to see if the PCA scores of the ALMOND model were well correlated to the y variable. The first components had a correlation coefficient of 0.9 and the second components had an  $r^2$  value of 0.56. Since the introduction of the y-variable enhanced the structure of the x-matrix seen in the PCA model the PLS model was further explored.

One iteration of GOLPE variable selection [37] was applied to the initial model. A predictive model was obtained using two components ( $r^2 = 0.74$ ,  $q^2_{(LOO)} = 0.64$   $q^2_{(5 \text{ random groups})} = 0.63$ ) (Fig. 2). The model was externally predicted using 39 compounds (14 competitive inhibitors and 25 non-inhibitors)

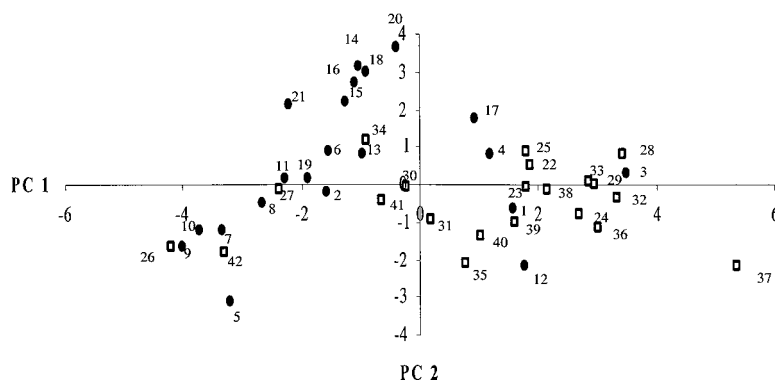


Figure 1. PCA score plot for inhibitors (●) and noninhibitors (□) based on GRIND descriptors in ALMOND.

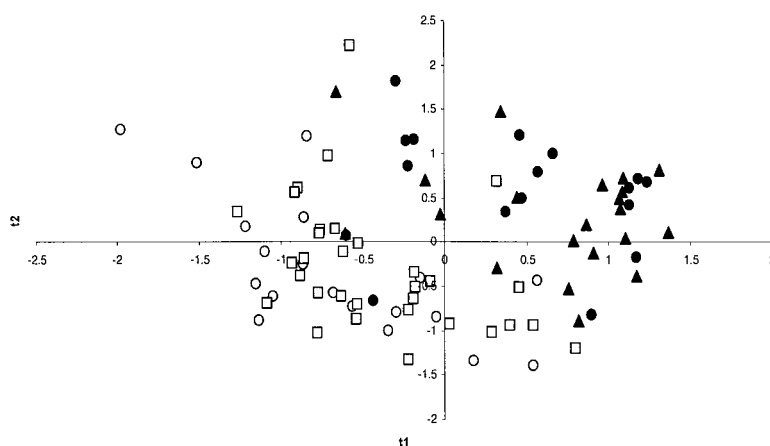


Figure 2. PLS t1 vs t2 score plot for the discriminative model. Inhibitors (▲) and noninhibitors (○) and for the externally predicted compounds, inhibitors (●) and noninhibitors (□).

yielding a  $SDEP_{\text{external}}$  of 0.38, which was an improvement, compared to the first model. This corresponds to 74% correct predictions and 13% as border cases. Altogether, only 13% of the 39 externally predicted compounds were false positives or negatives. The inhibitor quercetin (12) could not be well fitted to the model and the structurally similar compound kaempferol (45) was not predicted correctly. The DRY–DRY (hydrophobic–hydrophobic distances) interactions were comparable to the ones for the other inhibitors but the compounds' high ability to act as hydrogen bond donors and acceptors and their extensive ability to form polar interactions differentiate them from the other inhibitors in the test and training set. Three out of 25 externally predicted non-inhibitors were misclassified (cefoperazone (67), S-mephenytoin (79) and S-alprenolol (78)). In the case of S-mephenytoin, only a methyl group distinguishes the compound from the potent inhibitor phenytoin (6)

( $K_i = 6.0 \mu\text{M}$ ) and in literature S-mephenytoin is described as a poor inhibitor [38], though this poor affinity is not captured within the range of the experimental assay used here.

