Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Application of 3D QSAR CoMFA/CoMSIA and *in silico* docking studies on novel renin inhibitors against cardiovascular diseases

Aggeliki Politi [a,b], Serdar Durdagi [b,c], Panagiota Moutevelis-Minakakis [a], George Kokotos [a], Manthos G. Papadopoulos [b], Thomas Mavromoustakos [a,b,∗]

[a] Capodistrian Athens University, Department of Chemistry, Zographou, 15771 Athens, Greece
[b] Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, Athens, Greece
[c] Department of Biology, Chemistry and Pharmacy, Freie Universität Berlin, Germany

A B S T R A C T

For the first time, a set of renin inhibitors were subjected to the 3D QSAR/CoMFA and CoMSIA studies. The utility of renin inhibitors in the treatment of cardiovascular diseases has not been fully explored yet. At the moment, aliskiren is the first and only existing renin inhibitor in the drug market. The performed 3D QSAR/CoMFA and CoMSIA in combination with docking studies included aliskiren and 37 derivatives possessing a wide variety of bioactivity. The obtained results may aid in the design of novel bioactive renin inhibitors.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Raised blood pressure, especially systolic pressure (hypertension), confers a significant cardiovascular risk and public health concern and should be actively treated. One of the major systems involved in the elevation of the pressure is the renin–angiotensin system (RAS) and subsequently its inhibition will have beneficial effects to lower blood pressure and improve cardiovascular health [1–3]. The RAS is regulated by a series of highly specific enzymatic reactions. The first enzymatic reaction in the pathway starts with renal production of renin that cleaves angiotensinogen to generate angiotensin I. Angiotensin I is then cleaved by angiotensin–converting enzyme (ACE) to generate the active peptide vasoconstrictive hormone angiotensin II. The octapeptide hormone angiotensin II binds to the AT_1 receptor to exert tissue specific effects that control blood pressure (Fig. 1). Drugs that inhibit biological action of renin are expected to have several potential advantages. Firstly, since renin catalyzes the rate-limiting step in the cascade of enzymic reactions of RAS, its blockade is expected to be more efficient. Secondly, the only known substrate for renin is angiotensinogen, suggesting that drugs that block the enzymatic reaction in which angiotensinogen is converted to angiotensin I must have higher specificity and fewer side effects in comparison to ACE inhibitors that act on the subsequent biochemical step. Thirdly, ACE independent pathways for generation of angiotensin II such as chymase are not affecting the first step of the enzymatic cascade of reaction. Fourthly, other possible harmful peptides derived from angiotensin I will not be formed [4–8].

The above potential advantages of renin inhibitors triggered the intense research efforts for developing of new renin inhibitors. The first synthetic molecules were either peptide or peptide-like renin inhibitors (i.e., remikiren and zanikiren) and suffered from their poor oral bioavailability, rapid elimination, low efficacy, and high cost of synthesis. Recently, the use of a combination of molecular modeling and crystallographic structure analysis led to the development of aliskiren, an orally effective, non-peptide inhibitor of human renin, approved by Food and Drug Administration (FDA) in March 2007 by Novartis with the commercial name Tekturna [9–13].

The X-ray crystal structure of recombinant human renin was first determined by Sialecki and co-workers at 2.5 Å resolution [14]. Crystal structure of human renin shows that the enzyme consists of

Fig. 1. The renin–angiotensin system (RAS).

Fig. 2. Structure of aliskiren (template compound **26**) and the pockets that occupies (top) and its crystal structure (bottom).

two mainly β-sheet domains related by an approximately two-fold axis. The active site cleft is located between the two domains and extends over eight residues of the respective substrate [15]. The eight residues of the substrate occupy eight subsites of the active site (S5, S4, S3, S2, S1, S1′, S2′, S3′). Each domain contains one of the catalytic essential aspartic acid carboxylates (Asp32 and Asp215). In the peptidic inhibitor–enzyme complexes, the inhibitors bind in an extended conformation occupying the following subsites of the active site (S4, S3, S2, S1, S1′, S2′). On the contrary, aliskiren acts as a transition state mimetic, inhibiting renin via hydrogen bonding of both the central hydroxyl group and amino function of the catalytic Asp32 and Asp215 residues and does not interact with the S2 or S4 binding sites of renin, instead interacts with a new one called $S3^{sp}$ [16]. The structure of aliskiren and its favored interactions with binding sites are shown in Fig. 2. Moreover, aliskiren proved to be a highly potent inhibitor of human renin and in contrast to all previous inhibitors, it is much more potent and long acting *in vivo* after oral administration in sodium-depleted marmosets and in hypertensive patients [17]. Based on these results, Maibaum et al. [18,19] expanded the studies on renin inhibitors and synthesized 5(S)-amino-4(S)-hydroxy-8-phenyl-octanecarboxamide analogues that can be described as aliskiren derivatives. The thirty reported compounds of this research group [18,19] were used as database for performing three-dimensional quantitative structure–activity relationships (3D QSAR) studies. To our knowledge the present study is the first application for 3D QSAR on non-peptide renin inhibitors. The previous 2D QSAR studies cover mainly peptido-mimetic renin inhibitors and the developed models showed that their inhibitory activity was largely depended upon the molecular weight of the compounds, van der waals radius related parameters of the substituents and the localized electronic effects [20,21].

The aim of our study is to extract a correlation between the biological activity of a training set of molecules and their 3D structure. The importance of steric and electrostatic characteristics is revealed by aligning structurally similar analogues using phar-macophoric features as structural superimposition guides.

