

# Stereochemical and Steric Control of the UDP-Glucuronosyltransferase-Catalyzed Conjugation Reaction: A Rational Approach for the Design of Inhibitors for the Human UGT2B7

Ingo Bichlmaier,<sup>[a]</sup> Moshe Finel,<sup>[b]</sup> Wolfgang Sippl,<sup>[c]</sup> and Jari Yli-Kauhaluoma<sup>\*[a]</sup>

A set of 76 derivatives of the epimeric tricyclic sesquiterpenols longifolol and isolongifolol was subjected to inhibition and glucuronidation assays employing the human UDP-glucuronosyltransferase (UGT) 2B7. Detailed structure–activity relationships (SARs) with respect to functionality, stereochemical properties, and steric features were derived. To gain further insight into the SARs of UGT2B7 ligands herein, we have developed a 3D-quantitative structure–activity relationship (3D-QSAR) using Comparative Molecular Similarity Analysis (CoMSIA). The formation of the enzyme–inhibitor complex was predominantly controlled by spa-

tially directed hydrophobic interactions. The glucuronidation rate was significantly influenced by the steric demand of substituents in proximity of the nucleophilic hydroxy group. The glucuronidation of the compounds was prevented by the introduction of bulky substituents such as isopropyl, tert-butyl, and phenyl groups. The epimeric longifolol derivatives of series D were the best inhibitors displaying  $IC_{50}$  values as low as 4.6 nM. This study shows that high-potency substrates can be turned into potent inhibitors by addressing functional, stereochemical, and steric properties.

## Introduction

The UDP-glucuronosyltransferases (UGTs, EC 2.4.1.17) are important enzymes of the phase II metabolic system, catalyzing the transfer of glucuronic acid (GlcA) from UDP- $\alpha$ -D-glucuronic acid (UDPGlcA) to nucleophilic groups of xenobiotic and endobiotic compounds.<sup>[1]</sup> Common functional groups that are glucuronidated by UGT enzymes comprise hydroxy, carboxy, thiol, amino, amide, sulfonamide, and enolate groups. The resulting O-, N-, S-, and C-glucuronides display higher water solubility in comparison to their parent aglycones, hence, their urinary and biliary excretion is elevated.<sup>[1d]</sup> The detailed reaction mechanism of the enzyme-catalyzed conjugation reaction is essentially unknown because the full-length X-ray crystal structure of any of the UGT isoforms has not been resolved to date. However, it is assumed that the enzymatic glucuronidation resembles the bimolecular nucleophilic substitution reaction because of the inversion of configuration at the anomeric carbon atom in the GlcA moiety (Walden-inversion) yielding exclusively  $\beta$ -D-glucuronides.

Metabolic enzymes are commonly described as possessing a high degree of flexibility, or promiscuity, to depict their partly overlapping substrate selectivities.<sup>[2]</sup> These enzymes apparently evolved to possess broad-substrate selectivities so that they can detoxify a wide range of structurally different compounds. Nineteen UGT isoforms belonging either to subfamily 1A or 2B have been described, but merely seven UGT enzymes are assumed to be involved in the conjugation of drugs.<sup>[3]</sup> The single most important UGT enzyme responsible for drug glucuronidation is presumably UGT2B7, which catalyzes the glucuronidation of  $\approx 40\%$  of drugs that are metabolized by UGT enzymes.<sup>[3b]</sup>

The design of potent inhibitors for UGT enzymes is challenging because the full-length crystal structure of any of the UGT enzymes remains to be resolved. Yet another obstacle in the development of potent inhibitors is that common functional groups that promote water solubility such as hydroxy, amino, thiol, and carboxy groups serve as nucleophiles in the enzymatic glucuronidation reaction. Therefore, inhibitors that possess such chemical functionalities are at risk of being substrates. Moreover, UGTs in general exert low affinities toward their substrates, which renders it difficult to identify structural features that result in high potency.<sup>[3b]</sup> A recent study identified fatty acids such as oleic, linoleic, and arachidonic acid present in, or released during, incubations that might act as inhibitors of UGT isoforms resulting in the apparent high  $K_m$  values observed for UGT substrates.<sup>[4]</sup>

Potent and selective inhibitors are important as biochemical tools to identify the enzymes responsible for drug metabolism and these inhibitors can also be used in tissue preparations. They can be applied to elucidate drug–drug interactions, for

[a] I. Bichlmaier, Prof. J. Yli-Kauhaluoma  
Faculty of Pharmacy, Division of Pharmaceutical Chemistry  
University of Helsinki, P.O. Box 56, 00014 Helsinki (Finland)  
Fax: (+358) 9-19159556  
E-mail: jari.yli-kauhaluoma@helsinki.fi

[b] Dr. M. Finel  
Faculty of Pharmacy, Drug Discovery and Development Technology Center  
University of Helsinki, P.O. Box 56, 00014 Helsinki (Finland)

[c] Prof. Dr. W. Sippl  
Department of Pharmaceutical Chemistry  
Martin Luther University of Halle-Wittenberg, 06120 Halle (Germany)

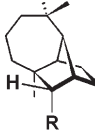
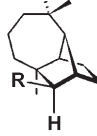
studying enzyme polymorphism, and to estimate the number of distinct binding-domains of a certain UGT enzyme.

Some inhibitors of UGT enzymes have been described, but most of them lack desirable levels of potency.<sup>[5]</sup> Even transition state (TS) mimics for UGT enzymes have been proposed.<sup>[6]</sup> However, these inhibitors lack the expected high potency levels that are commonly associated with TS analogues.<sup>[7]</sup> This is probably due to the shortfall in the elucidation of TS defining properties such as bond lengths, bond angles, charge distribution, electrostatic potential surfaces, and potency-promoting guest–host interactions.<sup>[8]</sup> Many attempts to design UGT inhibitors and TS analogues were based on the preparation of bisubstrate analogues by attaching a UDP-like moiety to a lipophilic aglycone-like entity.<sup>[9]</sup> Most of these inhibitors display a high degree of flexibility due to the incorporation of aliphatic moieties that possess many freely rotating  $\sigma$ -bonds. This structural flexibility results in numerous conformers, which renders it difficult to determine the biologically active conformation of the inhibitor also because the geometry of the UGT-binding site is not known. These approaches did not result in the development of potent UGT inhibitors or the derivation of applicable structure–activity relationships that could be used for further inhibitor design. The design of UGT inhibitors based on rational principles such as detailed structure–activity relationships employing whole series of compounds has not been conducted to date.

We have recently shown that the rate of the UGT2B7- and UGT2B17-catalyzed glucuronidation reaction was significantly influenced by the stereochemical properties of the substrate, that is, the spatial orientation of the nucleophilic functionality.<sup>[10]</sup> The epimeric tricyclic sesquiterpenoid alcohols isolongifolol (**1a**) and longifolol (**1b**) were identified as substrates of UGT2B7 and they displayed low competitive inhibition constants ( $K_i$ ) of 26 and 23 nM, respectively.<sup>[11]</sup> Therefore, these epimeric alcohols were chosen as lead compounds. Furthermore, these sesquiterpenols were appropriate lead compounds for the derivation of structure–activity relationships because of the rigidity of the tricyclo[5.4.0.0<sup>2,9</sup>]undecane framework.

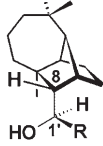
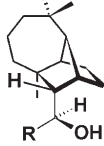
In this study, the potency toward UGT2B7 and the rate of the enzyme-catalyzed glucuronidation of 76 derivatives of the sesquiterpenols isolongifolol (**1a**) and longifolol (**1b**) were measured (Tables 1 and 2). The compounds are related as stereoisomers (epimers, diastereomers), and geometric isomers (*E*-, *Z*-isomers), and they bear different functional groups,

**Table 1.** Longifolol and isolongifolol derivatives.

	Isolongifolol derivatives	Longifolol derivatives
		
R		
Series A: primary alcohols		
CH <sub>2</sub> OH	<b>1a</b>	<b>1b</b>
(CH <sub>2</sub> ) <sub>2</sub> OH	<b>2a</b>	<b>2b</b>
(CH <sub>2</sub> ) <sub>3</sub> OH	<b>3a</b>	<b>2c</b>
CH=CHCH <sub>2</sub> OH	( <i>E</i> )- <b>4a</b>	
	( <i>Z</i> )- <b>4a</b>	
(CH <sub>2</sub> ) <sub>4</sub> OH	<b>5a</b>	<b>5b</b>
Series B: carboxylic/phosphonic acids, amines		
COOH	<b>6a</b>	<b>6b</b>
(CH <sub>2</sub> ) <sub>2</sub> COOH	<b>7a</b>	<b>7b</b>
CH=CHCOOH	( <i>E</i> )- <b>8a</b>	
	( <i>Z</i> )- <b>8a</b>	
(CH <sub>2</sub> ) <sub>3</sub> COOH	<b>9a</b>	<b>9b</b>
(CH <sub>2</sub> ) <sub>2</sub> P(O)(OH) <sub>2</sub>	<b>10a</b>	
CH=CHP(O)(OH) <sub>2</sub>	( <i>E</i> )- <b>11a</b>	
(CH <sub>2</sub> ) <sub>3</sub> P(O)(OH) <sub>2</sub>	<b>12a</b>	
CH <sub>2</sub> CH=	( <i>E</i> )- <b>13a</b>	
CHP(O)(OH) <sub>2</sub>		
(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	<b>14a</b>	<b>14b</b>

namely hydroxy, carboxy, phosphono, ester, amino, and alkyl groups of varying steric demand. Furthermore, the spatial orientation of these functional groups is different because they