The normal distribution SDEP plot of the results from a full factorial design (Fig. 3a) showed that in two components the initial predictive power was significantly improved (lower SDEP value) by using only a subset of the correlograms. The DRY autocorrelogram (model 2- denoted 2 in figure 3a) or the DRY and 0-N1 auto and crosscorrelograms (model 3- denoted 34 in figure 3a) increased the predictivity the most. The normal probability effect plot (Fig. 3b) show whether these are statistically significant effects [33]. The increased predictivity in model 2 was statistically significant but in the case of model 3 this was not the case. Therefore model 3 was not further explored. Model 2 (DRY autocorrelograms) predicted without any variable selection 67% of the 39 externally pre-

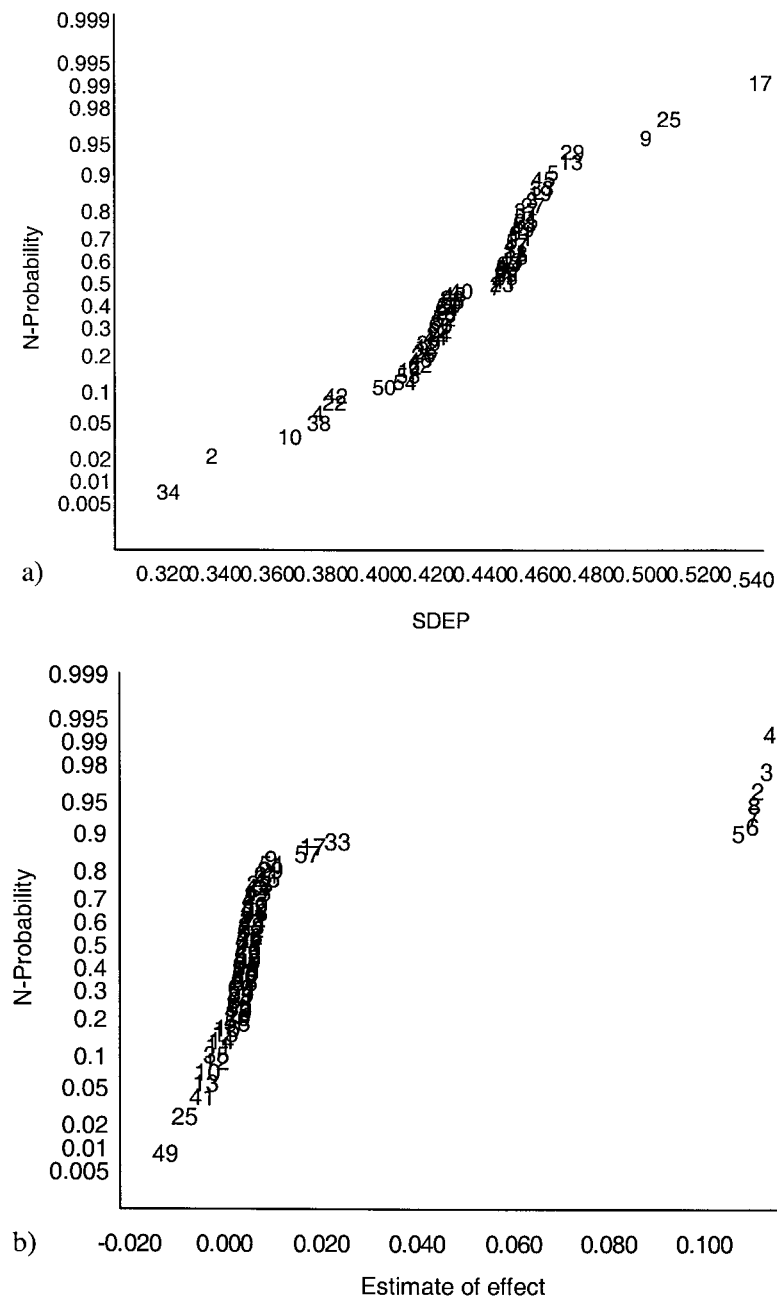


Figure 3. These plots have been drawn on normal probability scale where data that has a normal distribution will plot as a straight line. Figure 3a show the distribution of the SDEP values from the full factorial design to distinguish combinations of correlograms generating models that deviate from a normal distribution. Figure 3b show the estimate of effects based on Daniel's analysis to distinguish statistically significant changes [33].