The use of Comperative Molecular Field Analysis (CoMFA) and Comperative Molecular Similarity Indices Analysis (CoMSIA) approaches together provide better ability of visualization and interpretation of the obtained correlations in terms of field contributions [22,23]. Pharmacophore models are developed by using the most potent ligand of the training set as template. Each
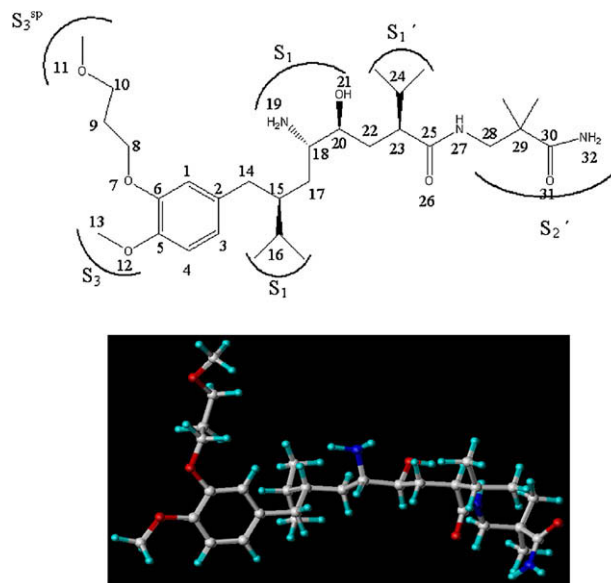
generated 3D QSAR model allows us to anticipate the predicted binding affinity values. To determine the linear correlation coeffi-cients between actual versus calculated binding affinities, partial least square (PLS) statistical analyses of the data were used. CoMFA and CoMSIA contour plots are used to explain different structural requirements for the inhibitors in the active site of renin. Contour results are used as pilot models for testing the designed novel analogues before their synthesis.

2. Results and discussion

CoMFA and CoMSIA techniques were used in order to derive stable 3D QSAR models for thirty 2,7-dialkyl-substituted 5(S)-amino-4(S)-hydroxy-8-phenyl-octanecarboxamides renin inhibi-tors used as a training set. The *in vitro* inhibitory concentrations (IC_{50}) of the molecules were converted into corresponding loga-rithmic values (pIC_{50}) and used as dependent variables. The produced QSAR models were further validated by performing a test set prediction using eight new molecules.
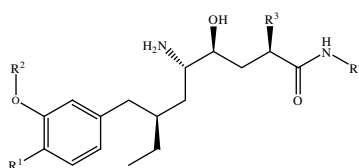
Table 1 lists all structures used in the training and test sets and their IC_{50} values obtained with purified human renin [18,19]. IC_{50} values differ up to 230-fold, thus there is a sufficient diversity in the data set in order to construct stable QSAR models. Among the used molecules for the QSAR study, the highly bioactive compound **26** (aliskiren) in Table 1, whose crystal structure is known and it has been resolved by co-crystallization with recombinant glycosylated human renin at 2.2 Å resolution [16] was selected as a template molecule.

The selected atoms of the template and common in all studied compounds for the superimposition during the alignment are C_1, C_2, C_3, C_4, C_5, C_6, O_7, C_8, C_{14}, C_{15}, C_{16}, C_{17}, C_{18}, N_{19}, C_{20}, O_{21}, C_{22}, C_{23}, C_{24}, C_{25}, O_{26}, N_{27}, C_{28} (Fig. 2). Fig. 3 illustrates the superimposition of all the molecules that consist of the training set.

The results obtained from the PLS analysis are summarized in Table 2. The predicted versus the experimental pIC_{50} values for the training set are listed in Table 3 and are depicted graphically in Fig. 4 for CoMFA and CoMSIA models, respectively. For CoMFA model, the leave-one out cross-validated r^2 value (r^2_{cv}) obtained was 0.628 and noncross-validated conventional r^2 value was 0.985
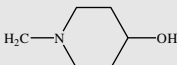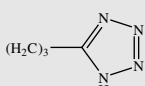
Table 1
The structure of **38** analogues used as training set (compounds **1–30**) and test set (compounds **31–38**), their inhibitory concentration (IC_{50}) values and the GoldScores of inhibitors in the training set.

Comp no.	R^1	R^2	R^3	R^4	IC_{50} (nM)	Chirality	GoldScore
Training set							
1	MeO	CH_2CONH_2	Me	$(CH_2)_3CH_3$	92		64.70
2	MeO	$CH_2CONHMe$	Me	$(CH_2)_3CH_3$	42		73.59
3	MeO	CH_2SO_2Me	Me	$(CH_2)_3CH_3$	50		72.28
4	MeO	$CH_2CH_2CH_2OH$	Me	$(CH_2)_3CH_3$	6		80.52
5	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3OH$	2		77.17
6	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2NMe_2$	18		75.33
7	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3NH_2$	16		83.28
8	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2$–N(piperidine)	22		55.75
9	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3$–N(morpholine)	11		65.36
10	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2CO_2H$	12		67.68
11	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3CO_2H$	3		70.86
12	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2CO_2Et$	4		73.32
13	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_4OCOMe$	6		70.59
14	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2CONH_2$	7		78.20
15	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3CONH_2$	1		93.46
16	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_4CONH_2$	2		68.37
17	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_2$–(2-pyridyl)	8		81.58
18	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$(CH_2)_3CONH_2$	0.7		64.65
19	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	CH_2CONH_2	3		77.32
20	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$CHMeCONH_2$	4	S	70.30
21	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$CH(CH_2OH)CONH_2$	2	S	67.25
22	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	CMe_2CONH_2	13		53.78
23	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$CHMeCH_2CONH_2$	6	R	86.47
24	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$CH_2CHMeCONH_2$	0.9	R	67.77
25	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$CH_2CHMeCONHMe$	1	S	79.52
26	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	CH_2–C(Me)$_2$–$CONH_2$	0.6		75.78
27	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	CH_2–C(Me)$_2$–$CONHMe$	0.4		72.85
28	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	$(CH_2)_3CONHMe$	0.8		77.07
29	*tert*-butyl	CH_2CONH_2	Me	$(CH_2)_3CH_3$	20		74.93
30	MeO	CH_2–C(Me)$_2$–$CONHMe$	Me	$(CH_2)_3CH_3$	5		92.50
Test set							
31	*tert*-butyl	CH_2CO_2Me	Me	$(CH_2)_3CH_3$	6		
32	*tert*-butyl	CH_2SO_2Me	Me	$(CH_2)_3CH_3$	13		
33	MeO	$CH_2CH_2CH_2OMe$	Me	$(CH_2)_3CH_3$	1		
34	MeO	$CH_2CH_2OCH_2CH_2OMe$	Me	$(CH_2)_3CH_3$	19		
35	MeO	$CH_2CH_2CH_2CH_2CH_3$	Me	$(CH_2)_3CH_3$	4		
36	MeO	$CH_2CH_2CH_2OMe$	*i*-propyl	C–CH$_2$–N(morpholine)	16		

Table 1 (continued)

