



# Molecular recognition of CYP26A1 binding pockets and structure–activity relationship studies for design of potent and selective retinoic acid metabolism blocking agents

Bin Sun, Shuai Song, Chen-Zhou Hao, Wan-Xu Huang, Chun-Chi Liu, Hong-Lei Xie, Bin Lin, Mao-Sheng Cheng, Dong-Mei Zhao\*

Key Laboratory of Structure-Based Drug Design & Discovery of Ministry of Education, School of Pharmaceutical Engineering, Shenyang Pharmaceutical University, 103 Wenhua Road, Shenhe District, Shenyang 110016, PR China

## ARTICLE INFO

### Article history:

Accepted 30 November 2014

Available online 8 December 2014

### Keywords:

All-trans-retinoic acid (ATRA)

CYP26A1

Retinoic acid metabolism blocking agents

(RAMBAs)

Homology modeling

Pharmacophore model

## ABSTRACT

All-trans-retinoic acid (ATRA), the biologically most active metabolite of vitamin A, plays a major role in the regulation of cellular differentiation and proliferation, and it is also an important pharmacological agent particularly used in the treatment of cancer, skin, neurodegenerative and autoimmune diseases. However, ATRA is very easy to be metabolized into 4-hydroxyl-RA in vivo by CYP26A1, an inducible cytochrome P450 enzyme, eventually into more polar metabolites. Therefore, it is vital to develop specific retinoic acid metabolism blocking agents (RAMBAs) to inhibit the metabolic enzyme CYP26A1 in the treatment of relevant diseases aforementioned.

In this study, CYP26A1 and its interactions with retinoic acid-competitive metabolism blocking agents were investigated by a combined ligand- and structure-based approach. First, since the crystal structure of CYP26A1 protein has not been determined, we constructed the 3D structure of CYP26A1 using homology modeling. In order to achieve a deeper insight into the mode of action of RAMBAs in the active site, the molecular superimposition model and the common feature pharmacophore model were constructed, and molecular docking was performed. The molecular superimposition model is composed of three features: the main chain groups, side chain groups, and azole groups. The common feature pharmacophore model consists of five chemical features: four hydrophobic groups and one hydrogen acceptor (HHHHA). The results of molecular docking show that the characteristic groups of RAMBAs were mapped into three different active pockets, respectively.

A structure–activity relationship (SAR) was obtained by a combination of the molecular superimposition and docking results with the pharmacophore model. This study gives more insight into the interaction model inside the CYP26A1 active site and provides guidance for the design of more potent and possibly more selective RAMBAs.

© 2014 Published by Elsevier Inc.

## 1. Introduction

All-trans-retinoic acid (ATRA), a natural metabolite of vitamin A, present in the multitude of human tissues, plays a crucial role in the regulation of cellular differentiation, proliferation and gene expression [1–3]. As a key signaling molecule, it has been shown that ATRA deficiency is associated with diseases such as acne, psoriasis, ichthyosis and cancer, especially in oncology against acute promyelocytic leukemia (APL), ATRA can change the prognosis of APL from a fatal leukemia to a highly curable disease [4–7]. Unfortunately,

although ATRA is useful in the treatment of cancer and skin-related diseases, its clinical applications have been significantly hampered by the emergence of resistance, and the fact that the ATRA is very easy to be metabolized into 4-hydroxyl-RA by P450 enzymes [8,9].

There are several microsomal cytochrome P450 (CYP) enzymes that are suggested to be involved in retinoic acid metabolism, e.g. CYP1A1, CYP4A11, CYP3A4/5/7 and CYP2C8/9 [10,11]. However, much attention has been paid to CYP26, a new family of cytochrome P450 enzymes, which is specifically responsible for retinoic acid metabolism. The CYP26 family consists of three members: CYP26A1, CYP26B1 and CYP26C1 [12–14]. Among them, CYP26A1 is the most extensively studied member, which metabolizes retinoic acids (RA) mainly into inactive derivatives or polar metabolites such as 4-OH-RA, 4-oxo-RA, 5, 8-epoxy-RA and

\* Corresponding author. Tel.: +86 24 2398 6413; fax: +86 24 2399 5043.

E-mail address: [medchemzhao@163.com](mailto:medchemzhao@163.com) (D.-M. Zhao).

**Table 1**  
Detailed BLAST results.

Entry name	Organism	SEcore	Identity	Length	PDB code
CP120_SYNY3	<i>Synechocystis</i> sp.	641	33.0%	444	2VE3
CP51A_HUMAN	<i>Homo sapiens</i>	290	23.0%	503	3JUS
EIZFM_STRCO	<i>Streptomyces coelicolor</i>	276	24.0%	461	3DBG
CP51A_HUMAN	<i>Homo sapiens</i>	257	24.0%	404	3JUS
CP51_MYCTU	<i>Mycobacterium tuberculosis</i>	241	25.0%	451	1E9X
CP46A_HUMAN	<i>Homo sapiens</i>	235	23.0%	500	2Q9F
CP3A4_HUMAN	<i>Homo sapiens</i>	235	24.0%	503	1TQN
CPXB_BACME	<i>Bacillus megaterium</i>	237	26.0%	1049	1BU7

18-OH-RA [15,16]. Moreover, it was found that CYP26A1 is over-expressed in cells containing exogenous retinoic acid, which is an important negative feedback loop controlling RA concentrations and limiting biological action within cells, and which also has the potential to reduce the therapeutic efficacy of ATRA in a wide variety of cancer cell lines [17]. Therefore, it is a reasonable strategy to develop specific retinoic acid metabolism blocking agents (RAM-BAs) targeting CYP26A1.

To date, several chemical families of RAMBAs targeting CYP26A1 have been described [18–23]. The antimycotic substance such as ketoconazole, itraconazole, miconazole and fluconazole were the first generation compounds [18–20], which were evaluated as the potent blocking agents to inhibit CYP26A1 activity. Unfortunately, there are adverse side effects for those compounds that have been attributed to the lack of CYP26 isoform specificity. Subsequent extensive structure–activity relationship (SAR) studies on imidazole derivatives have led to the discovery of CYP26A1 inhibitors with new scaffolds, such as liarozole, triazole derivatives, 2,2-dimethyl-3-(4-(naphthalen-2-ylamino)phenyl)propyl azole derivatives [21–23]. New RAMBAs with higher potency and better specificity against CYP26A1 have been reported, and some of them have appeared in clinical studies and have shown some encouraging preclinical and clinical results – improving the specificity and activity [24–26]. However, although great efforts have been made in developing CYP26A1 inhibitors, there is still lack of satisfactory potency, selectivity for CYP26A1. Therefore, it is essential to understand the structure–activity relationship of existing RAMBAs in the context of CYP26A1 before new RAMBAs could be discovered.

In this study, we constructed a homology model of CYP26A1 and systematically analyzed its spatial structure and topological features of the binding pockets in the catalytic cleft. Based on those analyses, a molecular superimposition model and a common feature pharmacophore model were constructed, and molecular docking was performed. Finally, the results of molecular superimposition and docking results with the pharmacophore model allowed us to deduce the SAR model of RAMBAs. The study provides a useful strategy for designing novel and possibly specific inhibitors against CYP26A1.