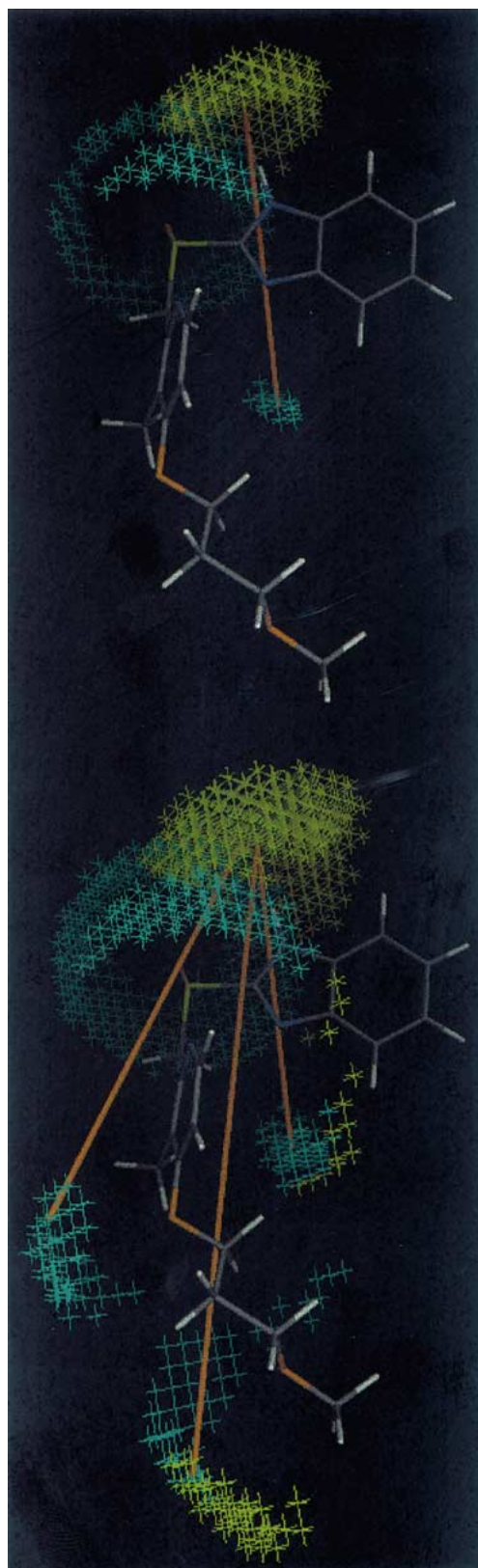


Figure 4. Filtered fields for S-rabeprazole selecting 150 points (left) and 500 points (right) comparing the different distances selected in the filtering process.

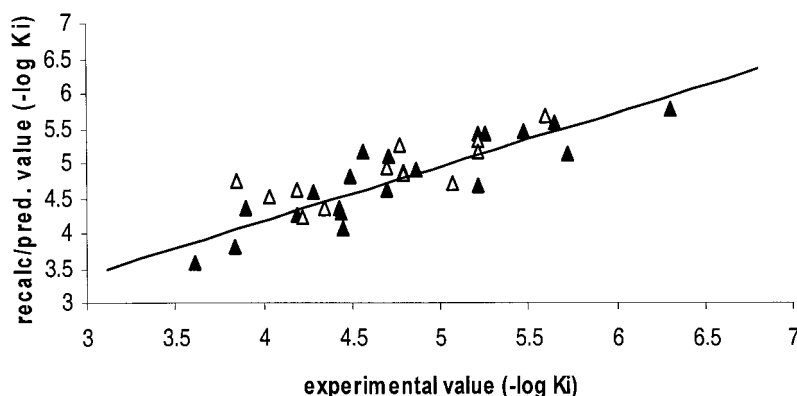


Figure 5. Quantitative CYP2C9 inhibitor model. Non-crossvalidated activities for training set (▲) and externally predicted activities of test set (Δ).

dicted compounds correctly, 15% as border cases and 18% were mispredicted. ( $r^2 = 0.70$ ,  $q^2_{(\text{LOO})} = 0.63$ ,  $q^2_{(5 \text{ random groups})} = 0.60$ ). The model was recalculated by increasing the number of points in the second step of the GRIND descriptors generation up to 600 but this did not improve the quality of the model. Quercetin, which was an outlier in the previous model, was now fitted well since its hydrophobic pattern correlates well to the other inhibitors. Instead SFID (19), progesterone (5) and 3R, 13S-pyranocoumarin (9) were less well described, since their ability to accept and donate hydrogen bonds and the polarity of the compounds were not taken into account in the model.