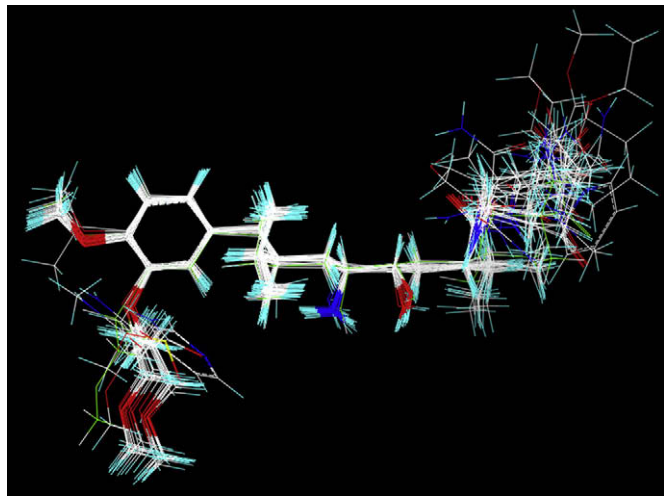| Comp no.	R^1	R^2	R^3	R^4	IC$_{50}$ (nM)	Chirality	GoldScore
37	MeO	CH$_2$CH$_2$CH$_2$OMe	Me	H$_2$C—N⌬—OH	10		
38	MeO	CH$_2$CH$_2$CH$_2$OMe	Me	(H$_2$C)$_3$—tetrazole	3		

Fig. 3. Structural alignment of the compounds consisting of the training set for constructing the 3D QSAR/CoMFA and CoMSIA models.

while PLS analysis for CoMSIA model showed that r^2_{cv} was 0.666 and r^2 was 0.971. Table 3 shows that average residual of the experimental and predicted values are 0.06 and 0.08 for CoMFA and CoMSIA models, respectively. Consequently, both models provide almost equal statistical confidence and productivity, as the linear plots of Fig. 4 confirm.

In order to validate constructed 3D QSAR models, eight compounds (**31–38**) were used as a test set. Their pIC$_{50}$ values ranged between 7.72 and 9.00 and their biological activities were predicted from the PLS equations derived from CoMFA and CoMSIA models (Table 4). The obtained average deviation of pIC$_{50}$ values from the study is 0.64 for CoMFA model and 0.42 for CoMSIA model suggesting that both models give a good prediction for the estimation of affinities of the novel compounds prior to their synthesis (Table 4).

Figs. 5 and 6 depict the stereoelectronic contour plots for compound **1** possessing the lowest bioactivity and **26** the orally potent inhibitor (Aliskiren) in the data set for the CoMFA and CoMSIA models. The steric field defined by the green colored contours represent regions of high steric tolerance (80% contribution), while yellow colored contours represent regions of unfavorable steric effect (20% contribution). The blue colored contours represent the regions where positively charged groups enhance the activity (80% contribution) and red colored contours where the negatively charged groups enhance the activity (20% contribution).

The CoMFA steric map displays yellow colored contours bordering the regions A and C. CoMFA electrostatic map displays red colored contours at the southern part of the region A and within region C. Blue colored contours are observed at the northern part of

Table 2
Summary of CoMFA and CoMSIA results.

	CoMFA	CoMSIA
Statistical results		
r^2_{cv} [a]	0.628	0.666
r^2	0.985	0.971
SEE[b]	0.082	0.117
Probability of r^2_{ncv}	0.000	0.00
F	315.292	128.684
Components	5	6
Fraction (%)		
Steric	0.495	0.776
Electrostatic	0.505	0.224
$r^2_{bootstrapping}$	0.993	0.988
SEE$_{bootstrapping}$	0.005	0.006

[a] Cross-validated r^2 by leave-one-out method.
[b] Standard error of estimate.

Table 3
Experimental and predicted activities (pIC$_{50}$) of the training set molecules.

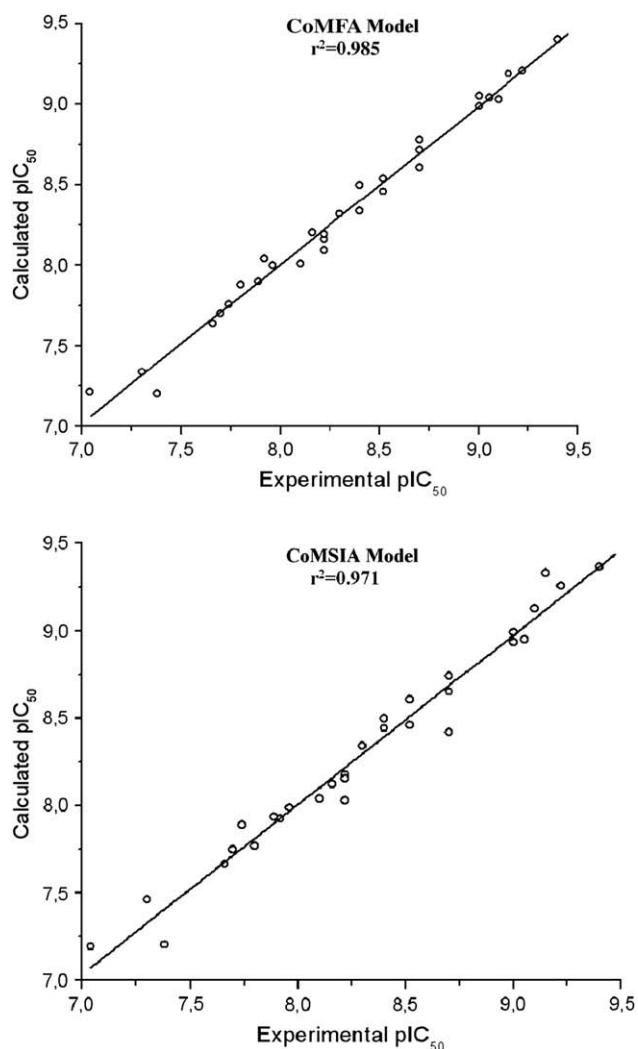| No.	CoMFA model			CoMSIA model		
	pIC$_{50}$ (experimental)	pIC$_{50}$ (predicted)	Residual	pIC$_{50}$ (experimental)	pIC$_{50}$ (predicted)	Residual
1	7.04	7.209	−0.17	7.04	7.195	−0.16
2	7.38	7.197	0.18	7.38	7.205	0.18
3	7.30	7.339	−0.04	7.30	7.463	−0.16
4	8.22	8.086	0.13	8.22	8.033	0.19
5	8.70	8.612	0.09	8.70	8.423	0.28
6	7.74	7.762	−0.02	7.74	7.89	−0.15
7	7.80	7.878	−0.08	7.80	7.769	0.03
8	7.66	7.641	0.02	7.66	7.664	0
9	7.96	8.002	−0.04	7.96	7.987	−0.03
10	7.92	8.035	−0.11	7.92	7.925	0
11	8.52	8.536	−0.02	8.52	8.609	−0.09
12	8.40	8.499	−0.1	8.40	8.499	−0.1
13	8.22	8.157	0.06	8.22	8.179	0.04
14	8.16	8.2	−0.04	8.16	8.122	0.04
15	9.00	8.988	0.01	9.00	8.937	0.06
16	8.70	8.784	−0.08	8.70	8.743	−0.04
17	8.10	8.006	0.09	8.10	8.041	0.06
18	9.15	9.188	−0.04	9.15	9.335	−0.19
19	8.52	8.457	0.06	8.52	8.461	0.06
20	8.40	8.338	0.06	8.40	8.444	−0.04
21	8.70	8.724	−0.02	8.70	8.653	0.05
22	7.89	7.897	−0.01	7.89	7.936	−0.05
23	8.22	8.189	0.03	8.22	8.156	0.06
24	9.05	9.041	0.01	9.05	8.954	0.1
25	9.00	9.053	−0.05	9.00	8.995	0
26	9.22	9.205	0.01	9.22	9.26	−0.04
27	9.40	9.403	0	9.40	9.368	0.03
28	9.10	9.026	0.07	9.10	9.132	−0.03
29	7.70	7.696	0	7.70	7.748	−0.05
30	8.30	8.321	−0.02	8.30	8.344	−0.04
	CoMFA residual average value: 0.06			CoMSIA residual average value: 0.08		