**Table 2.** Longifolol and isolongifolol derivatives (continued).

	Isolongifolol derivatives	Longifolol derivatives
		
R		
Series C: hydroxy acids, hydroxy esters, diols		
CH <sub>2</sub> COOH	<b>15a</b>	<b>15b</b>
CH <sub>2</sub> COOEt	<b>16a</b>	<b>16b</b>
CH <sub>2</sub> P(O)(OH) <sub>2</sub>	<b>17a</b>	<b>17b</b>
CH <sub>2</sub> P(O)(OEt) <sub>2</sub>	<b>18a</b>	<b>18b</b>
(CH <sub>2</sub> ) <sub>2</sub> P(O)(OH) <sub>2</sub>	<b>19a</b>	<b>19b</b>
(CH <sub>2</sub> ) <sub>2</sub> P(O)(OMe) <sub>2</sub>	<b>20a</b>	<b>20b</b>
(CH <sub>2</sub> ) <sub>2</sub> P(O)(OEt) <sub>2</sub>	<b>21a</b>	<b>21b</b>
(CH <sub>2</sub> ) <sub>2</sub> OH	<b>22a</b>	<b>22b</b>
		<b>22c</b>
		<b>22d</b>
Series D: compounds bearing lipophilic substituents		
CH <sub>2</sub> Cl	<b>23a</b>	<b>23b</b>
Me	<b>24a</b>	<b>24b</b>
Et	<b>25a</b>	<b>25b</b>
CH=CH <sub>2</sub>	<b>26a</b>	<b>26b</b>
CH <sub>2</sub> CH=CH <sub>2</sub>	<b>27a</b>	<b>27b</b>
<i>i</i> Pr	<b>28a</b>	<b>28b</b>
<i>t</i> BU	<b>29a</b>	<b>29b</b>
Ph	<b>30a</b>	<b>30b</b>
		<b>24c</b>
		<b>25c</b>
		<b>26c</b>
		<b>27c</b>
		<b>28c</b>
		<b>29c</b>
		<b>30c</b>

are either *endo*-positioned (isolongifolol derivatives) or *exo*-oriented (longifolol derivatives). In the case of the compounds of series C and D, an additional asymmetric carbon atom was introduced to the side chain at C1' resulting in two epimers for each isolongifolol (**1a**) and longifolol (**1b**) derivative. These variations in functionality, steric demand, and spatial orientation enabled the derivation of detailed structure–activity relationships with respect to potency and glucuronidation rate to develop high-potency inhibitors for the human UGT2B7 enzyme.

To gain further insight into the structure–activity relationships (SARs) of longifolol and isolongifolol derivatives, herein we have carried out a 3D-QSAR study on the data set of 76 UGT2B7 inhibitors. For this purpose we used a ligand-based alignment and the CoMSIA methodology.<sup>[12]</sup> The SARs for the compounds are interpreted using contour maps of the PLS coefficients obtained from the 3D-QSAR model. It should be mentioned here that even though relevant QSAR studies for some UGT isoforms have been conducted, these were not employed for the development of inhibitors for UGT enzymes.<sup>[13]</sup>

## Results and Discussion

### Potency assays

The IC<sub>50</sub> value of each longifolol and isolongifolol derivative toward UGT2B7 was determined by the use of the reference substrate estriol. The IC<sub>50</sub> was employed as a measure of the potency toward the binding site of the enzyme because it relates to the dissociation (formation) of the enzyme-inhibitor complex. The UGT2B7-substrates longifolol (**1a**) and isolongifolol (**1b**) were employed as reference compounds to indicate high potency (*vide supra*). The IC<sub>50</sub> values along with the standard deviations are displayed in Table 3. The IC<sub>50</sub> ratio was calculated from these data and it reflects the preferential dissociation of the enzyme-inhibitor complex of one compound over its respective stereoisomer or homologue.

The IC<sub>50</sub> values for series A ranged from 76 nM (**1b**) to 1.5 μM (**5a**) and the mean was 0.81 μM (Table 3 and Figure 1a). The potency toward UGT2B7 decreased with increasing length of the side chain, and, accordingly, the hydroxybutyl derivatives **5a** and **5b** displayed the highest IC<sub>50</sub> values of 1.5 and 1.4 μM, respectively. The spatial arrangement of the side chain had no significant effect on the potency because the IC<sub>50</sub> ratios of the epimeric alcohols were close to unity (IC<sub>50</sub> ratios = 1.1–1.5, Table 3 and Figure 1b), and also the geometric isomers (*E*)-**4a** and (*Z*)-**4a** displayed similar potency toward the binding domain of UGT2B7 (IC<sub>50</sub> ratio = 1.2).

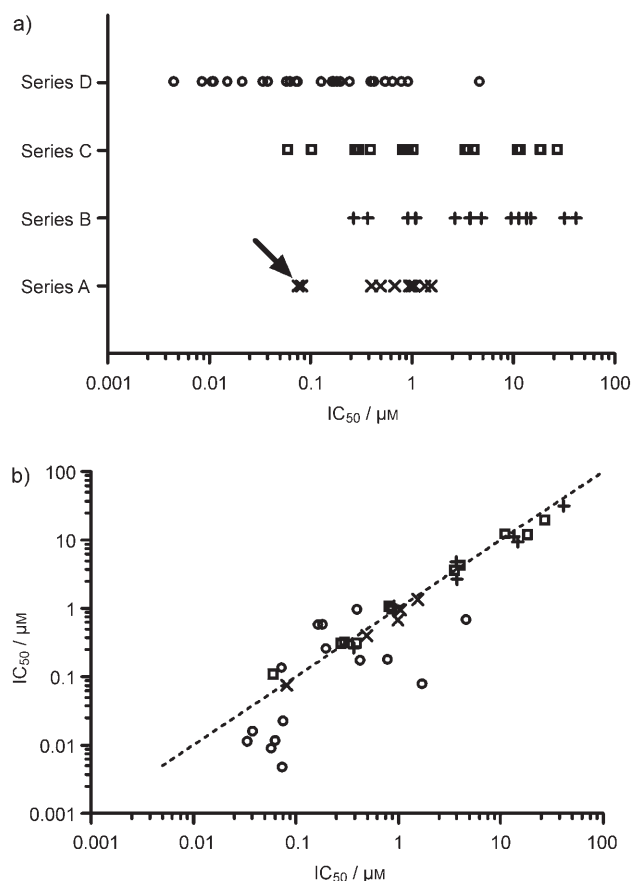
The mean of the IC<sub>50</sub> values for series B, 9.7 μM, was significantly higher than that of series A (ttest:  $P < 0.05$ ; Table 3 and Figure 1a). The IC<sub>50</sub> values ranged from 0.27 μM for compound **6b** to 41 μM for compound (*E*)-**13a** and the potency toward the enzyme decreased with increasing length of the side chain, in agreement with the results for series A. The best inhibitor in series B was compound **6b**. However, its IC<sub>50</sub> was 3.6-fold higher than that of its homologous alcohol **1b**. Also the other carboxylic and phosphonic acids, as well as the amines

**Table 3.** IC<sub>50</sub> values and IC<sub>50</sub> ratios.

Compd	IC <sub>50</sub> [μM] <sup>[a]</sup>	Compd	IC <sub>50</sub> [μM] <sup>[a]</sup>	IC <sub>50</sub> ratio <sup>[b]</sup>
Series A				
<b>1a</b>	0.082 ± 0.0057	<b>1b</b>	0.076 ± 0.046	1.1
<b>2a</b>	0.49 ± 0.044	<b>2b</b>	0.40 ± 0.038	1.2
<b>3a</b>	0.99 ± 0.080	<b>3b</b>	0.68 ± 0.036	1.5
( <i>E</i> )- <b>4a</b>	1.1 ± 0.098	( <i>Z</i> )- <b>4a</b>	0.95 ± 0.06	1.2
<b>5a</b>	1.5 ± 0.16	<b>5b</b>	1.4 ± 0.14	1.1
Series B				
<b>6a</b>	0.37 ± 0.029	<b>6b</b>	0.27 ± 0.028	1.4
<b>7a</b>	0.91 ± 0.059	<b>7b</b>	1.1 ± 0.089	1.2
( <i>E</i> )- <b>8a</b>	1.7 ± 0.14	( <i>Z</i> )- <b>8a</b>	2.1 ± 0.29	1.2
<b>9a</b>	14 ± 0.85	<b>9b</b>	11 ± 1.3	1.2
<b>10a</b>	15 ± 1.5	( <i>E</i> )- <b>11a</b>	9.5 ± 0.70	1.5 <sup>[c]</sup>
<b>12a</b>	41 ± 3.0	( <i>E</i> )- <b>13a</b>	32 ± 1.8	1.3 <sup>[c]</sup>
<b>14a</b>	3.8 ± 0.37	<b>14b</b>	2.7 ± 0.19	1.4
Series C				
<b>15a</b>	0.84 ± 0.070	<b>15b</b>	1.0 ± 0.081	1.2
<b>15c</b>	0.89 ± 0.053	<b>15d</b>	0.95 ± 0.096	1.1
<b>16a</b>	0.31 ± 0.025	<b>16b</b>	0.31 ± 0.018	1.0
<b>16c</b>	0.40 ± 0.027	<b>16d</b>	0.30 ± 0.024	1.3
<b>17a</b>	19 ± 1.8	<b>17b</b>	12 ± 0.77	1.6
<b>18a</b>	7.3 ± 0.82	<b>18b</b>	6.2 ± 0.49	1.2
<b>19a</b>	28 ± 3.0	<b>19b</b>	19 ± 1.4	1.5
<b>20a</b>	3.6 ± 0.23	<b>20b</b>	3.4 ± 0.28	1.1
<b>21a</b>	4.2 ± 0.34	<b>21b</b>	4.1 ± 0.25	1.0
<b>22a</b>	0.083 ± 0.0037	<b>22b</b>	0.10 ± 0.016	1.2
<b>22c</b>	0.14 ± 0.018	<b>22d</b>	0.17 ± 0.025	1.2
Series D				
<b>23a</b>	0.074 ± 0.0082	<b>23b</b>	0.13 ± 0.0074	1.8
<b>24a</b>	0.40 ± 0.039	<b>24b</b>	0.94 ± 0.087	2.4
<b>24c</b>	0.077 ± 0.0053	<b>24d</b>	0.022 ± 0.0018	3.5
<b>25a</b>	0.17 ± 0.0066	<b>25b</b>	0.56 ± 0.041	3.3
<b>25c</b>	0.065 ± 0.0054	<b>25d</b>	0.011 ± 0.0070	5.9
<b>26a</b>	0.19 ± 0.015	<b>26b</b>	0.56 ± 0.037	2.9
<b>26c</b>	0.059 ± 0.0048	<b>26d</b>	0.0087 ± 0.0010	6.8
<b>27a</b>	0.20 ± 0.026	<b>27b</b>	0.25 ± 0.014	1.3
<b>27c</b>	0.039 ± 0.0024	<b>27d</b>	0.016 ± 0.0010	2.4
<b>28a</b>	0.43 ± 0.027	<b>28b</b>	0.17 ± 0.015	2.5
<b>28c</b>	0.034 ± 0.0022	<b>28d</b>	0.011 ± 0.0010	3.1
<b>29a</b>	0.81 ± 0.052	<b>29b</b>	0.17 ± 0.011	4.8
<b>29c</b>	4.7 ± 0.11	<b>29d</b>	0.66 ± 0.026	7.1
<b>30a</b>	0.84 ± 0.064	<b>30b</b>	0.072 ± 0.0019	11
<b>30c</b>	0.075 ± 0.0089	<b>30d</b>	0.0046 ± 0.0005	16