## 2. Homology modeling of CYP26A1 and analysis of retinoic acid binding site

Since at present time the crystallographic structure of CYP26A1 from human is not available, we had to generate a homology model which can help us explain how the structurally diverse ligands bind to a common receptor site. The primary sequence of human CYP26A1 was retrieved from UniProtKB/Swiss-Prot [27] with Accession Number O43174. This was subject to BLAST-P [28] using PDB database [29] utilizing BLAST Server of NCBI in order to search the homologous templates. Eight hits with comparable e-value and bit score were obtained (Table 1), and the PDB codes of the three-dimensional structures of all these isoforms have been determined. An analysis of these structures showed that these enzymes

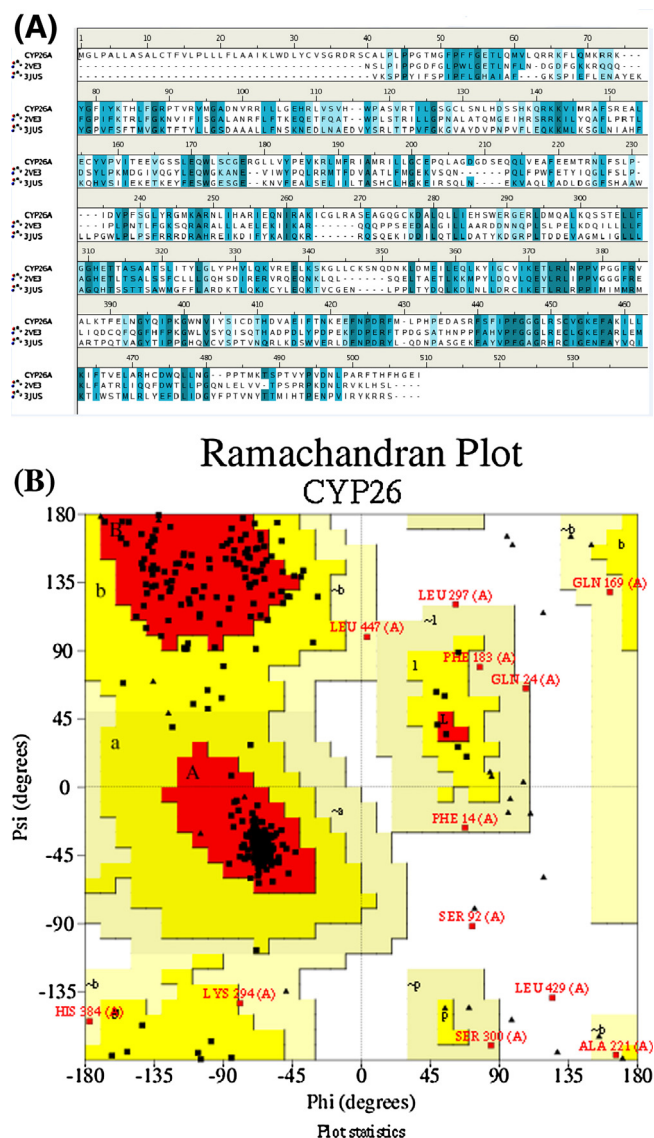
have the sequence alignment homology, and that they present similar types of structure. We chose two sequences with the highest scoring values: CYP120 (PDB code 2VE3) [30a] and CYP51A (PDB code 3JUS) [30b] as templates. Homology modeling of the CYP26A1 three dimensional structures was carried out using the program MODELLER [31]. Because the co-crystallized ligand retinoic acid of CYP120 was located in the active site, which was the first to be reported, the heme group and retinoic acid of CYP120 were inserted and positioned in the same coordinates of CYP26A1. Then the homology model so produced was evaluated using the structure evaluation program (<http://nihserver.mbi.ucla.edu>) through the analysis of bond lengths, bond angles, dihedral angles and other stereochemical characters (such as mainchain phi/psi conformational angle, side chain torsion angle, etc.) [32]. The resulting model was visualized and further assessed by Ramachandran plot (Fig. 1B). 99.2% residues are in the favorable region and allowed region.

The sequence alignment of CYP26A1 from CYP120 and CYP51A indicates the presence of three sequentially conserved regions. These regions corresponding to 422EFNPDRF428, 440FSFIPFGG448 and 452CVGKEFA458 from CYP26A1 contain highly conserved amino acid residues around the active site (Fig. 1A), suggesting that these metabolic enzymes may have similar binding mode with substrates.

The homology model of CYP26A1 with RA provides a clear and deep insight into the binding mode of CYP26A1 (Fig. 2), and reveals that the active site has three binding pockets which were denoted A, B, C. Pocket A is a region for substrate recognition and binding, which consists of six highly hydrophobic amino acid residues (Trp70, Phe183, Phe260, Val331, Phe335, Val331 and Pro439) and two basic amino acid residues (Arg51, Lys436). Since RA is also a very hydrophobic molecule, the conjugated branched-chain group of RA interacts almost entirely through van der Waals contacts with amino acid residues in Pocket A. The only exception is the ionic interaction formed between the carboxyl group of RA and Lys436. In addition, Arg51 is the gate keeper of the binding pocket. RA cannot pass through these narrow gaps on its way to the active site. Therefore, the helix containing Lys436 must move to adopt a more open conformation for substrate binding. Pocket B is a region for substrate oxidation, which includes an active center consisting of heme group and amino acid residues (Arg336, Cys403 and Arg336). The cyclohexene in RA is positioned in the pocket B and is bound with the heme, which will be oxidized into 4-hydroxy-RA [33]. A relatively small pocket C is located in the top region of pocket B, which also consists of three hydrophobic amino acid residues (Pro74, Val77 and Leu182), and the methyl group of RA occupies the pocket (Fig. 2).

## 3. Molecular superimposition

Despite the low sequence identity among CYPs from different organisms, the active sites of their three-dimensional structures have conserved sequence and are ubiquitously heme-containing, which is involved in oxidative metabolism that catalyzes the



**Fig. 1.** (A) Sequence alignment of CYP26A1, CYP120 and CYP51A with conserved constitution, deep blue; semiconserved constitution, blue. (B) Model evaluation by Ramachandran plot.

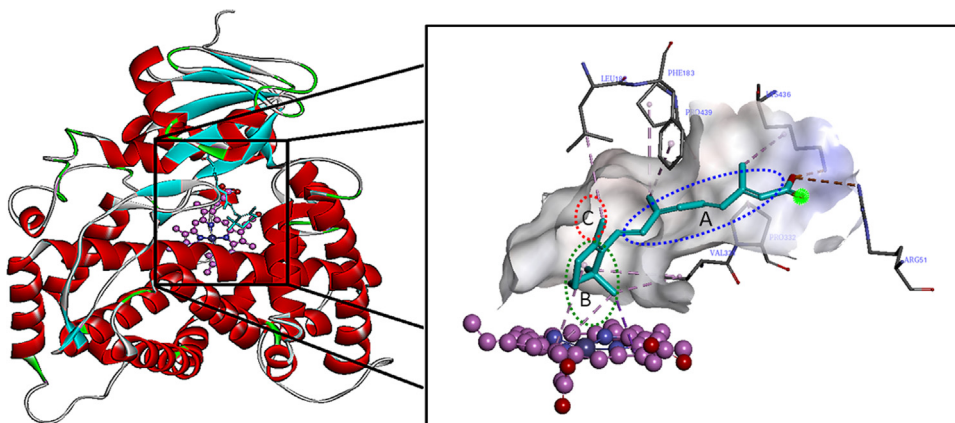
hydroxylation of activated substrates [34]. Therefore, they must have similar binding mode to metabolize exogenous substances. The antimycotic substance ketoconazole, itraconazole, miconazole, fluconazole and econazole are the first generation retinoic

acid metabolism blocking agents [18–20], which act not only on CYP26A1, but also on other homologous P450 enzymes. In order to understand how the first generation RAMBAs interacts with homologous P450 enzymes, we chose the crystal structures of P450 enzymes in complex with those RAMBAs for further analysis. Those P450 enzymes have high homology with CYP26A1, such as 3LD6 [30c], 2WX2 [30d], 4G3J [30e], 4G7G [30e], 2V0M [30f], 3MDT [30g]. The results are shown in Tables 2 and 3. The binding sites in the CYP120, CYP51, CYP3A4 and CYP46A1 complexes consist of broad and extended active chambers. A plurality of hydrophobic amino acids residues (Trp, Leu, Phe and Ile, etc.) were observed in the same conserved active site, which were involved in hydrophobic interactions with ligands. The heme group is required for oxidation metabolism, and it mediated coordinate bond interactions between RAMBAs and homologous P450s in the active chambers of the crystal structures.