Fig. 4. Calculated pIC_{50} versus experimental pIC_{50} values for the 30 molecules of the training set obtained by PLS analysis using CoMFA and CoMSIA models.

Table 4
Experimental and predicted activities (pIC_{50}) of the test set molecules.

No.	CoMFA model			CoMSIA model		
	pIC_{50} (experimental)	pIC_{50} (predicted)	Residual	pIC_{50} (experimental)	pIC_{50} (predicted)	Residual
31	8.22	7.413	−0.807	8.22	8.256	0.036
32	7.89	7.385	−0.505	7.89	8.361	0.471
33	9.00	7.317	−1.683	9.00	8.334	−0.666
34	7.72	7.243	−0.477	7.72	8.294	0.574
35	8.40	7.352	−1.048	8.40	8.505	0.105
36	7.80	7.595	−0.205	7.80	8.485	0.685
37	8.00	8.121	0.121	8.00	9.358	1.358
38	8.52	7.985	−0.535	8.52	9.338	0.818
	CoMFA residual average value: 0.64			CoMSIA residual average value: 0.42		

the region A, close to R_2 group (Fig. 5). The CoMSIA steric and electrostatic contours were placed in an almost identical topographical location to that of the CoMFA model.

From the SAR in Table 1, emerges that the more active compounds contain an alkyl ether chain in the R_2 group (compounds **26–28**) while the less active molecules possess $-CONH_2$ or $-SO_2Me$ (compounds **1** and **3**). More specifically, in

compound **1** the carbonyl group is mapped near blue colored contour (negative charged groups disfavored) explaining the decreased activity. On the other side, in region C, blue (positive charged groups favored) and red colored (negative charged groups favored) contours are embedded into the yellow (bulky groups disfavored) fields in CoMFA and CoMSIA models, indicating that the orientation of the R_4 group may play a critical role in determining the bioaffinity as negatively and positively charged favorable polyhedra coexist in this region. Furthermore, from the SAR evaluation, it appears that the length of the tail of the chain affects activity, no matter if the terminal group is electropositive or electronegative (compounds **10–11** and **14–16**).

The green colored contours around the region B show that bulky groups enhance the binding affinity for the renin in both CoMFA and CoMSIA models (Figs. 5 and 6). Specifically, the presence of *i*-propyl group instead of methyl group on R_3 position leads in greater pIC_{50} values. Furthermore, in case of highly active compounds (e.g., compounds **18**, **24–28**) the *i*-propyl group falls within the green colored contour of the region B, while less active compounds such as **5–7**, **15** tend to locate the methyl group away from this contour and approache the sterically unfavorable yellow contour especially in the CoMSIA model (Fig. 6). Therefore, the replacement of *i*-propyl with a bulkier group such as *t*-butyl is expected to enhance further the renin binding affinities.

A second smaller green colored contour is presented in the region C (Fig. 6). This contour is larger and more visible in CoMSIA model while in CoMFA model is embedded in the blue colored contour. The green contour located in region C favors molecules with bulkier groups (Fig. 6). The groups like methyl or the geminal-dimethyl-substitution as in compounds **24** and **26** demonstrated higher activity compared to unsubstituted molecule **14**.

Three yellow colored contour maps were observed in both models; one is in region A and the other two in the region C, suggesting that less bulky groups at these positions favor activity. Most
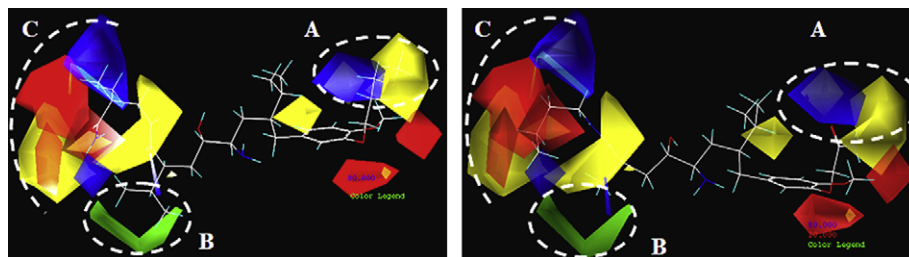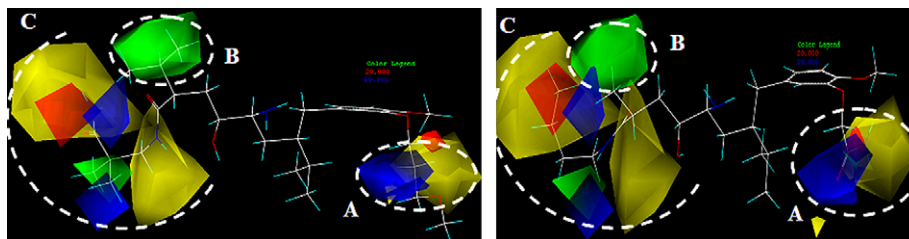