[a] Mean values ± SD ( $n = 15$ ). [b] The higher IC<sub>50</sub> value in the numerator. [c] Comparison between the saturated phosphonic acid and its α,β-unsaturated homologue.

**14a** and **14b**, displayed markedly lower potency toward UGT2B7 in comparison to their corresponding homologous alcohols. This result indicated that the polar acidic and basic groups exerted weaker attractive interactions toward the binding domain of the enzyme. The saturated and α,β-unsaturated phosphonic acids **12a** and (*E*)-**13a** displayed the highest IC<sub>50</sub> values, which were significantly higher than those of their corresponding butanoic acid homologue **9a** (ttest:  $P < 0.05$ ). Accordingly, the phosphonic acids **10a** and (*E*)-**11a** also exerted lower potency than their corresponding homologous carboxylic acids **7a** and (*E*)-**8a**, respectively (ttest:  $P < 0.05$ ). These re-



**Figure 1.** a) Distribution of the  $IC_{50}$  values for each series of compounds. The  $IC_{50}$  values are displayed on a logarithmic scale for clarity. The  $IC_{50}$  values of longifolol and isolongifolol are indicated by the arrow; b)  $IC_{50}$  values of the tricyclic compounds versus the  $IC_{50}$  value of their corresponding isomers. The  $IC_{50}$  values are displayed on a logarithmic scale for clarity. The bisecting dashed line corresponds to an  $IC_{50}$  ratio of unity. Only the mean values are displayed for clarity.

sults might stem from the higher acidity and increased steric demand of the phosphono group compared to the smaller and less acidic carboxy group. The spatial arrangement of the side chain and geometric isomerism had no significant influence on the  $IC_{50}$  values because the  $IC_{50}$  ratios of the corresponding isomers were close to unity ( $IC_{50}$  ratios = 1.2–1.5, Table 3 and Figure 1b). These findings were consistent with the results for series A.

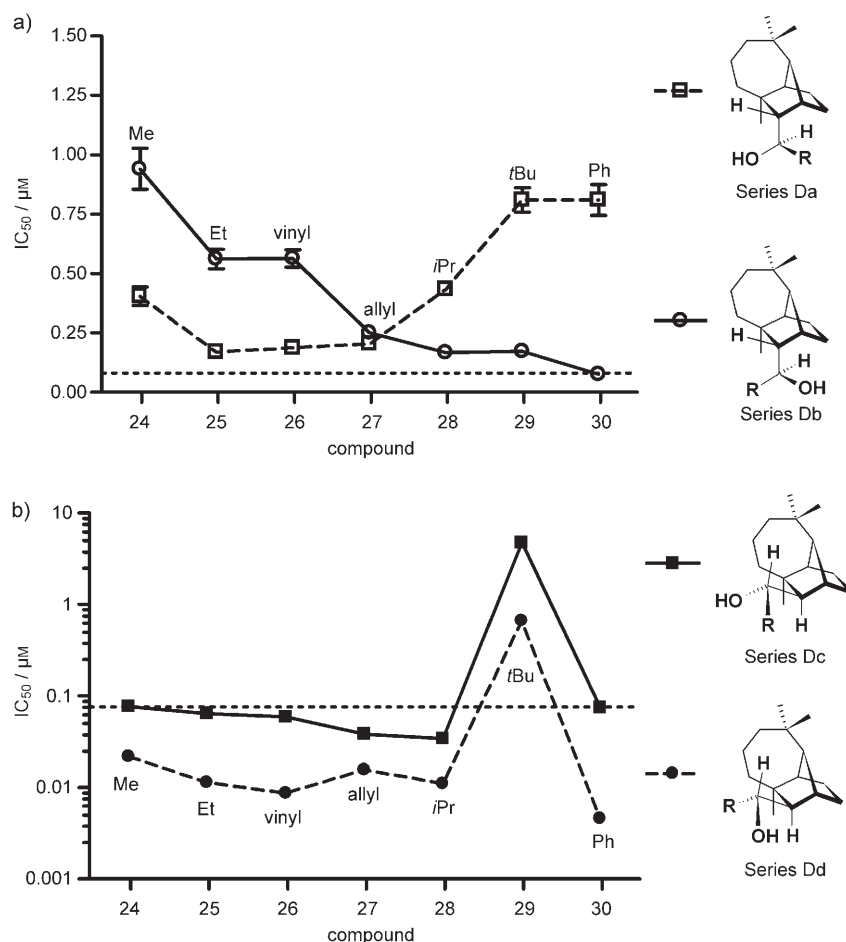
The results of series A and B revealed that the elongation of the side chain resulted in decreased potency toward the binding site of UGT2B7. This might stem from an increased flexibility due to a higher number of rotating  $\sigma$ -bonds, resulting in a higher entropy cost during the formation of the encounter complex. Furthermore, the introduction of the polar carboxy, phosphono, and amino groups into the side chain lowered the potency toward the enzyme. This might indicate that lipophilic interactions are favored during the formation of the enzyme–inhibitor complex. Interestingly, the spatial arrangement of the side chain bearing the different functional groups had no significant effect on the potency, indicating that there was no significant recognition of the functional groups during the forma-

tion of the encounter complex. Isolongifolol (**1a**) and longifolol (**1b**) displayed the lowest  $IC_{50}$  values in series A and B and it was therefore tested whether the hydroxy group in this position had a significant influence on the potency. Therefore, the epimers and diastereomers of series C and D, in which the hydroxy group was retained, were assayed.

The  $IC_{50}$  values of the compounds of series C ranged from 62 nM for compound **22a** to 28  $\mu$ M for compound **19a** (Table 3 and Figure 1a). The *endo*- or *exo*-arrangement of the side chain had no significant effect on the inhibition level because the epimeric and diastereomeric compounds displayed  $IC_{50}$  ratios close to unity. Interestingly, also the absolute configuration of the asymmetric carbon atom C1' had no significant effect on the inhibition levels. In this respect, the epimeric and diastereomeric  $\beta$ -hydroxy carboxylic acids **15a–d** displayed  $IC_{50}$  values ranging from 0.84 to 1.0  $\mu$ M and also the diols **22a–d** exerted similar potency levels toward the enzyme. Consistent with the results of series B, the  $\beta$ - and  $\gamma$ -phosphonic acids **17a**, **17b**, **19a**, and **19b** possessed the highest  $IC_{50}$  values (12–19  $\mu$ M). The epimeric and diastereomeric carboxylic and phosphonic acid esters (**16a–d**, **18a**, **18b**, **20a**, **20b**, **21a**, **21b**) had smaller  $IC_{50}$  values than their corresponding free acids, suggesting that lipophilic interactions are favored during the formation of the enzyme–inhibitor complex. The  $IC_{50}$  values of the bifunctional derivatives were comparable to those of the monofunctional derivatives of series B and the results indicate that the hydroxy group at the asymmetric carbon atom in the side chain did not markedly contribute to the potency toward UGT2B7.