These ligands were superimposed in the active site of homology model CYP26A1, and the result can be used directly as the analysis of ligand binding conformation. In the model of molecular superimposition, the steric and electrostatic were set to 50% respectively using the program Molecular Overlay [36]. Retinoic acid in its CYP120 bound conformation was set as the reference molecule. Compared with retinoic acid, the ligands overlay rates are greater than 70%, which suggests the binding modes of RAMBAs have similar spatial structure (Fig. 3). What is more, the main chain groups of RAMBAs contain the hydrophobic aromatic groups, which may have strong hydrophobic interactions with nonpolar amino acids residues (ketoconazole in 3LD6 or 2V0M, fluconazole in 2WX2, econazole in 3JUS, VNI derivative in 4G3J or 4G7G, voriconazole in 3MDT). In addition, the relatively small azole groups are located in the core of receptors (P450 enzymes), which can produce coordinate bond with the ferrous ion in heme, so the first generation RAMBAs might inhibit the retinoic acid metabolism through competitive binding with the ferroheme. Furthermore, the side chain groups of first generation RAMBAs were composed of a substituted benzene ring, which occupy the internal hydrophobic cavity. Further investigation revealed that the groups might play an important role in the affinity of the P450 enzymes (see below).

#### 4. Common feature pharmacophore model and validation

A common feature pharmacophore model explains how structurally diverse ligands can bind to a common receptor site. Among the reported CYP26A1 inhibitors, eight representative CYP26A1 inhibitors with diverse scaffolds were chose as a training set (Tables 3 and 4), which act on the same target CYP26A1 with the same mode of action. The common feature pharmacophore was generated using the HipHop program of Discovery Studio



**Fig. 2.** Three-dimensional representation of the model obtained for CYP26A1 associated with retinoic acid. Pocket A, blue; pocket B, green; pocket C, red.



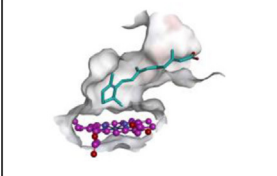
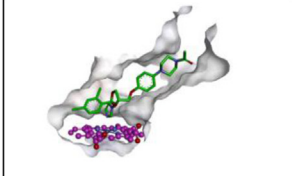
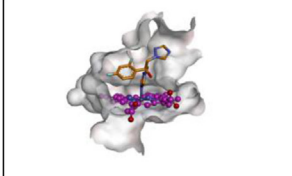
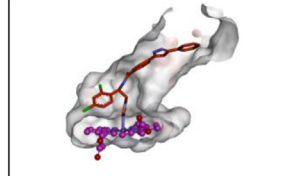


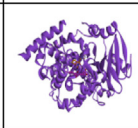

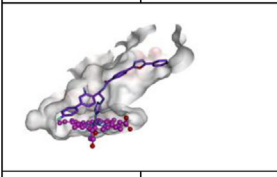
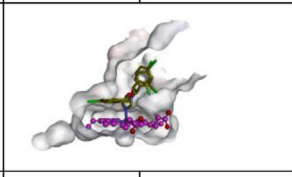
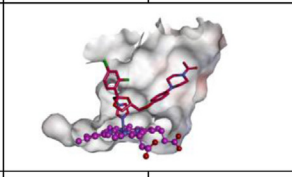
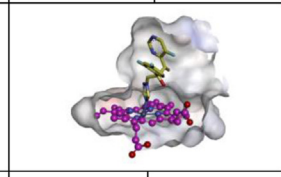
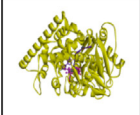
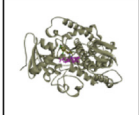
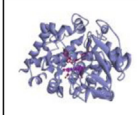
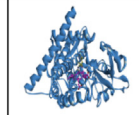
**Table 2**

Summary of homologous P450 crystal molecules with RAMBAs.

Index	PDB code	Ligand	Resolution	Overlay similarity scores	Hydrophobic residues
1	2VE3(CYP120)	Retinoic acid	2.1	1	Trp80, Leu88, Phe182, Phe253, Val426
2	3LD6 (CYP51)	Ketoconazole	2.8	0.826	Tyr131, Thr135, Phe139, Trp239, Ile377, Leu310
3	2WX2 (CYP51)	Fluconazole	2.27	0.872	Tyr103, Leu356, Val461
4	4G3J (CYP51)	VNI derivative1	1.83	0.860	Val102, Thr295, Leu357, Val213
5	4G7G (CYP51)	VNI derivative2	2.05	0.828	Val102, Phe105, Met106, Ala287, Phe290
6	3JUS (CYP51)	Econazole	2.9	0.738	Thr135, Val143, Phe234, Thr315
7	2V0M(CYP3A4)	Ketoconazole	2.8	0.799	Phe57, Ala370, Leu482, Leu210
8	3MDT (CYP46A1)	Voriconazole	2.3	0.857	Leu112, Phe121, Phe299, Ile301, Ala302

**Table 3**

Crystal structures of RAMBAs-bound homologous cytochrome P450s.

			
 retinoic acid CYP120 PDB:2VE3	 CYP51A PDB:3LD6	 CYP51 PDB:2WX2	 CYP51 PDB:4G3J
			
 CYP51 PDB:4G7G	 CYP51 PDB:3JUS	 CYP3A4 PDB:2V0M	 CYP46A PDB:3MDT

4.0 software [35,36]. In the model generation methodology, the highest weighting was assigned to the most active compounds in the training set (compound **2**, IC<sub>50</sub> 3 nM; compound **6**, IC<sub>50</sub> 5 nM), which was achieved by putting “2” in principal value, and it can ensure that all of the chemical features in the compound will be considered in building hypotheses space, and “0” in MaxOmitFeat value, which forces mapping of all features of compound, respectively. For the most active compounds (**1**, **5**, **7** and **8**) put “1” in principal value and MaxOmitFeat value, ensuring that at least one mapping for each of generated hypotheses will be found. The compounds (**3** and **4**) with “0” and “2” in principal value and MaxOmitFeat value, respectively, which indicates that the compounds will not be considered in building pharmacophore model feature

**Table 4**

Activities of selected compounds derived from the CYP26A1 domain enzymatic assays.

Index	Name	IC <sub>50</sub> (nm) <sup>a</sup>	Ref.
1	OSI Pharma 15C	50	[23]
2	Naphthyl compound	3	[37a]
3	Benzimidazolyl compound	2500	[37b]
4	Iiazozole	6000	[37c]
5	Phenyl compound	450	[37d]
6	R115866	5	[24]
7	R116010	10	[8]
8	OSI Pharma 1C	34.7	[37e]

<sup>a</sup> The imidazole and triazole derivatives were evaluated for their ATRA metabolism (CYP26A1) inhibitory activity with a cell-free microsomal assay, using radiolabeled ATRA as the substrate, each inhibitor was calculated on the basis of multiple experiments at different inhibitor concentrations.

and mapping of the features of compound. All other parameters were kept at default.