The relatively small change in interpretation and predictivity upon changing different parameters, such as number of correlograms or points, and the high correlation to the PCA scores signifies the robustness of the model. The external predictive capabilities of the two models (all correlograms (model 64) and DRY autocorrelograms only (model 2) are equal though they miss-predict different compounds.

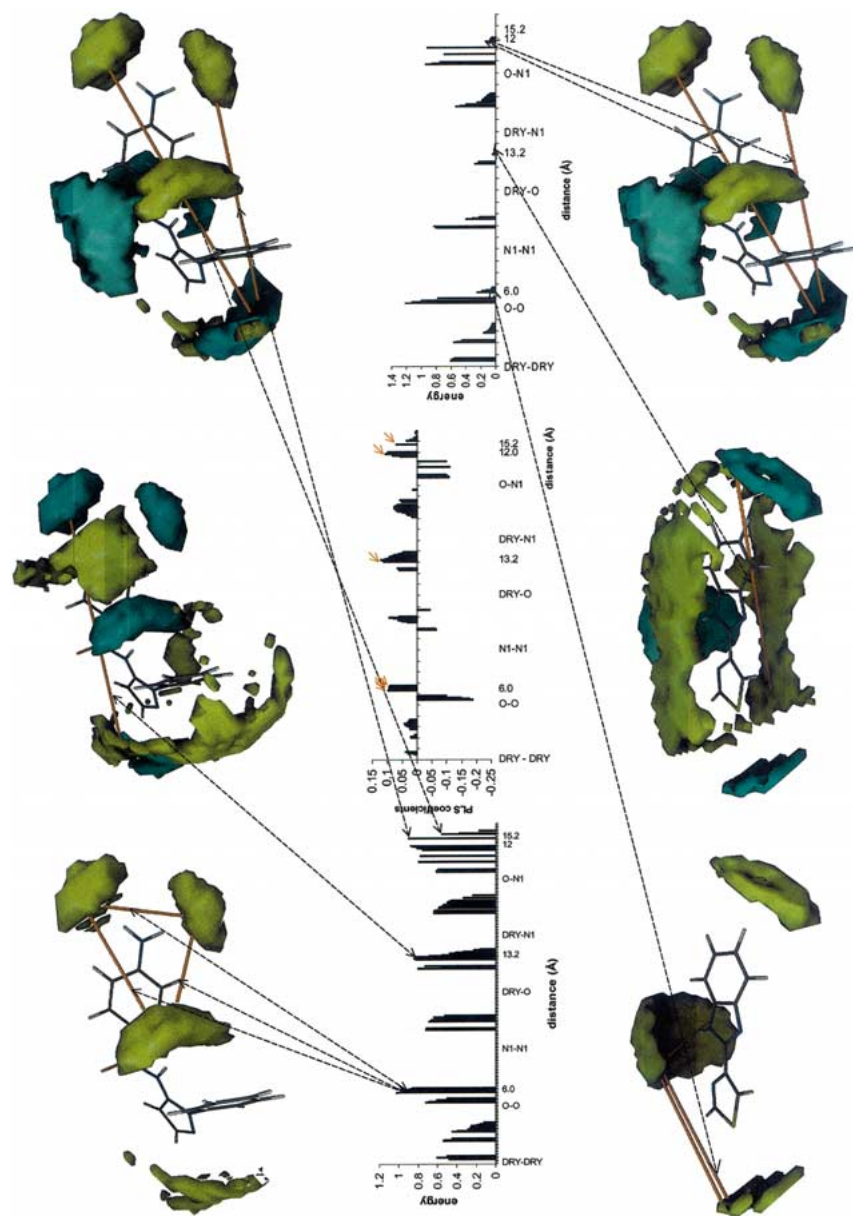
#### Quantitative model

In order to calculate a quantitative PLS model the x-matrix was correlated to experimentally determined  $K_i$ -values. Instead of trying to find common discriminant patterns like in PLS-DA, the model should explain as much variance as possible to enable the prediction of diverse compounds. The data set consisted of 33 structurally diverse competitive inhibitors (Tables 1 and 2) that describe a large chemical space [8]. The PLS coefficients for the model calculated using the DRY, O and N1 probe showed that the two first components contributed significantly to the information in the model. ( $r^2 = 0.78$ ,  $q^2_{(\text{LOO})} = 0.24$

,  $q^2_{(5 \text{ random groups})} = 0.24$ ). A full factorial design was performed in the same way as for the discriminant model to distinguish between statistical significant changes in SDEP values and the effect of interactions between correlograms but no combinations gave a significantly more predictive model. Therefore, no models with reduced number of correlograms were built.