Fig. 5. CoMFA contour map of template compound **26** (highly active, on the left) and the respective analogue **1** (inactive analogue, on the right). Sterically favored areas are shown in green colored contours (contribution level of 80%). Sterically unfavored areas are shown in yellow colored contours (contribution level of 20%). Positive potential favored areas are shown in blue colored contours (contribution level of 80%). Positive potential unfavored areas are shown in red colored contours (contribution level of 20%). (Regions A–C show contour maps around R^2, R^3, R^4 groups of the ligands, respectively. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. CoMSIA contour map of template compound **26** (highly active, on the left) and the respective analogue **1** (inactive analogue, on the right). Sterically favored areas are shown in green colored contours (contribution level of 80%). Sterically unfavored areas are shown in yellow colored contours (contribution level of 20%). Positive potential favored areas are shown in blue colored contours (contribution level of 80%). Positive potential unfavored areas are shown in red colored contours (contribution level of 20%). (Regions A–C show contour maps around R^2, R^3, R^4 groups of the ligands, respectively. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

potent compounds have a methoxy group at the R_1 position (region A). This group is more favorable than the *tert*-butyl not only because a yellow contour is in the vicinity of R_1 but also because of its more favorable interactions with pocket S3. Furthermore, the yellow colored contours in region A indicate that there is a small available space in the active site for interactions. This is in correlation with previous published reports suggesting that R_2 group interacts with a narrow subsite (S3 sub-pocket) [18,19]. Large yellow colored contours on the region C (Figs. 5 and 6) suggest increase of bulkiness at the terminal segment of alkyl chain of R4 is unfavorable. This observation is confirmed by compounds **8** and **9**.

Since the crystal structure of renin is known, we found interesting to examine the stability of 3D QSAR model with the renin inhibitors target. As CoMFA model emerged with slightly better predictive ability, we superimposed the CoMFA contour maps with the active site of renin. The MOLCAD surface of active site was developed and displayed with lipophilic and electrostatic potential. In Fig. 7a the brown color shows lipophilic regions and blue color shows polar regions of active site. The CoMFA yellow contours appearing around A and C regions indicate that steric bulk substituents might have negative effect on the activity of the molecules. These zones approach to residues Gln13, Tyr14, Gln128, Ser76 and Ser35 which include hydrophilic groups. Similarly the green contour appearing around R3 group approaches the lipophilic region near Pro 292 and Leu 73 suggesting that bulkier group will increase the biological activity, an observation that is in agreement with CoMFA model. Fig. 7b shows the MOLCAD surface of active site with superimposed electrostatic contour maps of CoMFA model. At the electrostatic map shown in Fig. 7b, the red color shows the electropositive zone and purple color shows electronegative zone. Red contour appears around the positive charge density of renin while blue colored contour maps are oriented towards the electronegative groups of renin amino acids confirming the QSAR results.