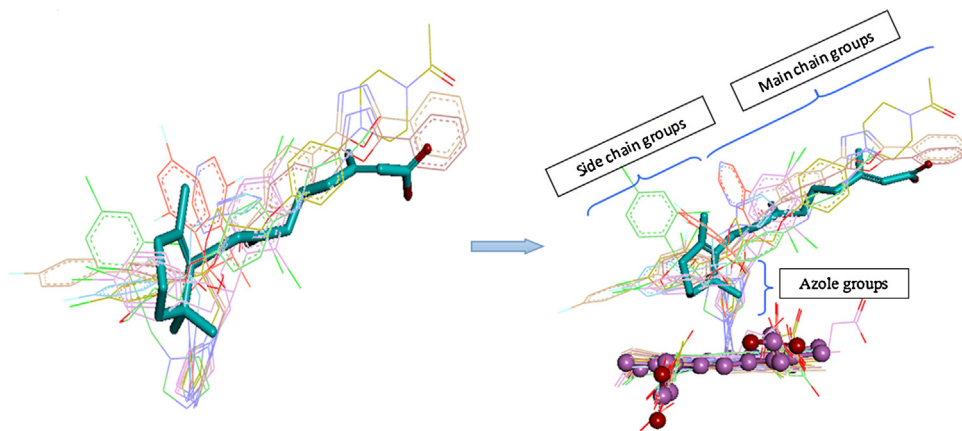
The ten pharmacophore hypotheses generated had scores ranging from 66.263 to 71.192 (Table 5). This small range of ranking score suggested that all 10 feature models in space arrangement have similar fashion. Cluster analysis was used to estimate and sort the difference between the components and the locations of chemical features. All 10 hypotheses contain the similar features, of which four hydrophobic (H) and one hydrogen bond acceptor (A) features are common for all. The hypotheses were differentiated by the location of features, or the direction of hydrogen bond vectors, or both. Unreasonable hypotheses that specified the characteristics of a small number of compounds and had identical chemical characteristics and nearly the same distances between these features were deleted. Hypotheses with diverse configurations were selected according to the ranking scores and fit values. Finally, based on the analysis results, the hypotheses 01, 02, 04 and 08 were selected for further analysis (Table 6).

From the results of fit scores in Table 6, it is difficult to select the overall best hypo based on this analysis. However, hypo 02 shows the overall better fit score to map the training set compounds than the corresponding other hypotheses. Thus, hypo 02 was indicated as the best and final ligand-based model, which contains five chemical features: four hydrophobic (H1, H2, H3 and H4) and one hydrogen bond acceptor (A) features (Fig. 4). The hydrophobic groups (H1, H2, H3 and H4) are defined by either aliphatic groups, such as methyl, dimethyl, ethyl, tertiary butyl, and aromatic moieties. And the hydrogen bond acceptor (A) consists of either oxygen atoms in ester groups or nitrogen atoms with sp<sup>2</sup> hybridization in

**Table 5**  
Summary of the pharmacophore models generated by HipHop for CYP26A1 inhibitors.

Hypothesis	Features <sup>a</sup>	Rank <sup>b</sup>	Direct Hit <sup>c</sup>	Partial Hit <sup>d</sup>	Max Fit
01	HHHHA	71.192	111111	000000	5
02	HHHHA	69.931	111111	000000	5
03	HHHHA	69.433	111111	000000	5
04	HHHHA	68.917	111111	000000	5
05	HHHHA	67.956	111111	000000	5
06	HHHHA	67.796	111111	000000	5
07	HHHHA	67.204	111111	000000	5
08	HHHHA	66.820	111111	000000	5
09	HHHHA	66.653	111111	000000	5
10	HHHHA	66.263	111111	000000	5

<sup>a</sup> H, hydrophobic group; A, hydrogen bond acceptor.  
<sup>b</sup> The ranking score of training set compounds fitting the hypothesis.  
<sup>c</sup> Direct Hit indicates whether ("1") or not ("0") a molecule in the training set mapped every feature in the hypothesis.  
<sup>d</sup> Partial Hit indicates whether ("1") or not ("0") a particular molecule in the training set mapped all but one feature in the hypothesis. Numeration of molecules is from right to left in both Direct Hit and Partial Hit.

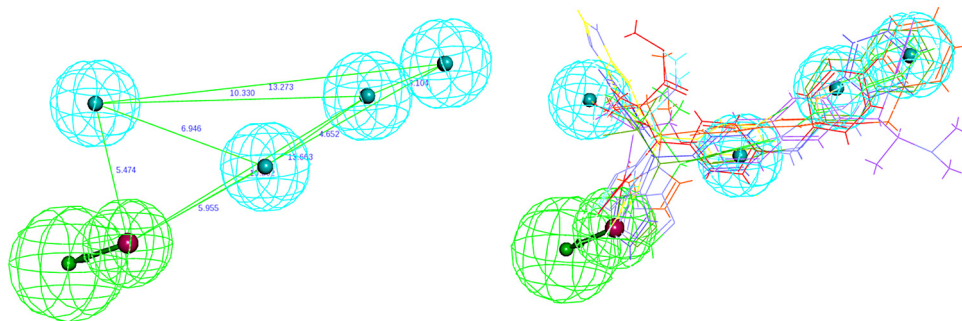


**Fig. 3.** The molecular superimposition of the first generation RAMBAs.

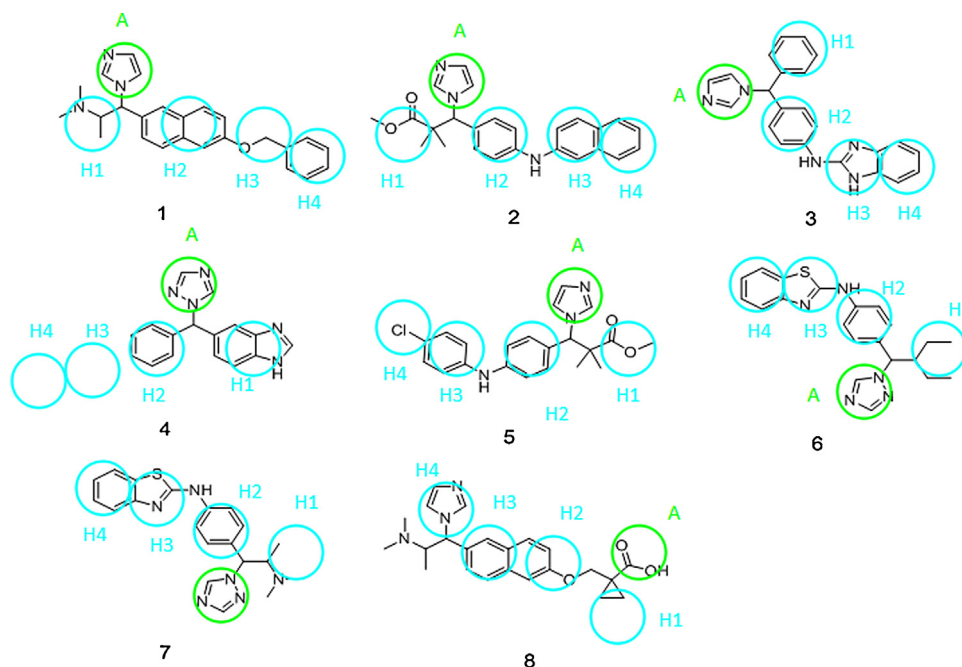
**Table 6**  
The validation of hypotheses 01, 02, 04 and 08 by scaling the fit values, and conformation values of the compounds in training set.

Training set compounds	Bioactivity values		Scaled-Fit Values <sup>[b]</sup>			
	IC <sub>50</sub> (nm)	Confon <sup>a</sup>	Hypo 01	Hypo 02	Hypo 04	Hypo 08
1	50	42	1.747	1.298	1.719	1.706
2	3	25	4.999	5.000	3.603	4.153
3	2500	12	1.062	3.010	0.487	0.509
4	6000	27	2.019	0.448	1.922	1.740
5	450	65	4.207	3.701	3.599	4.319
6	5	107	4.066	4.127	4.999	4.999
7	10	11	3.191	3.661	2.888	3.411
8	34.7	103	1.742	2.465	2.798	2.960

<sup>a</sup> Diverse set of conformations for all molecules was generated by using 'Best conformer generation' option with 20 kcal/mol energy cutoff.  
<sup>b</sup> Compare fit analysis of all compounds with maximum omitting feature value as 1.



**Fig. 4.** Selected common feature pharmacophore model hypos 02 consisting of four hydrophobic groups (H), and a hydrogen bond acceptor (A). Distances between the features are expressed in Å, with a tolerance sphere of radius  $\pm 0.9$  Å. For hydrogen bond acceptor A, the small sphere represents the hydrogen bond acceptor on training set compound, and the big sphere represents the corresponding hydrogen bond donor on the protein residues.