The initial model was not internally or externally well predictive and no variables could be excluded by variable selection without model deterioration. To ensure that the number of points picked (150) was enough the model was compared to a model built from 500 points and weighted towards distance (25%) in the distance versus energy filtering step. Some of the larger compounds in the data set had specific long interactions that were neglected in the 150 points derived model since they were relatively low in energy values. In the case of S-rabeprazole (Fig. 4) two ether functions, that are able to form weak hydrogen bonds, are not described when only 150 points are picked in the filtering process. In the GRIND filtering process, selection of points is based on distances and energies and the benefits of including low energy interactions and thereby decreasing the signal-noise ratio has to be evaluated. In this case, expanding the number of points captures significant information that is not included in the former case. From the 263 x-variables one iteration of variable selection was applied which generated a model ( $r^2 = 0.77$ ,  $q^2_{(\text{LOO})} = 0.60$ ,  $q^2_{(5 \text{ random groups})} = 0.57$ ,  $\text{SDEP}_{\text{external}} = 0.37$  (Fig. 5) where 11 out of 12 compounds in the external test set were predicted within 0.5 log units.

In order to interpret the model the most potent compound, sulphaphenazole, was compared to the less



*Figure 6.* Upper row: The filtered GRID fields for the most potent inhibitor sulphaphenazole for the O-O (left), the DRY-O (center) and the O-N1 (right) molecular interaction fields. Lower row: The corresponding GRID fields for thiabendazole, the least potent inhibitor in the data set. The arrows pointing at the most important PLS coefficients of the model (center) correlate to the intramolecular distances shown in red. The energy of interaction and the relative importance of a certain interaction can be elucidated from the compound correlograms (center left – sulphaphenazole, center right – thiadiazole). The compound correlograms plot energy (y) over distance (x). The energy of interaction is very favorable for the potent inhibitor sulphaphenazole (high peaks) while the same distances are of little importance for thiadiazole, the least potent inhibitor (low peaks).

potent, thiabendazole (Fig. 6). From the PLS pseudo-coefficients of the model (center Fig. 6) the most relevant pharmacophoric distances can be determined (denoted in red arrows). An important hydrophobic region can be distinguished near the site of oxidation/ nitrogen-interaction for all template molecules and a hydrogen bond donor feature at a distance of 13 Å away. These results are in good agreement with present SAR [17]. Also a hydrogen bond acceptor possibility is observed in rather close proximity to the hydrophobic site. The interactions are also present for the less potent inhibitor thiabendazole but from the correlograms it can be concluded that they are not very energetically favorable. The distances found in the pharmacophore model can be well correlated to the homology model.

## Conclusion

In this study we have generated a model capable of distinguishing between CYP2C9 competitive inhibitors and non-inhibitors. We have also derived a pharmacophore model for quantitative prediction of competitive CYP2C9 inhibition. These models could be useful in the screening of chemical libraries for CYP2C9 inhibition and for predicting potential drug-drug interactions.

The discrimination between inhibitors and non-inhibitors was strongly correlated to clogP in the 2D space and lipophilicity could therefore be used as a simple descriptor to identify CYP2C9 inhibitors. In the 3D space the DRY probe was highly contributing to the discrimination. The two discriminative models that were generated were similar in predictive capability but mispredicted different compounds.

In the quantitative model, all correlograms were needed to explain structure-activity relationships. Examination of the filtered fields and correlograms revealed that the number of selected points needed to be increased in order to describe low energy, long distance interactions for large compounds. The model derived using the latter conditions explained better the 3D space covered by the inhibitors. The external predictions of the quantitative model are within 0.5 log units for all but one compound, which can be considered satisfactory from a drug design point of view. The external data set comprises both highly flexible and rigid compounds and also differences in molecular size, indicating that the predictive power is valid over

a wide range, which is desirable when predicting the properties of unknown compounds.