The complementary information from docking studies is vital for the design of new molecules. The performed docking studies not only highlight the similarities between the QSAR and binding conformations but also provide the crucial interactions with the active site for enhanced activity. Fig. 8 represents the most stable binding conformation for **26** and **1**. In these docking studies both molecules are placed well in the active site of renin and demonstrate the following interactions. The –OH group of compound **26** forms H-bond both with oxygens of Asp32 (∼2.0 and 2.6 Å), the –NH₂ forms H-bond with both amino acids of Gly217 (∼2.2 Å) and Asp32 (∼2.1 Å), while the methoxy group of the side chain forms a H-bond with the distance of 1.9 Å with the –NH group of Tyr14. The primary –CONH forms H-bond with Ser76 (∼2.0 Å) and has a 3.4 Å distance from Gly34. The terminal –NH₂ interacts with H-bond with –C=O group of Arg74 (∼2.3 Å). In compound **1** the
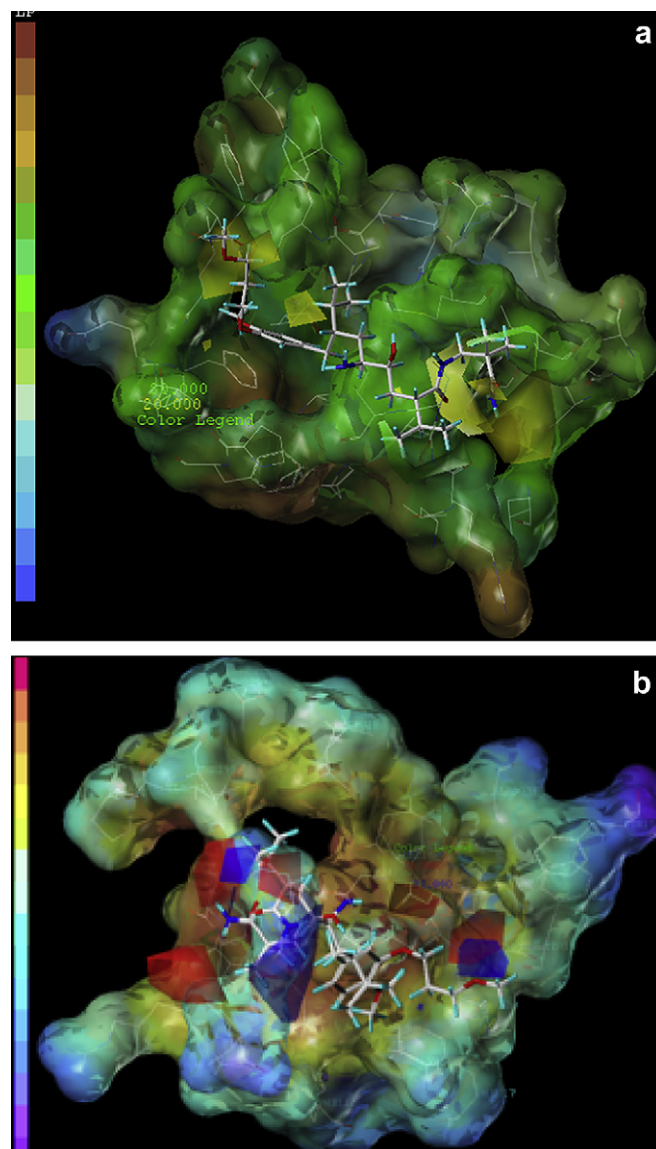
Fig. 7. a) The CoMFA steric contour maps within the active site of renin (lipophilic potential, MOLCAD lipophilic potential surface was calculated for the receptor with the Connolly method; brown color denotes the most lipophilic areas and blue color denotes the most polar areas), b) the CoMFA electrostatic contour maps within the active site of rennin (electrostatic potential, red colors denotes the most electropositive areas whereas purple colors denote the most electronegative fields). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
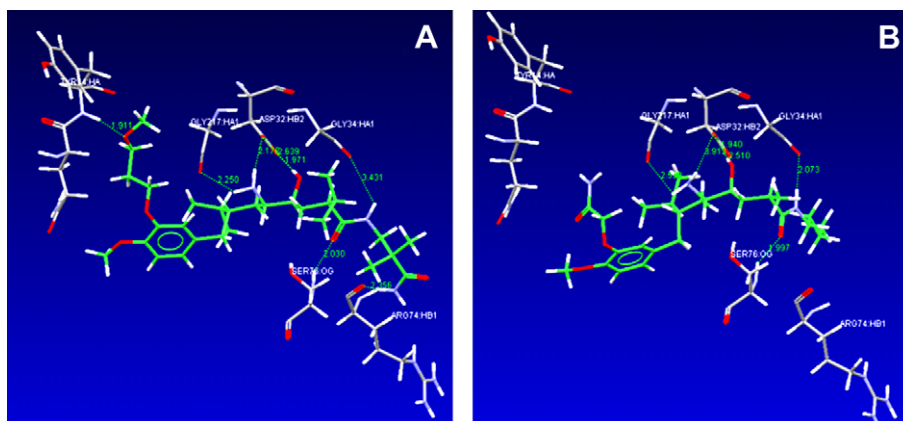
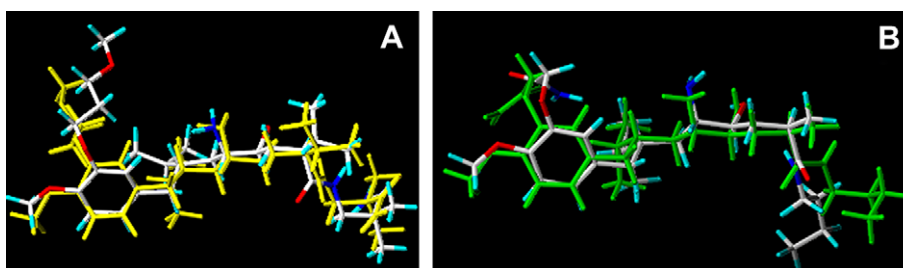**Fig. 8.** Docking interactions of compound **26**(A) and compound **1**(B) with active site residues.

Fig. 9. Superimposition of the QSAR conformation (designated with yellow and green colors respectively) with the binding conformation of compounds **26**(A) and **1**(B). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

–OH group forms H-bonds with oxygens of Asp32 ($\sim$1.9 and 2.5 Å). The –NH$_2$ group has a 2.5 Å distance from the –C=O group of Gly217 and $\sim$4.0 Å distance from the oxygen of Asp32. The –CONH group forms H-bonds with –C=O group of Gly34 ($\sim$2.1 Å) and –NH of Ser76 ($\sim$1.9 Å). There are no interactions with Arg74 as compound **1** does not have a terminal –CONH$_2$.

The GoldScore Fitness from docking simulations using the GOLD program [24] for compounds **26** and **1** were calculated to be 75.78 and 64.70 respectively, showing a better affinity for **26**. The major differences between the two compounds, that may explain their variation in their bioactivity are the following. The less active compound does not interact with the S$_3$ sub-pocket which is composed by the following amino acids: Gln13, Tyr14, Val30, Val120, Tyr155, Gly217, Ala218, Ser 219, Ala 303. More specifically compound **1** does not form H-bond with amino acid Tyr14, an interaction that is likely to be critical for bioactivity as it is previously reported [25,26]. The absence of this interaction can be attributed to the substitution of compound **1** in R$_2$ position. The short chain prevents compound **1** from approaching inside the active sub-pocket which lies only in its entrance. Additionally, the absence of interaction with Arg74 for compound **1**, due to its smaller dimensions can explain the inferior docking score and here upon the inferior biological activity.

As it is observed from the superimposition (Fig. 9) there is a general correlation between the conformations of compounds used in QSAR and the best docking binding conformation. For example, RMSD values between these two conformations were calculated to be 0.934 for compound **26** and 0.985 for compound **1**. However, it should be noted that some small differences have been observed between the conformations used in QSAR and binding orientations of the terminal R$_4$ group defined by N27–C28–C29–C30 atoms in compound **26**. The QSAR conformation is *gauche*

(53.1°) while the binding conformation is *trans* (175.6°). Additionally, there are small differences in the dihedral angles of alkyl ether chain between the QSAR and binding conformations of compound **26**. Regarding compound **1** the terminal chain of QSAR conformation has a *trans* conformation while the conformation of the bound conformation is *gauche*. Moreover the –NH$_2$ group in the R$_1$ position is oriented differently as it is derived from QSAR and binding conformations. Specifically, the –NH$_2$ group in QSAR conformation orients towards the west side while in binding conformation it orients towards the east side.

3. Conclusions

The present 3D QSAR study aims at targeting the following points: a) the comparison of the obtained results from CoMFA and CoMSIA models; b) the validation of the constructed models; c) the explanation of the different activities based on the constructed contour maps.

Although the molecules used in the present 3D QSAR study are flexible, the constructed models revealed statistical significance and good predictive abilities by using CoMFA and CoMSIA. These results confirm that these methodologies can be successfully applied in flexible molecules in agreement with our previous studies [27,28].