The  $IC_{50}$  values of the derivatives bearing lipophilic substituents (series D) were between 4.6 nM (**30d**) and 4.7  $\mu$ M (**29c**) (Table 3 and Figure 1a). Most of the  $IC_{50}$  values of the compounds of series D were distinctly smaller than those of the other series indicating that lipophilic interactions toward the binding site of UGT2B7 were favored. Interestingly, the epimeric and diastereomeric derivatives also displayed significantly different  $IC_{50}$  values resulting in  $IC_{50}$  ratios that deviated significantly from unity (Table 3 and Figure 1b). As can be seen from Figure 1b, the compounds of series A, B, and C displayed  $IC_{50}$  ratios close to 1.0, whereas those of the derivatives from series D ranged from 1.3 to 16 and only two out of 15 values were below an  $IC_{50}$  ratio of 2.0. It was therefore concluded that the spatial orientation of lipophilic substituents had a marked effect on the formation of the enzyme–inhibitor complex indicating specific attractive interactions toward the UGT2B7-binding site.

As can be seen from Figure 2a, the isolongifolol (**1a**) derivatives (series Da and Db) to an extent displayed the inverse trend. In the case of series Db, the  $IC_{50}$  values decreased with increasing steric demand of the substituent R. In contrast, the compounds of series Da displayed lower affinities when the steric demand of the substituent R was increased. The longifolol (**1b**) derivatives of series Dc and Dd displayed a parallel trend (Figure 2b) and the  $IC_{50}$  values of the compounds of series Dd were constantly lower than those of their corresponding epimers of series Dc. Interestingly, the *tert*-butyl-substituted derivatives **29c** and **29d** displayed high  $IC_{50}$  values of 4.7 and



**Figure 2.** IC<sub>50</sub> values of the stereoisomers of series D. The mean values and standard deviations are shown ( $n = 15$ ). The IC<sub>50</sub> values in panel b are displayed on a logarithmic scale for clarity. The dashed lines depict the IC<sub>50</sub> values of isolongifolol **1a** (panel a) and longifolol **1b** (panel b).

0.66 μM, respectively. This indicated that the bulky *tert*-butyl group induced repulsive interactions toward the binding site of UGT2B7. In contrast, the isopropyl and phenyl derivatives displayed significantly lower IC<sub>50</sub> values ( $t$ -test:  $P < 0.05$ ). The epimeric longifolol derivatives (series Dc and Dd) were better inhibitors than the epimeric isolongifolol derivatives (series Da and Db)—with the exception of compounds **29c** and **29d**. The longifolol derivatives displayed higher IC<sub>50</sub> values than isolongifolol (**1a**), with the only exception of compound **30b** (IC<sub>50</sub> = 72 nM). In contrast, the epimeric longifolol derivatives (series Dc and Dd) displayed higher affinities toward UGT2B7 than longifolol—with the exception of compounds **29c** and **29d**. The best inhibitor was the phenyl-substituted secondary alcohol **30d** (IC<sub>50</sub> = 4.6 nM). This analysis showed that the spatial orientation of lipophilic substituents had an appreciable effect on the inhibition level. In this respect, the observed trends and the rather large differences in IC<sub>50</sub> values (IC<sub>50</sub> ratios > 2.0) indicated specific and spatially directed interactions between the lipophilic substituent R and the binding site of UGT2B7. The longifolol derivatives of series Dd were found to exert the highest potency levels toward UGT2B7.

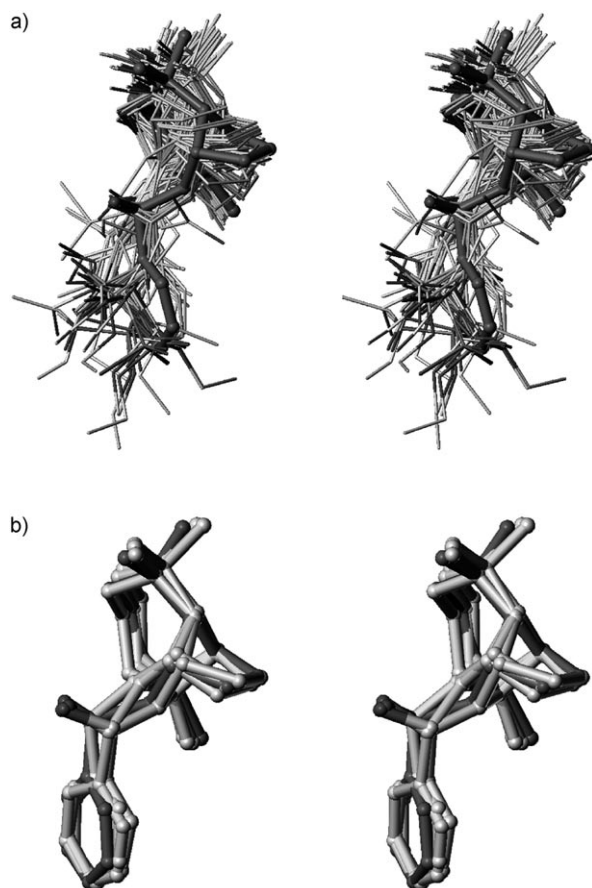
The potency studies revealed that the potency depends on functionality, length of the side chain, and spatial arrangement of the substituents. The introduction of hydrophilic groups into the side chain resulted in lower potency toward UGT2B7 independent of their spatial orientation. Also the elongation of the side chain resulted in lower affinities. However, the introduction of lipophilic alkyl substituents resulted in distinct IC<sub>50</sub> ratios for the stereoisomers and the IC<sub>50</sub> values were influenced by the spatial orientation of those substituents. It was therefore concluded that the differences in IC<sub>50</sub> values stemmed from specific and spatially directed lipophilic interactions toward the binding site of UGT2B7. The spatial arrangement of the substituents, as in series Dd, resulted in high potency towards the enzyme.

### 3D-QSAR study

In the absence of knowledge of the three-dimensional structure of UGTs, ligand-based three-dimensional quantitative structure–activity relationship (3D-QSAR) techniques such as Comparative Molecular Field Analysis (CoMFA)<sup>[12]</sup> or Comparative Molecular Similarity Indices (CoMSIA)<sup>[14]</sup> may be useful in identifying molecular features that improve activity. This technique requires the use of a template conformation upon which each 3D-QSAR model is based. For systems in which no information exists about the binding site, it is a well established assumption to consider that a similar class of molecules binds to the putative binding site adopting a similar orientation and conformation. CoMSIA, as the related CoMFA, compares a series of molecules in terms of molecular interaction fields and correlates field differences with differences in the dependent biological activity. In CoMSIA, Gaussian functions are used to describe steric, electrostatic, hydrophobic and hydrogen bond similarities. Regarding the interpretation of 3D-QSAR results in terms of field contribution maps, CoMSIA denotes those areas within the regions occupied by the ligands that favor or disfavor the presence of a particular group. In addition, it has been shown that CoMSIA fields are more contiguous, facilitating the interpretation compared to the classical CoMFA method.<sup>[15]</sup>

In the present study a CoMSIA model was developed for 76 ligands of human UGT2B7. The global minimum conformation of **30d** (which shows the highest inhibitory activity) was

chosen as the template conformation and all other molecules were aligned to it using the Flex-Align module in MOE.<sup>[16]</sup> In this way, the remaining compounds of the data set were flexibly fitted on **30d** maximizing the steric and electronic overlap (Figure 3). The resulting alignment which is shown in Figure 3a shows that the hydrophobic tricyclic core structure is well



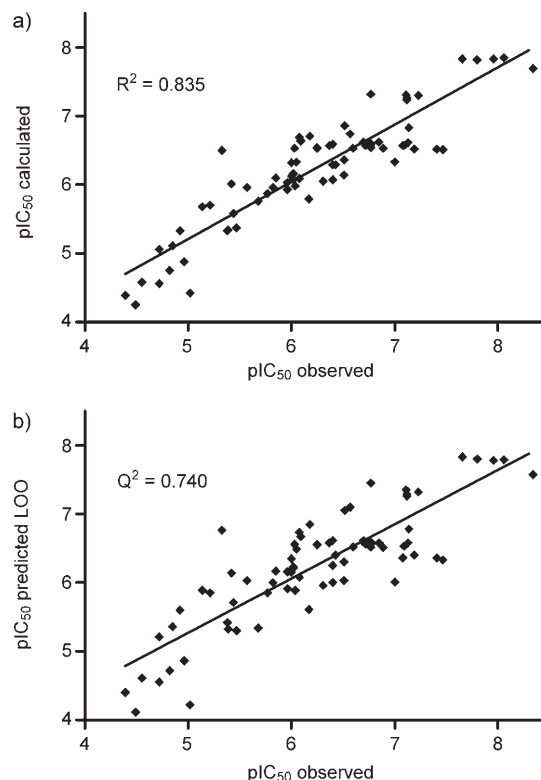
**Figure 3.** a) Stereoview of the ligand-based alignment of all 76 inhibitors as used for the CoMSIA study. b) Alignment of the four stereoisomers **30a–30d**. (**30d** is colored dark gray.)

aligned, whereas the hydrophilic groups are less precisely superimposed. For the most potent inhibitors the alignment shows that the distance between hydrophilic group and tricyclic core is similar. This is exemplarily shown for the four stereoisomers of **30d** in Figure 3b.

