ELSEVIER

Combined 3D QSAR and molecular docking studies to reveal novel cannabinoid ligands with optimum binding activity

Serdar Durdagi,[a,b] Manthos G. Papadopoulos,[a]
Demetris P. Papahatjis[a] and Thomas Mavromoustakos[a,c,*]

[a]*Institute of Organic and Pharmaceutical Chemistry, The National Hellenic Research Foundation,
48 Vas. Constantinou Avenue, 11635 Athens, Greece*
[b]*Department of Biology Chemistry and Pharmacy, Freie Universität Berlin, Takustr. 3, 14195 Berlin, Germany*
[c]*Department of Chemistry, University of Athens, Zographou, 15784 Athens, Greece*

Abstract—The combination of NMR spectroscopy and molecular modeling studies provided the putative bioactive conformation for the analgesic cannabinoid (CB) ligand (−)-2-(6a,7,10,10a-tetrahydro-6,6,9-trimethylhydroxy-6*H*-dibenzo[*b,d*]pyranyl)-2-hexyl 1,3-dithiolane which served as a template in reported three-dimensional quantitative structure–activity relationship (3D QSAR) studies [Durdagi et al., *J. Med. Chem.* **2007**, *50*, 2875]. The reported 3D models of the CB1 receptor allowed us to construct a new 3D QSAR model based on theoretical calculations and molecular docking studies. Statistical comparison of the constructed two 3D QSAR studies showed the improvement of the new model. In addition, the new model can explain more effectively the experimental data and thus it can serve more efficiently in the rational drug design of pharmacologically optimized CB analogues.
© 2007 Elsevier Ltd. All rights reserved.

In a previous study,[1] we reported the three-dimensional quantitative structure–activity relationship (3D QSAR) results of novel cannabinoid (CB) analogues using conformer **a** (left in Fig. 1) of (−)-2-(6a,7,10,10a-tetrahydro-6,6,9-trimethylhydroxy-6*H*-dibenzo[*b,d*]pyranyl)-2-hexyl 1,3-dithiolane (**12** in Table 1) as a template ligand. Determination of the conformation of the template compound is one of the critical steps in 3D QSAR studies. In the above-mentioned study, the conformation of **12** used as a template was derived by applying a combination of NMR and molecular modeling studies.[2]

The knowledge of the receptor structure is not a prerequisite for 3D QSAR analysis, however, the availability of its crystal structure or 3D model facilitates the structure alignment, and can provide statistically more reliable models.[3,4] 3D models of the CB1 receptor were constructed by Shim et al.[5] and Tuccinardi et al.[6] with a molecular modeling procedure using the X-ray structure of bovine rhodopsin[7] as the initial template and

taking into account the available site-directed mutagenesis data. These groups studied different CB classes, however, they found them to interact at an active site with similar homologies. Although the more recent 3D model of the CB1 receptor was used for the molecular docking studies, the active site residues are determined considering both models mentioned above. The active site in the docking runs included all atoms within a radius of 5.5 Å around the critical amino acids (Phe174, Leu190, Lys192, Leu193, Gly195, Val196, Thr197, Phe200, Thr201, Pro251, Trp356, Leu359, Ser383, Cys386, and Leu387).