The two models CoMFA and CoMSIA did not record significant differences neither in their statistical analysis nor in the contour plots. In order to test the stability of the obtained PLS models for every conventional CoMFA and CoMSIA PLS run, bootstrapping was also performed. Obtained results support the reliability of the created models. Moreover, the contour plots of CoMFA and CoMSIA models, were located in the same space, but differed in volumes. These differences can be attributed to the energy

functions used to calculate the field values in CoMFA and CoM-SIA. In CoMFA the energy function is very sensitive with the changes in position. In CoMSIA, these fields are calculated using much smoother potentials which are not as steep as the Lennard–Jones and Coulombic functions and have a finite value even at the atomic positions. In test set, the predictions of the activities in CoMFA and CoMSIA provided statistically significant results. Average deviation from experimental pIC_{50} for both CoMFA and CoMSIA models were found less than one logarithmic unit.

3D QSAR studies are applied in order to address the following question. What molecular features can be modified to enhance bioactive properties of drug molecules? 3D QSAR studies are usually applied when the receptor is unknown. However, in our study, the X-ray structure of renin is known, therefore, QSAR studies are combined with docking results. This combination as it is shown in our previous studies is very effective to produce novel potential drugs [29,30].

From the GoldScore fitness function it cannot be extracted total reliance on pIC_{50} values (Table 1). This is not a surprising result as the docking programme mainly scores the interactions with the amino acids of the active site and the compared compounds have small differences in their structure. However, the difference in the activity can be attributed to the fact that some of the studied renin inhibitors (e.g., compound **1** is not placed deeply in the S3 subpocket), while in others (compounds **2**, **6–8**, **13**, **17–21**, **24**, **29**) an interaction is not observed with the Tyr14 amino acid of S3 subpocket, that has been characterized as essential for high binding affinity.

The correlation of the results obtained from docking and QSAR studies lead to better understanding of the structural requirements for enhanced activity. The obtained results can be used as a guideline to design and predict new and more potent renin inhibitors.

4. Computational methods

3D QSAR/CoMFA and CoMSIA studies were performed using Sybyl molecular modeling package [23], while docking studies were performed using the GOLD program [24].

4.1. Data set and molecular modeling

The thirty compounds used in this study are originated from the work of Goschke et al. [18] and Maibaum et al. [19] and are derivatives of aliskiren. The selection of the training set was done based on structural diversity and wide range of activity. Crystal structure of aliskiren is known and it was used to build the basic skeleton of all the molecules in data set. The structures of the compounds were designed in Sybyl6.8 molecular modeling package and Gasteiger–Huckel charges [31] were assigned to the atoms of the compounds. All the compounds were subjected to minimization until converged to a maximum derivative of 0.001 $kcal^{-1} A^{-1}$, using Tripos force field [23].

To obtain reliable results, the total compounds were divided into two sets, the training set consisting of 30 compounds and the test set consisting of 8 compounds. In the 3D QSAR study the renin inhibitory activities of the studied inhibitors were expressed in IC_{50} values and converted into pIC_{50} values using the equation $pIC_{50} = -\log IC_{50}$. The IC_{50} values used in the training set has 230-fold difference between the lowest and highest active compounds. In the selection of test set compounds, similar range of diversity with training set has been taking into account.

4.2. CoMFA and CoMSIA settings

CoMFA was performed using the QSAR option of Sybyl. The steric and electrostatic field energies were calculated using the Lennard–Jones and the Coulomb potentials, respectively, with a $1/r$ distance-dependent dielectric constant in all intersections of a regularly spaced (0.2 nm) grid [30]. An sp^3 carbon atom with a radius of 1.53 Å and a charge of $+1.0$ was used as a probe to calculate the steric and electrostatic energies between the probe and the molecules using the Tripos force field [23]. The truncation for both the steric and the electrostatic energies were set to 30 kcal mol^{-1}. This indicates that any steric or electrostatic field value that exceeds this value will be replaced with 30 kcal mol^{-1}, thus making a plateau of the fields close to the center of any atom.

CoMSIA was performed using the QSAR option of Sybyl. An sp^3 carbon atom with a radius of 1.53 Å and a charge of $+1.0$ was used as the probe to calculate the CoMSIA similarity indices. Steric and electrostatic similarity indices were evaluated at the intersections of a similar grid using the same probe atom according to the standard implementation of CoMFA in Sybyl. The similarity indices between the compounds and the probe atom are calculated according to:

$$A_{F,k}^{q}(j) = -\sum_{i=1}^{n} w_{\text{probe},k} w_{ik} e^{-\alpha r_{iq}^2}$$

where A is the similarity index at the grid point q, summed over all atoms i of the molecule j under investigation; $w_{\text{probe},k}$ is the probe atom; w_{ik} is the actual value of the physicochemical property k of atom i; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule; and R is the attenuation factor [32]. The default value of attenuation factor α was set to 0.3. Larger values of α will result in a steeper Gaussian function and increasing attenuation of the distance-dependent effects of molecular similarity. On the other hand, reducing α to smaller values will result in a probe atom detecting molecular similarity of its neighborhood more globally. The optimal value of α is between 0.2 and 0.4 [22,33].

4.3. CoMFA and CoMSIA partial least-squares (PLS) analysis and validation

The initial PLS analysis was performed using the "leave-one-out" cross-validation method for all 3D QSAR analyses. A minimum column filtering value of 2.00 kcal mol^{-1} was set to improve the signal-to-noise ratio by omitting those grid points whose energy variation was below this threshold. In both CoMFA and CoMSIA analyses, descriptors were treated as independent variables, whereas the pIC50 values were treated as dependent variables in the PLS regression analyses to derive the 3D QSAR models. The final model (noncross-validated conventional analysis) was developed from the model with the highest r_{cv}^2, and the optimum number of components was set to equal, that yielding the highest r_{cv}^2. The noncross-validated models were assessed by the conventional correlation coefficient r^2, standard error of prediction, and F values. For the creation of the CoMFA field, "CoMFA standard" scaling was selected, while in the case of CoMSIA, the "none" option was selected in the Sybyl.