The most significant CoMSIA model generated with the SYBYL 7.2 program<sup>[17]</sup> was obtained by combining hydrogen-bond donors and acceptors, and electrostatic and hydrophobic features. The inclusion of the steric field did not increase the quality of the model. As summarized in Table 4, the CoMSIA model resulted in a leave-one-out cross-validated correlation coefficient  $q^2_{\text{LOO}}$  of 0.740 and standard error prediction (SDEP) of 0.490 (3 components), non-cross-validated PLS analysis yielded a correlation coefficient of 0.835 and a standard error of estimate (SDEE) of 0.388. Figure 4a and 4b show the experimental (observed)  $pIC_{50}$  compared to the calculated values,

**Table 4.** Summary of CoMSIA results.

Optimal number of components	3
Q2 (LOO)	0.740
SDEP	0.490
Q2 (L-20%-O)	0.706
Q2 (L-50%-O)	0.630
R2	0.835
SDEE	0.388
F value	107.1



**Figure 4.** a) Calculated versus observed activities ( $pIC_{50}$ ) for the 76 compounds under study obtained from the CoMSIA model. b) Predicted (LOO) versus observed activities ( $pIC_{50}$ ) for the 76 compounds under study obtained from the CoMSIA model.

and the correlation between observed and predicted (leave-one-out) activities. The leave-one-out cross-validation method might lead to high  $q^2$  values which do not necessarily reflect a general predictivity of a model.<sup>[18]</sup> Therefore, further cross-validation, selecting five and two groups of approximately the same size in which the objects were assigned randomly, was performed. In this method, 20% and 50% of the compounds were randomly selected and a model is generated, which is then used to predict the remaining compounds. With a  $q^2_{\text{L20\%O}}$  value of 0.706, we have shown that even after discarding every fifth compound from the training set the predictivity of the models is hardly impaired. Repeating this calculation  $n = 30$  times for each model assesses the liability of the respective model to random effects caused by different assignment of the ligands to the five groups. The leave-50%-out cross-validation showed that even when reducing the number of com-

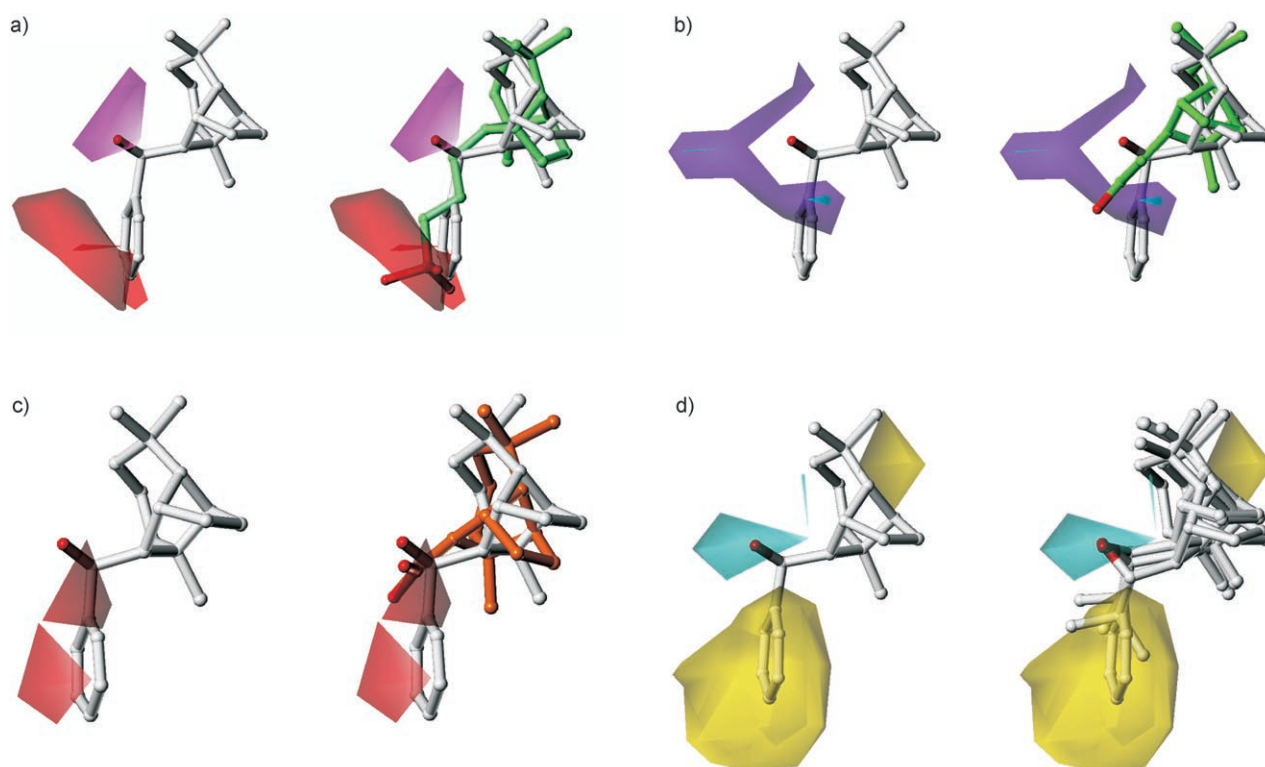
pounds used for model generation, significant CoMSIA models were obtained ( $q^2_{LS0\%O}$ -value 0.630). The leave-20%-out and leave-50%-out cross-validation further supported the robustness of the derived CoMSIA model.

### Graphical interpretation of the CoMSIA model

With the aim of interpreting the 3D-QSAR model, the three-dimensional representation of field contributions (as "stdev\*-coeff" contour plots) for the individual probes were plotted. The graphical representation is useful to detect important features, which contribute to interactions between the ligand and its target protein. The stdev\*-coeff contour plots for the individual fields are shown superimposed on the structure of the most potent inhibitor **30d** in Figure 5a–d. In Figure 5a the magenta contour map around the hydroxy group of **30d** indicates the favorable position of a hydrogen bond acceptor group, whereas the red contour map more distantly located shows that enlarging the distance between the hydrophobic tricyclic core and the polar side chain decrease the activity. As an example the less active compound **12a** with a longer side chain is shown for comparison purposes on the right hand site. For the hydrogen bond donor field (Figure 5b) a large negative contour plot near the hydroxy group indicates that

the position of the hydroxy group is essential and modifying its local position in the molecule decreases the activity. As an example the less active compound **2a** is shown on the right hand site with the hydroxy group located in the unfavorable area. The analysis of the electrostatic map (Figure 5c) shows that more negatively charged groups located in the red contour plot are predicted to decrease activity. This is in accordance with the finding that the carboxylic and phosphonic acids displayed significantly lower potency toward UGT2B7 in comparison to their corresponding homologous alcohols. This result indicated that the charged acidic groups exerted weaker attractive interactions toward the binding site of the enzyme. Cyan and yellow contours in Figure 5d indicate areas where hydrophilic and hydrophobic properties are preferred, respectively. The cyan contour plot around the hydroxy group underlines the importance of the unsubstituted polar hydroxy group in this particular region. Visual inspection of the most active molecules indicates that, in all cases, the aromatic or aliphatic moiety orients at least one carbon atom into the favored hydrophobic region predicted by the CoMSIA model. This is exemplarily shown in Figure 5d on the right for compounds **26d**, **27d**, **29d**, and **30d**.

Field contribution maps for hydrophobic- and hydrogen-bond donor and acceptor properties were generated for



**Figure 5.** Graphical representation (Stdev\*-coeff contour plots) of the CoMSIA (the most active inhibitor **30d** is displayed for comparison). a) Hydrogen bond acceptor plots. The magenta area (80%) represents the region where a hydrogen bond acceptor is predicted to increase activity, whereas the red area (20%) represents the region where hydrogen bond acceptors are predicted to decrease activity. This is highlighted on the right site, where the less active compound **12a** is shown in green. b) Hydrogen bond donor plots. The purple area (20%) represents the region where hydrogen bond donors are predicted to decrease activity. This is highlighted on the right site, where the less active compound **2a** is shown in green. c) Electrostatic contour plots. The red area (80%) represents the region where more positively charged groups are predicted to increase activity (and more negatively charged groups decrease activity). This is highlighted on the right site, where the less active compound **6a** bearing the carboxylic group is shown in orange. d) Hydrophobic contour plots. Yellow contours (80%) encompass regions favorable for hydrophobic groups whereas in cyan (20%) colored areas more hydrophilic groups are favorable for activity. On the right the most active inhibitors **25d**, **26d** and **28d** are displayed together with **30d**.

UGT2B7 ligands that allow a more detailed understanding of their activity in structural terms. The hydrogen bond property fields underline the importance of the correct position of the polar alcohol, carboxylic, or phosphonic acid group in the molecule structure.

### Glucuronidation assays

To investigate the influence of steric hindrance on the glucuronidation rate, the longifolol and isolongifolol derivatives (Tables 1 and 2) were subjected to glucuronidation assays. The sesquiterpenoid alcohols and their corresponding  $\beta$ -D-glucuronides could not be detected either by UV or fluorescence spectroscopy because of the lack of appropriate chromophores. Therefore, the formation of the terpenoid  $\beta$ -D-glucuronides was assayed by the use of radiolabeled [ $^{14}$ C]UDPGlcA after incubating the reaction buffer for 14 h at a high UGT2B7 concentration of 2.0 mg protein mL $^{-1}$ . The detection limit was 8.0 pmol terpenoid  $\beta$ -D-glucuronide at a signal-to-noise ratio of 10. The high concentration of enzyme and the long incubation time were necessary to unambiguously identify and detect the formed terpenoid [ $^{14}$ C] $\beta$ -D-glucuronides. The employed co-substrate concentration of 500  $\mu$ M was sufficient for the purpose of this study. The compounds were employed at saturating concentrations to ensure that depletion of terpenoid derivatives did not occur. The best conversion was 8.5% (**3b**), therefore, no depletion of terpenoid alcohol due to glucuronide formation took place. Hence, the amount of formed  $\beta$ -D-glucuronide was applied as a measure of the glucuronidation rate. The results of this study are displayed in Table 5 and some clear trends were observed.