**Fig. 5.** 2D mapping of the ligand-based pharmacophore features onto the training set compounds. The features are as follows: hydrophobic group (H) in cyan, hydrogen bond acceptor (A) in green.

2N-imidazole or triazole group. The chemical features mapped to the eight inhibitors are depicted in Fig. 5, and the result shows that the final ligand-based model hypo 02 was overlaid on a majority of training set compounds. The main-chain groups and side-chain groups of compounds can match the hydrophobic groups (H1, H2, H3 and H4), and the azole groups match the hydrogen bond acceptor (A).

The hypo 02 was further validated as the final ligand-based model through investigating the SAR according to the relevance between the fit values and biological activity values. For this purpose, a test set consisting of five reported CYP26A1 inhibitors and four inactive compounds were mapped onto the final ligand-based model 02, and a series of align/fit values from 0.2 to 0.948 were generated (Table 7). Analysis of these fit values showed that the active inhibitors could be distinguished from the inactive ones. For inhibitors with a carbomethoxy core (such as 1, 2 and 3), replacement of the benzofuranyl group (such as 5) increased the activity of inhibitors toward CYP26A1, which can be explained that the changes of the side chain have an impact on the activity of inhibitors. Moreover, the inactive compounds were unable to match onto the hydrogen bond donors and acceptors or the two hydrophobic features at the same time. From the results above, it can be concluded that the selection of hypo 02 is reasonable. The research provides confidence for the utility of the chemical feature-based pharmacophore model to distinguish the most active inhibitors from the inactive compounds.

## 5. Docking study and inhibitor binding analysis

Molecular docking was performed to investigate the binding modes of the four representative RAMBAs in the pharmacophore training set (liarazole, phenyl compound, R115866 and OSI Pharma 1C), which can help us rationalize the available SAR and evaluate the regional influence on biological activity. The study was carried out using CDocker [39] and the following factors were taken into account in molecular docking: van der Waals interactions, hydrogen bonding, deformation penalty, and hydrophobic effects

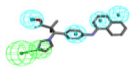
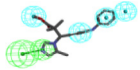
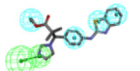
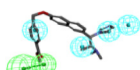
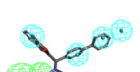
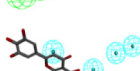
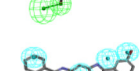

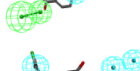
between the receptor and the ligand. Furthermore, CDocker can predict the binding free energy, and accompany an acceptable success rate in protein–ligand systems. The prepared homology model of CYP26A1 was imported into Discovery Studio 3.5, and was defined as the receptor structure excluding the co-crystallized ligand RA whose position determined the location of the active site.

The docking results were showed in Table 8. The binding interaction energy values of the compounds were predicted as –33.56, –45.37, –45.90 and –48.06 kcal/mol, respectively, which suggests that the blocking agents (phenyl compound, R115866 and OSI Pharma 1C) have better activity than the control group (liarazole). Moreover, the coordinate bond distances of compounds (2.72, 2.42, 2.45 and 2.34 Å) were significantly lower than that of the co-crystallized ligand retinoic acid (4.86 Å) in active site of CYP26A1, which suggests that the RAMBAs have stronger binding capacity than RA. RMS deviation values have also been calculated to measure docking reliability by comparing with the position in its protein structure. Analysis of the results by docking simulations showed that the values range from 0.24 to 1.43, which were considered as similar mode of docking of the ligands. The docked conformations in catalytic site of CYP26A1 were displayed in Fig. 6.

The RAMBAs (1, 2, 3 and 4) have a similar mode of action with the first generation RAMBAs (Fig. 6C and D). The aromatic groups of the main chains fit pocket A better than other parts of inhibitors. The pi–pi contacts were formed between the amino acid residues (Trp73, Phe183, Phe260 and Phe335) and aromatic groups of the blocking agents. It is worth noting that the introduction of nitrogen atoms between the aromatic ring of compounds (2 and 3) may form weaker hydrogen bonds with residue Pro439. Overall, these groups of the main chain may play important roles in the selectivity and affinity of inhibitors.

In addition, the first generation RAMBAs lacks specificity leading to unwanted side effects. Compounds (2, 3 and 4) are more potent than the first generation RAMBAs and have greater specificity toward CYP26A1. Through the molecular docking, it was found that the fragments of side chain have better match for the active pocket C than the substituted phenyl group, such as the ester

**Table 7**  
Mapping of the test set onto hypothesis 02. Compounds **1–5** are active CYP26A1 inhibitors, whereas compounds **6–9** are inactive.

Test set compounds	3D mapping onto hypo 02	$\Delta E$ (kcal/mol)	Fit value	IC <sub>50</sub> (nM)
<b>1</b>		12.678	0.779	6 [22]
<b>2</b>		8.017	0.789	90 [37e]
<b>3</b>		16.459	0.948	100 [22]
<b>4</b>		12.415	0.709	100 [22]
<b>5</b>		1.645	0.678	7000 [38a]
<b>6</b>		0	0.2	- [38b]
<b>7</b>		0.177	0.613	- [38b]
<b>8</b>		22.455	0.680	- [38b]
<b>9</b>		10.618	0.582	- [38b]

**Table 8**  
Summary of the docking studies of four representative RAMBAs in the training set.

Compounds	Activity	CDOCKER-interaction-energy	RMS deviation	Coordinate bond-distance
(1) Liarozole	6000	33.56	0.24	2.72
(2) Phenyl compound	10	45.37	1.21	2.42
(3) R115866	5	45.90	0.34	2.45
(4) OSI Pharma 1C	34.7	48.06	1.43	2.34

fragment of compound **2**, alkyl fragment of compound **3** and tertiary amino fragment of compound **4**. We deduce that those fragments of side chain might decrease the steric clashes of active pocket C, and augment the selectivity of blocking agents toward CYP26A1. Thus those conformations of compounds match the active site of CYP26A1 better than the first generation RAMBAs.

Through further comparing the interactive modes and biological activities of RAMBAs, we observed that the conserved heme molecule and pocket B also have important influences on the affinity of blocking agents toward CYP26A1. The imidazole group or triazole group occupies pocket B, and the charge–charge interactions are formed between the basic nitrogen atoms of the group and the heme in the bottom of pocket B. Thus the imidazole group or triazole group has an important role in achieving the high affinity to block the metabolism of retinoic acid.

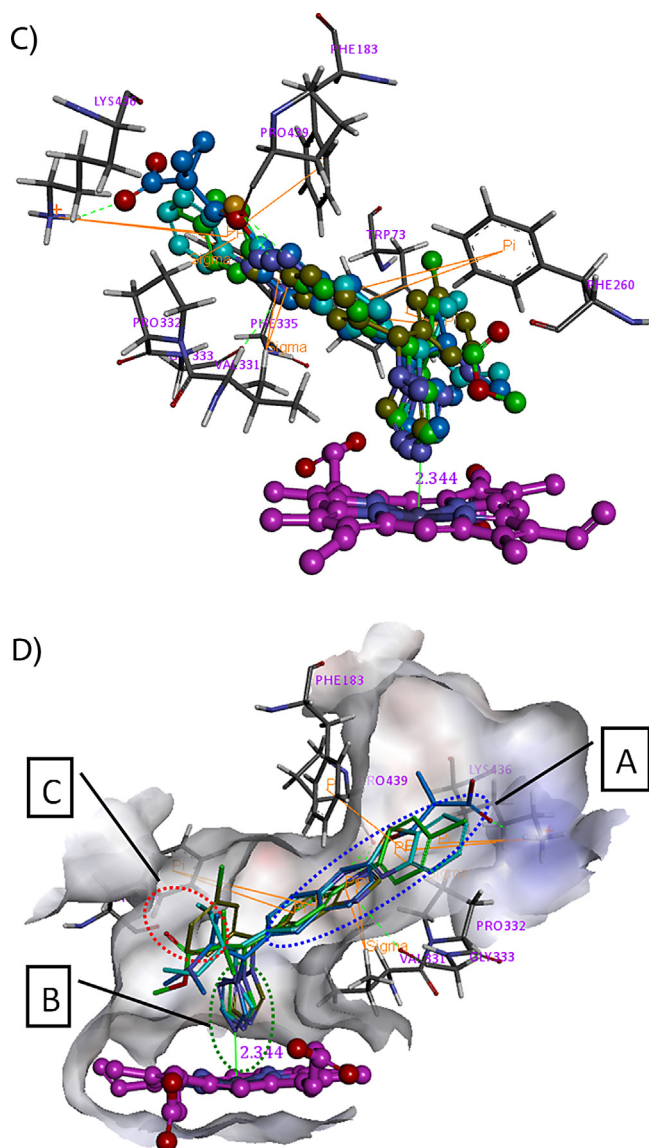
In conclusion, the imidazole group, main chain groups, and side chain groups are essential regions for CYP26A1 potent inhibitory activity. The main-chain groups contain aromatic groups that form pi–pi interactions with amino acid residues. The side chain groups can improve activity toward CYP26A1 by a better fitting into pocket C. For the imidazole group may generate a coordinate bond

with heme group, which has important role in achieving high affinity.