## References

1. Stockman, B.J., et al., *FEBS Lett.*, 283 (1991) 267–269.
2. Schaal, W., et al., *J. Med. Chem.*, 44 (2001) 155–169.
3. Martin, Y.C., et al., *J. Comput-Aided Mol. Design*, 7 (1992) 83–102.
4. CATALYST, Molecular Simulations: San Diego, CA.
5. Ekins, S., et al., *J. Pharmacol. Exp. Ther.*, 288 (1999) 21–29.
6. Ekins, S., M.J. de Groot, and J.P. Jones, *Drug Metab. Dis.*, 29 (2001) 936–944.
7. Jones, J.P., et al., *Drug Metab. Dis.*, 24 (1996) 1–6.
8. Afzelius, L., et al., *Mol. Pharmacol.*, 59 (2001) 909–919.
9. Goodford, P.J., *J. Med. Chem.*, 28 (1985) 849–857.
10. Cramer, R.D.r., D.E. Patterson, and J.D. Bunce, *J. Am. Chem. Soc.*, 110 (1988) 5959–5967.
11. Pastor, M., et al., *J. Med. Chem.*, 43 (2000) 3233–3243.
12. Riganelli, D., et al., *Pharm. Pharmacol. Lett.*, 3 (1993) 5–8.
13. Smith, D.A., M.J. Ackland, and B.C. Jones, *Drug Discovery Today*, 2 (1997) 406–414.
14. Murray, M., *Int. J. Mol. Med.*, 3 (1999) 227–238.
15. Rendic, S. and F.J. Di Carlo, *Drug Metab. Rev.*, 29 (1997) 413–580.
16. Masimirembwa, C.M., et al., *Meth. Enzymol.*, (2001) submitted.
17. Smith, D.A., M.J. Ackland, and B.C. Jones, *Drug Discovery Today*, 2 (1997) 479–485.
18. Sjöström, M., S. Wold, and B. Söderström. *Proc. PARC in Practice*, (1985) North-Holland.
19. Wold, S., et al., *Chemometrics: Mathematics and Statistics in Chemistry*, K.B. R, Editor. 1984, Reidel Publishing Company, Holland: Dordrecht.
20. Bardsley, W.G., et al., *Comput. Chem.*, 19 (1995) 75–84.
21. Masimirembwa, C.M., et al., *Drug Metabol. Dis.*, 27 (1999) 1117–1122.
22. Wold, S., K. Esbensen, and P. Geladi, *Chem. Intell. Lab. Syst.*, 2 (1987) 37–52.
23. Oprea, T.I. and J. Gottfries, *J. Combin. Chem.*, 3 (2001) 157–166.
24. Olsson, T. and V. Shcherbukhin, *Synthesis and Structure Administration (SaSA)*. 1997–2000, AstraZeneca, <http://www.astrazeneca.com>.
25. Cruciani, G., M. Pastor, and W. Guba, *Eur. J. Pharmaceut. Sci.*, 11 (2000) S29–39.
26. Miners, J.O. and D.J. Birkett, *B. J. Clin. Pharmacol.*, 45 (1998) 525–538.
27. Rao, S., et al., *J. Med. Chem.*, 43 (2000) 2789–2796.
28. Cruciani, G. and K.A. Watson, *J. Med. Chem.*, 37 (1994) 2589–2601.
29. Baroni, M., et al., *Quantit. Struct. Act. Rel.*, 12 (1993) 225–231.
30. Fedorov, V.V., *Theory of Optimal Experiments*. 1972, New York: Academic Press.
31. Albano, C., et al., *Anal. Chim. Acta*, 103 (1978) 429–443.
32. Cruciani, G., et al., *Predictive Ability of Regression Models. Part I: Standard Deviation of Prediction Errors (SDEP)*. *Journal of Chemometrics*, 1992. 6: p. 335–346.
33. Box, G.E.P., W.G. Hunter, and S.J. Hunter, *Statistics for experimenters*. 1978: John Wiley.
34. Baroni, M., et al., *Quantit. Struct. Act. Rel.*, 12 (1993) 9–20.

35. Zamora, I., T.I. Oprea, and A.L. Ungell, *J. Med. Chem.*, (2001) submitted.
36. Hansch, C., *Drug Metab. Rev.*, 1 (1972) 1–14.
37. Baroni, M., et al., *Quantit. Struct. Act. Rel.*, 12 (1993) 9–20.
38. Miners, J.O., et al., *Biochem. Pharmacol.*, 37 (1988) 1137–1144.