To obtain confidence limits and test the stability of obtained PLS models, for every conventional CoMFA and CoMSIA PLS run, bootstrapping was also performed (100 runs, column filtering: 2.00 kcal mol^{-1}). The idea is to simulate a statistical sampling procedure by assuming that the original data set is the true population and generating many new data sets from it. These new data sets (called bootstrap samplings) are of the same size as the original

data set and are obtained by randomly choosing samples (rows) from the original data, with repeated selection of the same row being allowed. The statistical calculation is performed on each of these bootstrap samplings, with new values being calculated for each of the parameters to be estimated. The difference between the parameters calculated from the original data set and the average of the parameters calculated from the many bootstrap samplings is a measure of the bias of the original calculation.

4.4. Molecular docking

The binding interactions and the active conformations are derived from the GOLD program, using the genetic algorithm [24]. The energy functions of the interactions are partly based on the conformational and non-bonded interactions. The scoring function used was GoldScore. For the docking studies, the crystal structure of renin with aliskiren (2v0z, pdb code) was downloaded from protein data bank. From the crystal structure, the inhibitor and water molecules were removed. An active site of 10 Å around the docked inhibitor was created. The maximum number of generic algorithm runs was set to 20 for each compound. The default generic algorithm parameters were selected (100 population size, 5 number of islands, 100,000 number of generic operations and 2 for the niche size).

Acknowledgement

References

[1] C. Werner, M. Baumhakel, K.K. Teo, R. Schmieder, J. Mann, T. Unger, S. Yusuf, M. Bohm, Clin. Res. Cardiol. 97 (2008) 418–431.
[2] S. Iravanian, S.C. Duley, Heart Rhythm 5 (6) (2008) S12–S17.
[3] A.H. Gradman, R. Kad, J. Am. Coll. Cardiol. 51 (5) (2008) 519–528.
[4] J.A. Staessen, Y. Li, T. Richart, Lancet 368 (9545) (2006) 1449–1456.
[5] C.M. Tice, Annu. Rep. Med. Chem. 41 (2006) 155–167.
[6] J.C. Hershey, B. Steiner, W. Fischli, G. Feuerstein, Drug Discov. Today Ther. Strateg. 2 (3) (2005) 181–185.
[7] N.D.L. Fisher, N.K. Hollenberg, J. Am. Soc Nephr. 16 (2005) 592–599.
[8] A. Ferro, R. Gilbert, H. Krum, Int. J. Clin. Pract. 60 (5) (2006) 577–581.
[9] J.M. Wood, J. Maibaum, J. Rahuel, M.G. Grütter, N.C. Cohen, et al., Biochem. Biophys. Res. Commun. 308 (2003) 698–705.
[10] M. Azizi, R. Webb, J. Nussberger, N.K. Hollenberg, J. Hypertens. 24 (2006) 243–256.
[11] K.K. Daugherty, Am. J. Health-System Pharm. 65 (14) (2008) 1323–1332.
[12] C. Jensen, P. Herold, H.R. Brunner, Nat. Rev. Drug Discov. 7 (2008) 399–410.
[13] N.C. Cohen, Chem. Biol. Drug Des. 70 (2007) 557–565.
[14] A.R. Sialecki, K. Hayakawa, M. Fujinaga, M. Murphy, M. Fraser, A. Muir, C. Carilli, J.A. Lewicki, J.D. Baxter, M.N.G. James, Science 243 (1989) 1346–1351.
[15] D.E. Epps, J. Cheney, H. Schostarez, T.K. Sawyer, M. Prairie, W.C. Krueger, F. Mandel, J. Med. Chem. 33 (8) (1990) 2080–2086.
[16] J.R. Rahuel, V. Rasetti, J. Maibaum, H. Rueger, R. Goschke, et al., Chem. Biol. 7 (2000) 493–504.
[17] J.L. Poll, JMPC 13 (8) (2007) S21–S33.
[18] R. Göschke, S. Stutz, V. Rasetti, N.-C. Cohen, J. Rahuel Rigollier, et al., J. Med. Chem. 50 (2007) 4818–4831.
[19] J. Maibaum, S. Stutz, R. Göschke, P. Rigollier, Y. Yamaguchi, et al., J. Med. Chem. 50 (2007) 4832–4844.
[20] P.G. Gupta, Mini Rev. Med. Chem. 3 (2003) 315–321.
[21] P.G. Gupta, J.K. Gupta, A.N. Nagappa, V. Jagannathan, D. Gangwal, Drug Des. Deliv. 5 (2) (1989) 73–81.
[22] S. Durdagi, A. Kapou, T. Kourouli, T. Andreou, S.P. Nikas, et al., J. Med. Chem. 50 (2007) 2875–2885.
[23] Sybyl, v. 6.8, Molecular Modeling Software Packages, Tripos Inc., St Louis, MO 63144, 2001.
[24] Gold, Version 2.1, Cambridge Crystallographic Data Center, Cambridge, UK, 1997.
[25] D.D. Holsworth, C. Cai, M. X-Cheng, W.L. Cody, D.M. Downing, et al., Bioorg. Med. Chem. Lett. 16 (2006) 2500–2504.
[26] H.P. Märki, A. Bingeli, B. Bittner, V. Bohner-Lang, V. Breu, et al., Il Farmaco 56 (1–2) (2001) 21–27.
[27] A. Kapou, N. Avlonitis, A. Detsi, M. Koufaki, T. Calogeropoulou, et al., Bioorg. Med. Chem. 15 (2007) 1252–1265.
[28] S. Durdagi, T. Mavromoustakos, N. Chronakis, M.G. Papadopoulos, Bioorg. Med. Chem. 16 (2008) 9957–9974.
[29] S. Durdagi, M.G. Papadopoulos, D.P. Papahatjis, T. Mavromoustakos, Bioorg. Med. Chem. Lett. 17 (24) (2007) 6754–6763.
[30] S. Durdagi, C. Koukoulitsa, A. Kapou, T. Kourouli, T. Andreou, et al., Drugs Future 32 (Suppl. A) (2007) 79.
[31] J. Gasteiger, M. Marsili, Tetrahedron 36 (1980) 3219–3228.
[32] A.R. Leach, V.J. Gillet, An Introduction to Chemoinformatics. Kluwer Academic Publishers, Dordrecht, 2003.
[33] M. Bohm, J. Sturzebecher, G. Klebe, J. Med. Chem. 42 (1999) 458–477.