The amount of formed longifolol (**1b**) glucuronide was 1.21 nmol, whereas isolongifolol (**1a**) was conjugated at a significantly lower rate and merely 109 pmol glucuronide was obtained (eudismic ratio = 11).<sup>[11, 19, 20]</sup> The carboxylic acids **6a** and **6b** displayed a similar behavior because compound **6b** was glucuronidated at a significantly higher rate than its epimer **6a** (eudismic ratio = 15). The eudismic ratios > 10 expressed the high preferential glucuronidation of the *exo*-positioned carboxy and hydroxymethyl groups (compounds **1b** and **6b**) in comparison to the *endo*-COOH and *endo*-CH $_2$ OH groups (compounds **1a** and **6a**). However, the eudismic ratios decreased with elongation of the side chain indicating that the higher flexibility of the side chain leveled the effect of the distinct spatial *endo*- and *exo*-arrangements. Consistently, there was no chiral distinction for the stereoisomers bearing  $\geq 3$  carbon atoms in the side chain (compounds **3a**, **3b**, **7a**, **7b**, **9a**, and **9b**) and it was concluded that UGT2B7 did not recognize the distant stereogenic center at C8 in these homologous compounds. Small differences in the glucuronidation rates were observed for the geometric isomers (*E*-**4a** and (*Z*)-**4a** as well as (*E*-**8a** and (*Z*-**8a**, which displayed eudismic ratios of 2.5 and 2.7, respectively, and the (*Z*)-isomers were better substrates than the (*E*)-isomers. Interestingly, the elongation of the side chain resulted in higher glucuronidation rates. For example, the amount of formed glucuronide increased in the order **1a** < **2a** < **5a** < **3a** for the homologous primary alcohols of ser-

**Table 5.** Amount of formed  $\beta$ -D-glucuronide.

Compd	Glucuronide [pmol] <sup>[a]</sup>	Compd	Glucuronide [pmol] <sup>[a]</sup>	Eudismic ratio <sup>[b]</sup>
Series A <sup>[c]</sup>				
<b>1a</b>	109	<b>1b</b>	1211	11
<b>2a</b>	387	<b>2b</b>	1921	5.0
<b>3a</b>	1843	<b>3b</b>	2116	1.1
( <i>E</i> )- <b>4a</b>	781	( <i>Z</i> )- <b>4a</b>	1959	2.5
<b>5a</b>	1186	<b>5b</b>	1423	1.2
Series B <sup>[d]</sup>				
<b>6a</b>	81	<b>6b</b>	1244	15
<b>7a</b>	1321	<b>7b</b>	1854	1.4
( <i>E</i> )- <b>8a</b>	877	( <i>Z</i> )- <b>8a</b>	2320	2.7
<b>9a</b>	1370	<b>9b</b>	1561	1.1
<b>10a</b>	N.D.	( <i>E</i> )- <b>11a</b>	N.D.	
<b>12a</b>	N.D.	( <i>E</i> )- <b>13a</b>	N.D.	
<b>14a</b>	N.D.	<b>14b</b>	N.D.	
Series C				
<b>15a</b> <sup>[d]</sup>	1105	<b>15b</b> <sup>[d]</sup>	1396	1.3
<b>15c</b> <sup>[d]</sup>	1577	<b>15d</b> <sup>[d]</sup>	1956	1.2
<b>16a</b>	N.D.	<b>16b</b>	N.D.	
<b>16c</b>	N.D.	<b>16d</b>	N.D.	
<b>17a</b>	N.D.	<b>17b</b>	N.D.	
<b>18a</b>	N.D.	<b>18b</b>	N.D.	
<b>19a</b>	N.D.	<b>19b</b>	N.D.	
<b>20a</b>	N.D.	<b>20b</b>	N.D.	
<b>21a</b>	N.D.	<b>21b</b>	N.D.	
<b>22a</b> <sup>[c]</sup>	1761	<b>22b</b> <sup>[c]</sup>	1453	1.2
<b>22c</b> <sup>[c]</sup>	1788	<b>22d</b> <sup>[c]</sup>	1890	1.1
Series D <sup>[e]</sup>				
<b>23a</b>	N.D.	<b>23b</b>	N.D.	
<b>24a</b>	26	<b>24b</b>	17	1.5
<b>24c</b>	12	<b>24d</b>	21	1.8
<b>25a</b>	15	<b>25b</b>	< 8	
<b>25c</b>	< 8	<b>25d</b>	< 8	
<b>26a</b>	17	<b>26b</b>	< 8	
<b>26c</b>	< 8	<b>26d</b>	< 8	
<b>27a</b>	N.D.	<b>27b</b>	N.D.	
<b>27c</b>	N.D.	<b>27d</b>	N.D.	
<b>28a</b>	N.D.	<b>28b</b>	N.D.	
<b>28c</b>	N.D.	<b>28d</b>	N.D.	
<b>29a</b>	N.D.	<b>29b</b>	N.D.	
<b>29c</b>	N.D.	<b>29d</b>	N.D.	
<b>30a</b>	N.D.	<b>30b</b>	N.D.	
<b>30c</b>	N.D.	<b>30d</b>	N.D.	

[a] Amount of formed glucuronide in pmol; N.D.: not detected; < 8: traces of glucuronide were detected, but quantification was not possible (below signal-to-noise ratio of 10). [b] The higher value in the numerator. [c] The compounds were employed at a saturating concentration of 250  $\mu$ M. [d] The compounds were employed at a saturating concentration of 1.0 mM. [e] The compounds were employed at a saturating concentration of 100  $\mu$ M.

ies A, presumably because the more distant nucleophilic carboxy and hydroxy groups experienced lower steric shielding by the tricyclic hydrocarbon scaffold. The phosphonic acids **10a**–(*E*)-**13a** and the tertiary amines **14a** and **14b** were not glucuronidated at measurable rates.

The introduction of a second substituent into the side chain at C1' decreased the glucuronidation rates significantly (ser-

ies D). The glucuronidation rates of the methyl derivatives **24a** and **24b** were decreased by 76 and 85 % in comparison to isolongifolol (**1b**), respectively. In particular, the conjugation rate of the longifolol derivatives was further decreased by the introduction of a methyl substituent. In this respect, the glucuronidation rate of **24c** and **24d** decreased by >99 % in comparison to longifolol. The results showed clearly that the glucuronidation rate was highly affected by the steric demand of the substituent in vicinity of the hydroxy group. In this respect, the glucuronidation rate decreased in the order Me > Et ≈ vinyl > allyl (Figure 6). In the cases of the allyl-, isopropyl-, *tert*-butyl, and phenyl-substituted derivatives, no glucuronide formation was observed. Consistently, also the bifunctional hydroxy ester

glucuronidation could be prevented by increasing the steric bulk resulting from a decreased accessibility to the nucleophilic hydroxy group.

## Conclusion

Detailed structure–activity relationships with respect to chemical functionality, stereochemistry, and steric properties were derived. The potency assays showed that lipophilic interactions between the small guest molecule and the binding site of UGT2B7 resulted in higher potency levels. The spatial arrangement of lipophilic substituents had a significant effect on the IC<sub>50</sub> values and the phenyl-substituted longifolol derivative **30d** was the best inhibitor (IC<sub>50</sub> = 4.6 nM). The rate of the UGT2B7-catalyzed glucuronidation reaction decreased with increasing steric demand of the substituent in proximity of the nucleophilic hydroxy group. Therefore, by introducing bulky substituents such as isopropyl, *tert*-butyl, and phenyl groups, the glucuronidation reaction was prevented due to the decreased accessibility to the hydroxy group. These findings reveal that high potency substrates such as isolongifolol (**1a**) and longifolol (**1b**) can be turned into potent true inhibitors by addressing functional, stereochemical, and steric properties.

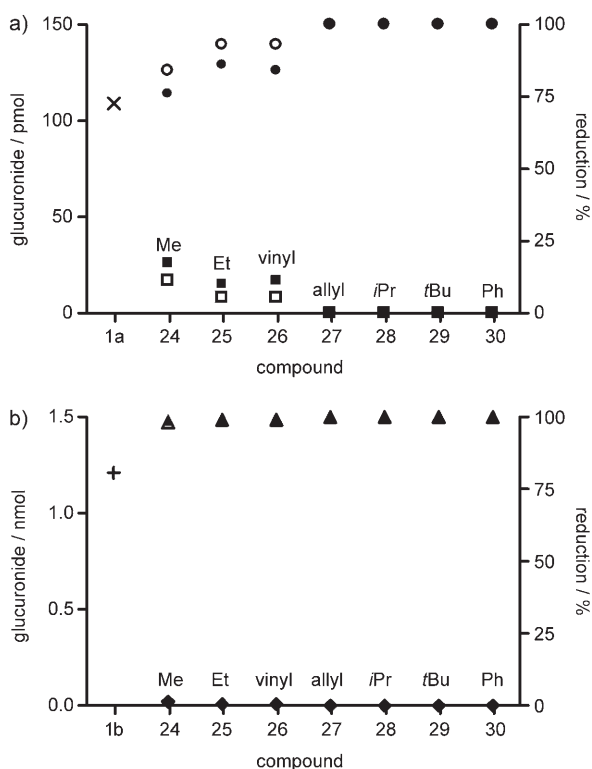
## Experimental Section

**Materials.** UDPGlcA (trisodium salt, CAS 63700-19-6), saccharic acid-1,4-lactone (CAS 61278-30-6), and estriol (CAS 50-27-1) were obtained from Sigma (St. Louis, MO, USA). Radiolabeled [<sup>14</sup>C]UDPGlcA was acquired from PerkinElmer Life and Analytical Sciences (Boston, MA, USA). HPLC grade solvents were used throughout the study. The recombinant human UGT2B7 was expressed as a His-tagged protein in baculovirus-infected insect cells as previously described.<sup>[21]</sup> The syntheses and characterizations of the sesquiterpenoid derivatives has been published previously.<sup>[22]</sup>