## 6. Comparison of the molecular superimposition and docking results with the pharmacophore model: toward an integrated model within the CYP26A1 active site

In order to evaluate the relevance of ligand-based pharmacophore model to the molecular superimposition and docking results, the molecular superimposition and molecular docking with the pharmacophore model were produced (Fig. 7E and F). These results suggest that the bioactive conformations of RAMBAs are similar, and that they match the ligand-based pharmacophore model well. Three hydrophobic groups (H2, H3 and H4) occupy the aromatic groups of main chain, the pi–pi interactions are formed between the hydrophobic pocket A and aromatic groups of the RAMBAs, which contribute to protein–ligand interactions. Hydrophobic groups H1 match the side-chain groups. Through further analyzing the structure of side chain groups and the active pocket C, it was found that the side chain fragments can fit into the pocket C more easily than the substituted phenyl group. The

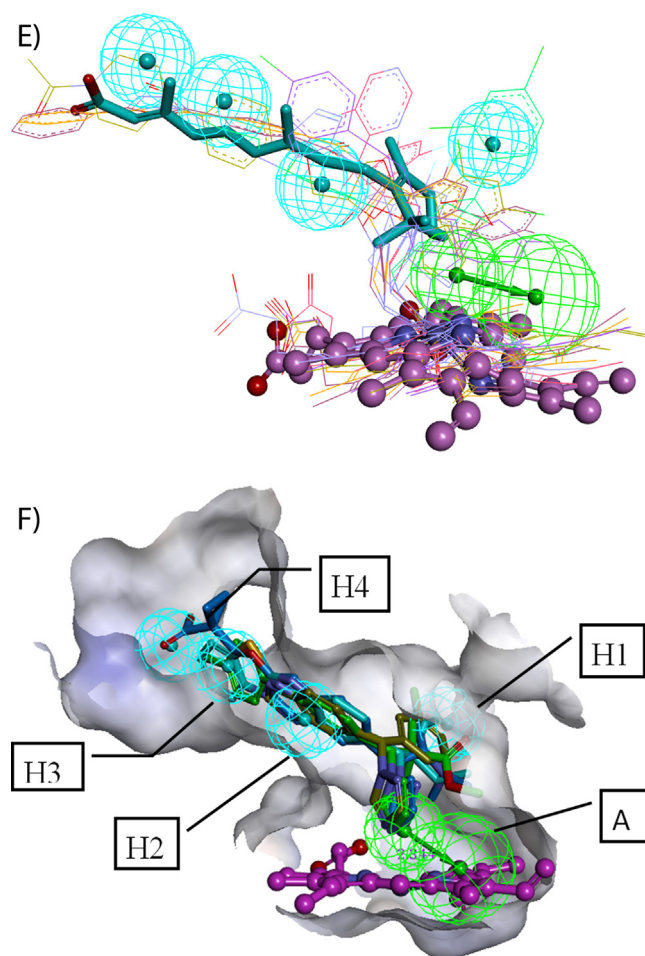




**Fig. 6.** (C) Docking of inhibitors (1, 2, 3 and 4) into the active site of CYP26A1. (D) The hydrophobic surface of the pockets are also indicated pocket A, blue; pocket B, green; pocket C, red.

hydrogen bond acceptor (A) is positioned to the imidazole groups, and a large number of studies have found the imidazole groups may form coordinate bond with the heme [30a–g]. Thus it is suggested that the imidazole groups produce pharmacological activity through the way of competitive inhibition retinoic acid binding.

The evidence in this study confirmed that the proposed ligand-based pharmacophore model was able to fit the binding cavity and match well with the topology of the active site. Based on the summary of molecular superimposition and docking results, common feature pharmacophore model of the RAMBAs are proposed as follows: (1) the main chain of RAMBAs contains a large aromatic group located at hydrophobic groups (H2, H3, H4) and forms pi–pi interactions with the hydrophobic amino acid residues; (2) the side chain of RAMBAs contains a flexible hydrophobic fragment occupying the hydrophobic group H1 and the active pocket C; (3) the nitrogen atom of the imidazole group forms the coordinate bond with the conserved heme molecule. These points were emphasized during the model generation, because they are shared by all the blocking



**Fig. 7.** (E) The molecular superimposition with the pharmacophore model. (F) Molecular docking results with the pharmacophore model. The five key pharmacophore features (H1, H2, H3, H4, A) appropriate matching RAMBAs.

agents. These interaction points could be used in the design of new CYP26A1 blocking agents.

## 7. Structure–activity relationship (SAR) analysis

Based on the analysis of molecular superimposition model, pharmacophore model and molecular docking results, the structure–activity relationship (SAR) was summarized as follows: (1) the main chain part of RAMBAs contain aryl fragments, such as combined benzene ring, pyrrole ring, thiazole ring, etc. It was observed that the main chain part interacted with amino acid residues at the active site by multiple hydrophobic interactions. (2) The potency of inhibitors would decrease if a benzene ring fragment is introduced into side chain part, such as R115866, R116010, 2,2-dimethyl-3-(4-(naphthalen-2-ylamino)phenyl) propylazole derivatives, etc. The potency of inhibitors would reduce if a benzene ring fragment is introduced into side chain part, such as econazole, ketoconazole, liarazole, etc. We suspect the flexible side chain fragments may reduce the steric hindrance in pocket C, and optimize spatial arrangements within the CYP26A1 active site. (3) The imidazole or triazole part is necessary basic drug group of RAMBAs, small changes. The 4-nitrogen atom of the imidazole or triazole fragment can form the coordinate bond with the conserved heme molecule. It is an essential part to inhibit CYP26A1 activity (Fig. 8).



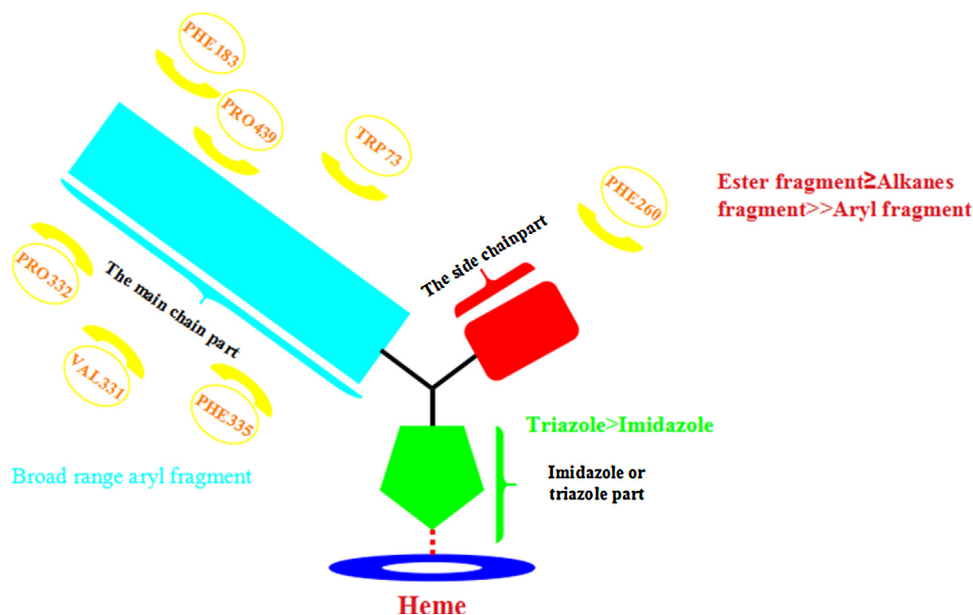


Fig. 8. Optimal SAR for RAMBAs.