**IC<sub>50</sub> values.** The IC<sub>50</sub> value of each longifolol and isolongifolol derivative was determined by the use of the reference substrate estriol at a concentration of 25 μM (*K<sub>m</sub>* ≈ 5 μM, [*S*]/*K<sub>m</sub>* ≈ 5). The IC<sub>50</sub> values were measured at five concentrations bracketing the apparent IC<sub>50</sub> (0.10, 0.25, 1.0, 4.0, and 10 × IC<sub>50</sub>). The reaction mixture consisted of phosphate buffer (50 mM, pH 7.4), MgCl<sub>2</sub> (10 mM), and saccharic acid-1,4-lactone (5.0 mM). The enzyme UGT2B7 was employed at 0.10 mg protein mL<sup>-1</sup>. Assays in the absence of inhibitor, blank runs in the absence of co-substrate, and control assays using either isolongifolol (**1a**) or longifolol (**1b**) at a fixed concentration of 0.10 μM were included. The enzyme reaction was initiated after a 5 min pre-incubation at 37 °C by the addition of a solution of UDPGlcA to a final concentration of 5.0 mM. The reactions were terminated after an incubation time of 15 min at 37 °C by the addition of ice-cold perchloric acid (4.0 M) and transfer to ice. The mixtures were centrifuged (16 000 g, 10 min) and aliquots of the supernatants were subjected to HPLC analysis. The data were analyzed by nonlinear regression applying Equation (1)

$$\% \text{inhibition} = \frac{100}{1 + \left(\frac{[I]}{[I]_{50}}\right)^h} \quad (1)$$

wherein [*I*] = inhibitor concentration and *h* = Hill coefficient. The results reflect a minimum of three replicate determinations.



**Figure 6.** Amount of formed glucuronide (left ordinate) and percent reduction in glucuronide formation (right ordinate) for the stereoisomers of series D in comparison to isolongifolol (**1a**, ×) and longifolol (**1b**, +). a) Amount of formed glucuronide and percent reduction for the compounds of series Aa (■, ●) and Ab (□, ○). b) Amount of formed glucuronide and percent reduction for the compounds of series Ac (◆, ▲) and Ad (◇, △).

and hydroxy phosphonic acid derivatives **16a–21b** of series C were not glucuronidated at detectable rates. To substantiate these findings, the stereoisomeric phenyl-substituted derivatives **30a–d** were subjected to glucuronidation assays employing higher concentrations of both radiolabeled co-substrate and enzyme. The results showed that these compounds were not glucuronidated at detectable rates and it was concluded that they were true inhibitors of UGT2B7. These results indicated that the UGT2B7-catalyzed glucuronidation reaction was highly sensitive to the steric demand of the substituents in proximity of the nucleophilic hydroxy group. The substrate

**Molecular modeling.** All molecular modeling and comparative molecular field evaluations were performed using SYBYL version 7.2 (Tripos Inc., St. Louis, USA)<sup>[17]</sup> and MOE2006.08 (Chemical Computing Group, Montreal, Canada)<sup>[16]</sup> running on a Linux Pentium PC. Atomic point charges were calculated using the MMFF94 method. The structures of the molecules were built with Builder module within MOE and energy minimized by using the MMFF94 force field with a distance-dependent dielectric term. The low-energy conformations of the most active compound **30d** were analyzed by means of a simulated annealing procedure. This was accomplished by heating to 700 K for 1 ps and slowly annealing to 200 K in steps of 100 K for 1 ps at each temperature, with a step size of 1 fs; and snapshots were captured every 5 fs. The global minimum structure from the MD trajectory was sent through a final round of minimization. The most crucial input for a 3D-QSAR is the alignment of the molecules. Molecule **30d** with the highest inhibitory activity was chosen as the template and all other molecules were aligned to it using the Flex-Align method in MOE. (Figure 3a). In this way, the remaining compounds of the data set were flexibly fitted on **30d** maximizing the steric and electronic overlap.

**3D-QSAR.** The 3D-QSAR study was carried out using the CoMSIA module in SYBYL7.2<sup>[17]</sup> and default settings. Atomic point charges were calculated using the MMFF94 method in MOE<sup>[16]</sup> (as used for the building of the inhibitors). All ligands were treated in the protonation state at pH 7.4. The biological data were transformed into  $pIC_{50}$  ( $-\log IC_{50}$ ) values. The ligand alignment was generated as described above. The five physicochemical properties for CoMSIA<sup>[14]</sup> (steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor) were evaluated using the standard probe atom with 1 Å radius, +1.0 charge, and hydrophobicity and hydrogen bond property values of +1. The value of the so-called attenuation factor  $R$  was 0.3. The size of the grid box considered for the calculation was defined in such a way that it extended approximately 4 Å beyond each of the molecules in each dimension (automated generation). To form the basis for a statistical significant model, the method of partial least squares (PLS) regression was used to analyze the 76 compounds by correlating variations in their biological activities with variations in their interaction fields. The optimum number of PLS components corresponding to the smallest standard error of prediction, was determined by the leave-one-out (LOO) cross-validation procedure. Using the optimal number of components, the final PLS analysis was carried out without cross-validation to generate a predictive model with a conventional correlation coefficient. A column filtering of  $2.0 \text{ kcal mol}^{-1}$  was applied to reduce the noise and to speed up the calculation. The LOO cross-validation method might lead to high  $q^2$  values which do not necessarily reflect a general predictivity of a model.<sup>[18]</sup> Therefore, further cross-validation, selecting five and two groups of approximately the same size in which the objects were assigned randomly, was performed. In this method, 20% and 50% of the compounds were randomly selected and a model is generated, which is then used to predict the remaining compounds (leave-20%-out, leave-50%-out). This cross-validation technique has been shown to yield better indices for the robustness of a model than the normal LOO procedure.<sup>[18]</sup> By repeating this procedure 50 times, 50 different test-sets are simulated by this procedure. A common test to check the consistency of the models is to scramble the biological data and repeat the model derivation process, thus allowing detection of possible chance correlations. After randomizing our data set, solely values of  $q^2 < 0.3$  were observed in the PLS analyses. To represent favorable and unfavorable areas for hydrogen bond acceptor and donor, electrostatic and hydrophobic contributions, 80%

and 20% stdev \* coeff contour plots were built and analyzed visually in SYBYL.

**Glucuronidation assays.** The formation of glucuronide was measured using radiolabeled [ $^{14}C$ ]UDPGlcA.<sup>[23]</sup> A solution of [ $^{14}C$ ]UDPGlcA (50  $\mu\text{L}$ ,  $196 \text{ mCi mmol}^{-1}$ ,  $0.02 \text{ mCi mL}^{-1}$  in EtOH/water 7:1 v/v) was transferred to 1 mL vials and the solvent was evaporated in vacuo. The residue was dissolved in reaction buffer (90  $\mu\text{L}$ ), which consisted of UGT2B7 (2.0 mg protein  $\text{mL}^{-1}$ ), cold UDPGlcA (450  $\mu\text{M}$ ),  $\text{MgCl}_2$  (10 mM), phosphate buffer (50 mM, pH 7.4) and saccharic acid-1,4-lactone (5.0 mM). The reaction was initiated by the addition of the sesquiterpenoid derivative to a final saturating concentration (cf. Table 4). After incubating for 14 h at  $37^\circ\text{C}$ , the reactions were terminated by the addition of perchloric acid (10  $\mu\text{L}$ , 4.0 M) and transfer to ice. The mixtures were centrifuged (16000 g, 10 min) and aliquots of the supernatants were subjected to HPLC analysis. Control assays in the presence of estriol (25  $\mu\text{M}$ ) and blank runs were included. The detection limit was determined by subjecting dilutions of reaction buffer containing longifolol [ $^{14}C$ ] $\beta$ -D-glucuronide to HPLC analysis. The results reflect a minimum of three replicate determinations.

**Glucuronidation assays for 30a–30d.** The formation of glucuronide was assayed by the use of radiolabeled [ $^{14}C$ ]UDPGlcA. A solution of [ $^{14}C$ ]UDPGlcA (400  $\mu\text{L}$ ,  $196 \text{ mCi mmol}^{-1}$ ,  $0.02 \text{ mCi mL}^{-1}$  in EtOH/water 7:1 v/v) was transferred to 2 mL vials and the solvent was evaporated in vacuo at room temperature. The residue was dissolved in reaction buffer (90  $\mu\text{L}$ ), which consisted of either UGT2B7 (2.0 mg protein  $\text{mL}^{-1}$ ) or human liver microsomes (100  $\mu\text{g}$ ),  $\text{MgCl}_2$  (10 mM), phosphate buffer (50 mM, pH 7.4) and saccharic acid-1,4-lactone (5.0 mM). The sesquiterpenoid derivative (10  $\mu\text{L}$  of a 1.0 mM solution in DMSO/water 50:50 v/v) was added to the reaction buffer to start the reaction. After incubating for 14 h at  $37^\circ\text{C}$ , the reaction mixture was centrifuged (16000 g, 10 min) and aliquots of the supernatants were subjected to HPLC analysis. Control assays in the presence of the substrate longifolol (100  $\mu\text{M}$ ) and blank runs were included. The detection limit (30 pmol, signal-to-noise ratio = 10:1) was determined by subjecting dilutions of reaction buffer containing longifolol [ $^{14}C$ ] $\beta$ -D-glucuronide to HPLC analysis. The results reflect a minimum of two replicate determinations.