## 8. Conclusions

In this study, we generated a homology model for human CYP26A1 using the crystal structures of CYP120 (PDB code 2VE3) and CYP51A (PDB code 3JUS) as templates. Further analysis found that the binding site of RA includes several crucial hydrophobic amino acid residues (Trp73, Phe183, Phe260, Val331, Phe335, Val331 and Pro439) and two basic amino acid residues (Arg51, Lys436). In order to achieve a deeper insight into the action mode of RAMBAs, the molecular superimposition model and the common feature pharmacophore model were constructed, and molecular docking was performed. These models provide more detailed information and description of ligand binding. The molecular superimposition model is composed of three features: the main chain groups, side chain groups, and imidazole groups. The common feature pharmacophore model consists of five chemical features: four hydrophobic groups and one hydrogen acceptor (HHHHA). The result of molecular docking shows that the characteristic groups of RAMBAs were combined into three different active pockets, respectively.

Finally, the structure–activity relationship was obtained by a combination of the molecular superimposition and docking results with the pharmacophore model: a main chain consisting of hydrophobic aromatic groups, a side chain consisting of a flexible hydrophobic fragment, and an imidazole group or triazole group. The main chain and side chain are stabilized inside the binding site mainly by pi–pi interactions with hydrophobic amino acid residues, and the imidazole group or triazole group is located in the core of receptors (CYP26A1), which forms coordinate bond with the heme group. The present study represents a step forward to understand the characterization of human CYP26A1 and its interaction with ligands. It provides more insight into discovery of new lead compounds inhibiting CYP26A1 and their optimizations.

## Acknowledgments

This work was supported by Grants from the National Natural Science Foundation of China (Grant no. 30973643). This work was supported by program for innovative research team of the ministry

of education and program for Liaoning innovative research team in university.

## References

- [1] M.A. Smith, D.R. Parkinson, B.D. Cheson, M.A. Friedman, Retinoids in cancer therapy, *J. Clin. Oncol.* 10 (1992) 839–864.
- [2] L.J. Gudas, J.A. Wagner, Retinoids regulate stem cell differentiation, *J. Cell. Physiol.* 226 (2011) 322–330.
- [3] J.S. Lee, R.A. Newman, S.M. Lippman, M.H. Huber, T. Minor, et al., Phase I evaluation of all-trans-retinoic acid in adults with solid tumors, *J. Clin. Oncol.* 11 (1993) 959–966.
- [4] J.L. Armstrong, M. Ruiz, A.V. Boddy, C.P.F. Redfern, A.D.J. Pearson, et al., Increasing the intracellular availability of all-trans-retinoic acid in neuroblastoma cells, *Br. J. Cancer* 92 (2005) 696–704.
- [5] B. Ozpolat, K. Mehta, G. Lopez-Berestein, Regulation of a highly specific retinoic acid-4-hydroxylase (CYP26A1) enzyme and all-trans-retinoic acid metabolism in human intestinal, liver, endothelial, and acute promyelocytic leukemia cells, *Leuk. Lymphoma* 46 (2005) 1497–1506.
- [6] L.C. McSorley, A.K. Daly, Identification of human cytochrome P450 isoforms that contribute to all-trans-retinoic acid 4-hydroxylation, *Biochem. Pharmacol.* 60 (2000) 517–526.
- [7] S.M. Thacher, J. Vasudevan, K.Y. Tsang, S. Nagpal, R.A. Chandraratna, New dermatological agents for the treatment of psoriasis, *J. Med. Chem.* 44 (2001) 281–297.
- [8] J.L. Armstrong, A.V. Boddy, C.P.F. Redfern, G.J. Veal, Molecular targeting of retinoic acid metabolism in neuroblastoma: the role of the CYP26 inhibitor R116010 in vitro and in vivo, *Br. J. Cancer* 96 (2007) 1675–1683.
- [9] J. Marill, N. Idres, C.C. Capron, E. Nguyen, G.G. Chabot, Retinoic acid metabolism and mechanism of action: a review, *Curr. Drug Metab.* 4 (2003) 1–10.
- [10] H. Chen, W.N. Howald, M.R. Juchau, Biosynthesis of all-trans-retinoic acid from all-trans-retinol: catalysis of all-trans-retinol oxidation by human P-450 cytochromes, *Drug Metab. Dispos.* 28 (2000) 315–322.
- [11] J. Marill, T. Cresteil, M. Lanotte, G.G. Chabot, Identification of human cytochrome P450s involved in the formation of all-trans-retinoic acid principal metabolites, *Mol. Pharmacol.* 58 (2000) 1341–1348.
- [12] W.J. Ray, G. Bain, M. Yao, D.I. Gottlieb, A novel mammalian cytochrome P450, is induced by retinoic acid and defines a new family, *J. Biol. Chem.* 272 (1997) 18702–18708.
- [13] G. MacLean, S. Abu-Abed, P. Dolle, A. Tahayato, P. Chambon, et al., Cloning of a novel retinoic-acid metabolizing cytochrome P450, Cyp26B1, and comparative expression analysis with Cyp26A1 during early murine development, *Mech. Dev.* 107 (2001) 195–201.
- [14] M. Taimi, C. Helvig, J. Wisniewski, H. Ramshaw, J. White, et al., A novel human cytochrome P450, CYP26C1, involved in metabolism of 9-cis and all-trans isomers of retinoic acid, *J. Biol. Chem.* 279 (2004) 77–85.
- [15] S. Abu-Abed, P. Dollé, D. Metzger, B. Beckett, P. Chambon, et al., The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures, *Genes Dev.* 15 (2001) 226–240.

- [16] R. Zolfaghari, C.J. Cifelli, S.O. Lieu, Q. Chen, N.Q. Li, et al., Lipopolysaccharide opposes the induction of CYP26A1 and CYP26B1 gene expression by retinoic acid in the rat liver in vivo, *Am. J. Physiol. – Gastrointest. Liver Physiol.* 292 (2007) G1029–G1036.
- [17] O. Loudig, G.A. MacLean, N.L. Dore, L. Luu, M. Petkovich, Transcriptional co-operativity between distant retinoic acid response elements in regulation of Cyp26A1 inducibility, *Biochem. J.* 392 (2005) 241–248.
- [18] J.P. Van Wauwe, P.A.J. Janssen, Is there a case for P-450 inhibitors in cancer treatment? *J. Med. Chem.* 32 (1989) 2231–2239.
- [19] M. Ahmad, Study on cytochrome P-450 dependent retinoic acid metabolism and its inhibitors as potential agents for cancer therapy, *Sci. Pharm.* 79 (2011) 921–935.
- [20] E.L. Schwartz, S. Hallam, R.E. Gallagher, P.H. Wiernik, Inhibition of all-trans-retinoic acid metabolism by fluconazole in vitro and in patients with acute promyelocytic leukemia, *Biochem. Pharmacol.* 50 (1995) 923–928.
- [21] R. De Coster, W. Wouters, R. Van Ginckel, D. End, M. Krekels, et al., Experimental studies with Icarazole (R 75, 251): an antitumoral agent which inhibits retinoic acid breakdown, *J. Steroid Biochem. Mol. Biol.* 43 (1992) 197–201.
- [22] M.S. Goma, C.E. Bridgens, A.S. Aboraia, G.J. Veal, C.P. Redfern, et al., Small molecule inhibitors of retinoic acid 4-hydroxylase (CYP26): synthesis and biological evaluation of imidazole methyl 3-(4-(aryl-2-ylamino)phenyl)propanoates, *J. Med. Chem.* 54 (2011) 2778–2791.
- [23] M.J. Mulvihill, J.L. Kan, P. Beck, M. Bittner, C. Cesario, et al., Potent and selective [2-imidazol-1-yl-2-(6-alkoxy-naphthalen-2-yl)-1-methyl-ethyl]-dimethylamines as retinoic acid metabolic blocking agents (RAMBAs), *Bioorg. Med. Chem. Lett.* 15 (2005) 1669–1673.
- [24] C.J. Verfaillie, C.A. Thissen, H.J. Bovenschen, J. Mertens, P.M. Steiljen, et al., Oral R115866 in the treatment of moderate to severe plaque-type psoriasis, *J. Eur. Acad. Dermatol. Venereol.* 21 (2007) 1038–1046.
- [25] C.J. Verfaillie, M. Coel, I.H. Boersma, J. Mertens, M. Borgers, et al., Oral R115866 in the treatment of moderate to severe facial acne vulgaris: an exploratory study, *Br. J. Dermatol.* 157 (2007) 122–126.
- [26] A.N. Geria, N.S. Scheinfeld, Talarozole, a selective inhibitor of P450-mediated all-trans retinoic acid for the treatment of psoriasis and acne, *Curr. Opin. Invest. Drugs* 9 (2008) 1228–1237.
- [27] B. Boeckmann, A. Bairoch, R. Apweiler, M.C. Blatter, A. Estreicher, et al., The Swiss-Prot protein knowledgebase and its supplement TrEMBL in 2003, *Nucleic Acids Res.* 31 (2003) 365–370.
- [28] S.F. Altschul, J.C. Wootton, E.M. Gertz, R. Agarwala, A. Morgulis, et al., Protein database searches using compositionally adjusted substitution matrices, *FEBS J.* 272 (2005) 5101–5109.
- [29] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, et al., Protein data bank, *Nucleic Acids Res.* 28 (2000) 235–242.
- [30] (a) K. Kühnel, N. Ke, M.J. Cryle, S.G. Sligar, M.A. Schuler, et al., Crystal structures of substrate-free and retinoic acid-bound cyanobacterial cytochrome P450 CYP120A1, *Biochem. J.* 47 (2008) 6552–6559;  
(b) N. Strushkevich, S.A. Usanov, H. Park, Crystal structure of human lanosterol 14 $\alpha$ -demethylase (CYP51) in complex with econazole (2012), <http://dx.doi.org/10.2210/pdb3jus/pdb>;  
(c) N. Strushkevich, S.A. Usanov, H.W. Park, Structural basis of human CYP51 inhibition by antifungal azoles, *J. Mol. Biol.* 397 (2010) 1067–1078;  
(d) C.-K. Chen, S.S.F. Leung, C. Guilbert, M. Jacobson, J.H. Mckerrow, et al., Structural characterization of CYP51 from *Trypanosoma cruzi* and *Trypanosoma brucei* bound to the antifungal drugs posaconazole and fluconazole, *PLOS Negl. Trop. Dis.* 4 (2010) e651;
- (e) T.Y. Hargrove, K. Kim, M. de Nazaré Correia Soeiro, C.F. da Silva, D. da Gama Jaen Batista, et al., CYP51 structure-based VNI scaffold development, *Int. J. Parasitol. Drugs Drug Resist.* 2 (2012) 178–186;
- (f) M. Kroos, T. Sjogren, Structural basis for ligand promiscuity in cytochrome P450 3A4, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 13682–13687;
- (g) N. Mast, C. Charvet, I.A. Pikuleva, C.D. Stout, Structural basis of drug binding to CYP46A1, an enzyme that controls cholesterol turnover in the brain, *J. Biol. Chem.* 285 (2010) 31783–31795.
- [31] A. Fiser, A. Sali, Modeller: generation and refinement of homology-based protein structure models, *Methods Enzymol.* 374 (2003) 461–491.
- [32] A. Šali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [33] J. Marill, C.C. Capron, N. Idres, G.G. Chabot, Human cytochrome P450s involved in the metabolism of 9-cis- and 13-cis-retinoic acids, *Biochem. Pharmacol.* 63 (2002) 933–943.
- [34] F. Yamashita, C. Feng, S. Yoshida, T. Itoh, M. Hashida, Automated information extraction and structure? activity relationship analysis of cytochrome P450 substrates, *J. Chem. Inform. Model.* 51 (2011) 378–385.
- [35] O.O. Clement, A. Trope-Mehl, HipHop: Pharmacophores based on multiple common-feature alignments, in: O.F. Güner (Ed.), *Pharmacophore Perception, Development, and Use in Drug Design*, International University Line, La Jolla, CA, 2000, pp. 69–83.
- [36] Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, Accelrys Software, Inc., San Diego, 2013.
- [37] (a) M.S. Goma, C.E. Bridgens, G.J. Veal, C.P. Redfern, A. Brancale, et al., Synthesis and biological evaluation of 3-(1H-imidazol- and triazol-1-yl)-2,2-dimethyl-3-[4-(naphthalen-2-ylamino)phenyl]propyl derivatives as small molecule inhibitors of retinoic acid 4-hydroxylase (CYP26), *J. Med. Chem.* 54 (2011) 6803–6811;  
(b) M.S. Goma, C.E. Bridgens, A.S. Aboraia, G.J. Veal, C.P.F. Redfern, et al., Small molecule inhibitors of retinoic acid 4-hydroxylase (CYP26): synthesis and biological evaluation of imidazole methyl 3-(4-(aryl-2-ylamino)phenyl)propanoates, *J. Med. Chem.* 54 (2011) 2778–2791;  
(c) V.C. Njar, L. Gediya, P. Purushottamachar, P. Chopra, T.S. Vasaitis, et al., Retinoic acid metabolism blocking agents (RAMBAs) for treatment of cancer and dermatological diseases, *Bioorg. Med. Chem.* 14 (2006) 4323–4340;  
(d) M.S. Goma, A.S. Lim, S.C. Wilson Lau, A.M. Watts, N.A. Illingworth, et al., Synthesis and CYP26A1 inhibitory activity of novel methyl 3-[4-(arylamino)phenyl]-3-(azole)-2,2-dimethylpropanoates, *Bioorg. Med. Chem.* 20 (2012) 6080–6088;  
(e) M.J. Mulvihill, J.L. Kan, A. Cooke, S. Bhagwat, P. Beck, et al., 3-[6-(2-Dimethylamino-1-imidazol-1-yl-butyl)-naphthalen-2-yloxy]-2,2-dimethylpropionic acid as a highly potent and selective retinoic acid metabolic blocking agent, *Bioorg. Med. Chem. Lett.* 16 (2006) 2729–2733.
- [38] (a) S. Pautus, S.W. Yee, M. Jayne, M.P. Coogan, C. Simons, Synthesis and CYP26A1 inhibitory activity of 1-[benzofuran-2-yl-(4-alkyl/aryl-phenyl)-methyl]-1H-triazoles, *Bioorg. Med. Chem.* 14 (2006) 3643–3653;  
(b) F. Li, D. Zhao, J. Ren, F. Hao, G. Liu, S. Jin, M. Cheng, 2-(2-Methylfuran-3-carboxamido)-3-phenylpropanoic acid, a potential CYP26A1 inhibitor to enhance all-trans retinoic acid-induced leukemia cell differentiation based on virtual screening and biological evaluation, *Bioorg. Med. Chem.* 21 (2013) 3256–3261.
- [39] I.D. Kuntz, Structure-based strategies for drug design and discovery, *Science* 257 (1992) 1078–1082.