**HPLC methods.** The HPLC system consisted of the Agilent 1100 series degasser, binary pump, autosampler, thermostated column compartment, multiple wavelength detector, and fluorescence detector (Agilent Technologies, Palo Alto, CA, USA). The resulting spectra were analyzed with Agilent ChemStation software (Rev B.01.01). Glucuronidation reaction products were separated and detected as follows. Estriol  $\beta$ -D-glucuronide: Hypersil BDS-C18 (150  $\times$  4.6 mm; Agilent Technologies, Palo Alto, CA, USA); 35% MeOH in phosphate buffer (50 mM, pH 3.0); flow rate  $1.0 \text{ mL min}^{-1}$ ; detection by fluorescence spectroscopy (excitation at  $\lambda = 335 \text{ nm}$ , emission at  $\lambda = 455 \text{ nm}$ ); retention time 4.2 min. Terpenoid [ $^{14}C$ ] $\beta$ -D-glucuronides: Chromolith SpeedROD (50  $\times$  4.6 mm; Merck, Darmstadt, Germany); gradient run 5% MeOH in phosphate buffer (50 mM, pH 3.0; 0.0 to 3.5 min), 5%–80% MeOH (3.5 to 8.0 min), 80% MeOH (8.0 to 13 min), 80%–5% MeOH (13 to 15 min), 5% MeOH (15 to 20 min); flow rate  $1.0 \text{ mL min}^{-1}$ ; retention times 8.1 to 10.2 min; detection by use of a 9701 HPLC Radioactivity Monitor (Reeve Analytical, Glasgow, Scotland).

## Acknowledgements

This work was supported by a fellowship from the Finnish Cultural Foundation (IB), a grant from Walter och Lisi Wahls Stiftelse

för Naturvetenskaplig Forskning, and the Academy of Finland (projects 207535 and 210933).

**Keywords:** 3D-QSAR • eudismic analysis • inhibition • metabolism • UGT

- [1] For reviews, see: a) A. Radomska-Pandya, M. Ouzzine, S. Fournel-Gigleux, J. Magdalou, *Methods Enzymol.* **2005**, *400*, 116–147; b) P. G. Wells, P. I. Mackenzie, J. R. Chowdhury, C. Guillemette, P. A. Gregory, Y. Ishii, A. J. Hansen, F. K. Kessler, P. M. Kim, N. R. Chowdhury, J. K. Ritter, *Drug Metab. Dispos.* **2004**, *32*, 281–290; c) M. B. Fisher, M. F. Paine, T. J. Strelevitz, S. A. Wrighton, *Drug Metab. Rev.* **2001**, *33*, 273–297; d) C. D. King, G. R. Rios, M. D. Green, T. R. Tephly, *Curr. Drug Metab.* **2000**, *1*, 143–161.
- [2] The term catalytic promiscuity has been originally used to describe the ability of an enzyme to catalyze an adventitious secondary activity at the active site responsible for the primary activity (see S. D. Copley, *Curr. Opin. Chem. Biol.* **2003**, *7*, 265–272). However, in the case of metabolic enzymes like UGTs the term has been used to depict their overlapping substrate selectivities.
- [3] a) I. S. Owens, N. K. Basu, R. Banerjee, *Methods Enzymol.* **2005**, *400*, 1–22; b) J. A. Williams, R. Hyland, B. C. Jones, D. A. Smith, S. Hurst, T. C. Goosen, V. Peterkin, J. Koup, S. E. Ball, *Drug Metab. Dispos.* **2004**, *32*, 1201–1208; c) P. I. Mackenzie, K. W. Bock, B. Burchell, C. Guillemette, S.-I. Ikushiro, T. Iyanagi, J. O. Miners, I. S. Owens, D. W. Nebert, *Pharmacogenet. and Genomics* **2005**, *15*, 677–685.
- [4] A. Rowland, P. Gaganis, D. J. Elliot, P. I. Mackenzie, K. M. Knights, J. O. Miners, *J. Pharmacol. Exp. Ther.* **2007**, *321*, 137–147.
- [5] K. Grancharov, Z. Naydenova, S. Lozeva, E. Golovinsky, *Pharmacol. Ther.* **2001**, *89*, 171–186.
- [6] a) C. M. Timmers, M. Dekker, R. C. Buijsman, G. A. van der Marel, B. Ethell, G. Anderson, B. Burchell, G. J. Mulder, J. H. van Boom, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1501–1506; b) D. Noort, E. A. Meijer, T. J. Visser, J. H. Meerman, G. A. van der Marel, J. H. van Boom, G. J. Mulder, *Mol. Pharmacol.* **1991**, *40*, 316–320.
- [7] V. L. Schramm, *Curr. Opin. Struct. Biol.* **2005**, *15*, 604–613.
- [8] V. L. Schramm, *Annu. Rev. Biochem.* **1998**, *67*, 693–720.
- [9] a) M. Said, D. Noort, J. Magdalou, J. C. Ziegler, G. A. van der Marel, J. H. van Boom, G. J. Mulder, G. Siest, *Biochem. Biophys. Res. Commun.* **1992**, *187*, 140–145; b) D. Noort, M. W. Coughtrie, B. Burchell, G. A. van der Marel, J. H. van Boom, A. van der Gen, G. J. Mulder, *Eur. J. Biochem.* **1990**, *188*, 309–312; c) D. K. Alargov, R. G. Gugova, P. S. Denkova, G. Müller, E. V. Golovinsky, *Monatsh. Chem.* **1999**, *130*, 937–943; d) D. K. Alargov, Z. Naydenova, K. Grancharov, P. S. Denkova, E. V. Golovinsky, *Monatsh. Chem.* **1997**, *128*, 725–732; e) D. K. Alargov, Z. Naydenova, K. Grancharov, P. Denkova, E. Golovinsky, *Monatsh. Chem.* **1998**, *129*, 755–760.
- [10] a) I. Bichlmaier, A. Siiskonen, M. Finel, J. Yli-Kauhaluoma, *J. Med. Chem.* **2006**, *49*, 1818–1827; b) I. Bichlmaier, A. Siiskonen, M. Kurkela, M. Finel, J. Yli-Kauhaluoma, *Biol. Chem.* **2006**, *387*, 407–416; c) I. Bichlmaier, *Drug Metab. Rev.* **2006**, *38* (Suppl. 2), 185–186.
- [11] I. Bichlmaier, M. Kurkela, A. Siiskonen, M. Finel, J. Yli-Kauhaluoma, *Bioorg. Chem.* **2007**, *35*, 386–400.
- [12] R. D. Cramer, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- [13] a) M. J. Sorich, R. A. McKinnon, J. O. Miners, D. A. Winkler, P. A. Smith, *J. Med. Chem.* **2004**, *47*, 5311–5317; b) M. J. Sorich, J. O. Miners, R. A. McKinnon, D. A. Winkler, F. R. Burden, P. A. Smith, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2019–2024; c) M. J. Sorich, P. A. Smith, D. A. Winkler, F. R. Burden, R. A. McKinnon, J. O. Miners, *Drug Metab. Rev.* **2003**, *35* (Suppl. 2), 333.
- [14] G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem.* **1994**, *37*, 4130–4146.
- [15] M. Bohm, J. Stürzebecher, G. Klebe, *J. Med. Chem.* **1999**, *42*, 458–477.
- [16] *MOE Molecular Modeling Software*, version 2006.8; Chemical Computing Group, Montreal, Canada.
- [17] *SYBYL Molecular Modeling Software*, version 7.2; Tripos Inc., St. Louis, USA.
- [18] W. Sippl, *J. Comp.-Aided Mol. Des.* **2000**, *14*, 559–572.
- [19] Eudismic analysis investigates the interaction between chiral compounds and their biological targets. This analytical method is used in drug design and development to increase the selectivity and potency of a chiral lead compound toward the pharmacological target. The term eudismic analysis was derived from the expressions eutomer and distomer. The more biologically active stereoisomer is termed eutomer, the less active one is called distomer. See: a) R. Crossley, *Chirality and the biological activity of drugs*, CRC, Boca Raton, **1995**, p. 196; b) M. Eichelbaum, B. Testa, A. Somogyi, *Stereochemical aspects of drug action and disposition*, Springer-Verlag, Berlin, **2003**, p. 442.
- [20] The eudismic ratio was calculated by dividing the amount of formed eutomer glucuronide by the amount of formed distomer glucuronide. The eudismic ratio therefore expressed the preferential formation of the glucuronide of one compound over that of its respective stereoisomer.
- [21] M. Kurkela, A. Garcia-Horsman, L. Luukkanen, S. Mörsky, J. Taskinen, M. Baumann, R. Kostainen, J. Hirvonen, M. Finel, *J. Biol. Chem.* **2003**, *278*, 3536–3544.
- [22] a) I. Bichlmaier, M. Kurkela, T. Joshi, A. Siiskonen, T. Rüffer, H. Lang, B. Suchanová, M. Vahermo, M. Finel, J. Yli-Kauhaluoma, *J. Med. Chem.* **2007**, *50*, 2655–2664; b) I. Bichlmaier, M. Kurkela, T. Joshi, M. Finel, J. Yli-Kauhaluoma, *ChemMedChem* **2007**, *2*, 881–889.
- [23] S. Kaivosari, J. S. Salonen, J. Mortensen, J. Taskinen, *Anal. Biochem.* **2001**, *292*, 178–187.

Received: May 22, 2007

Revised: September 5, 2007

Published online on October 17, 2007