The receptor model obtained by Tuccinardi et al.[6] was complexed with the low energy conformers of high affinity CB ligand **12**. Low energy conformers of **12** were obtained using a Monte Carlo (MC) conformational search analysis. MC simulations were performed using QUANTA/CHARMm (version 4.2 from MSI) software[8,9] to investigate the complete conformational space of **12**. The application of MC analysis, which allows full angular window specification and random change of all flexible dihedral angles responsible for the flexibility, generated 1000 conformers of **12**. Generated conformers were consequently subjected to the full geometry optimi-
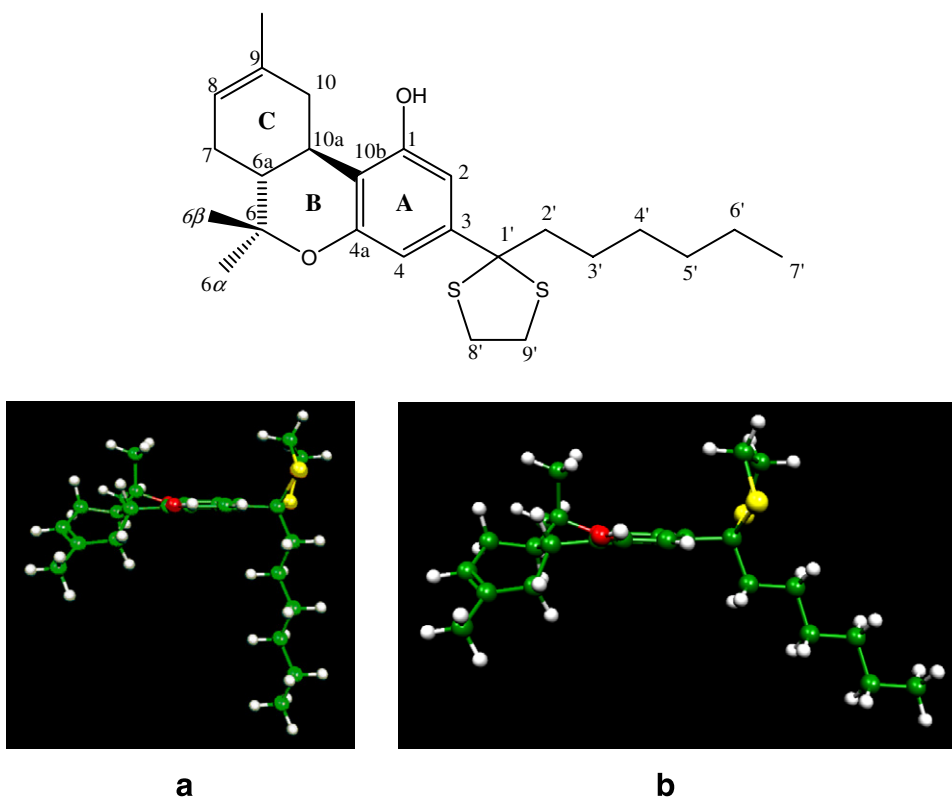
Figure 1. Molecular structure of **12** and its putative bioactive conformation **a** derived from a combination of molecular modeling and experimental NMR spectroscopy and conformation **b** derived from a combination of molecular modeling and molecular docking studies.

zation and these conformers are grouped into eight clusters. From each cluster, the lowest-energy conformer was selected and subjected to docking with the CB1 receptor.

The CB1 receptor has two available sites (S1 and S2) for accommodating CB ligands (Fig. 2). (i) S1 site: contains a cavity with a ~7 Å depth and accommodates the alkyl chain segment of the CB analogue. Our findings are in accordance with previous observations[10] which show that extension of five-carbon atom chain of tetrahydrocannabinol (THC) by one or two carbon atoms (optimum alkyl chain length is ~7 Å) improves binding, while further extension (>7 Å) is detrimental due to steric hindrance and (ii) S2 site: contains a cavity with a ~10 Å depth and accommodates the alkyl chain segment of the CB ligand.

Flexible docking has been employed to the lowest-energy conformers of **12** using the FlexX docking algorithm of SYBYL molecular modeling package.[11] FlexX is a docking method that uses an efficient incremental construction algorithm in order to optimize the interaction between a flexible ligand and rigid binding site residues of a receptor. Population analyses of docking results showed that conformer **b** (right in Fig. 1) had the highest percentage of conformation in the active site of receptor. Thus, we examined conformer **b** of compound **12** as a template ligand at the 3D QSAR analysis.

After acquiring the highest percentage of conformation (conformer **b**) in the active sites of the CB1 receptor,

we applied 3D QSAR/comparative molecular similarity indices analysis (CoMSIA)[12] studies using conformer **b** of template ligand **12**. The aim of applying the 3D QSAR/CoMSIA method is to derive indirect binding information from the correlation between the biological activity of a training set of compounds and their 3D structures.[1,13] The importance of steric and electrostatic characteristics is revealed by aligning structurally similar analogues using pharmacophoric features as structural superimposition guides.[10,13] CoMSIA calculates similarity indices around the molecules, with the similarity expressed in terms of different physicochemical properties, such as steric occupancy, partial atomic charges, local hydrophobicity, and hydrogen bond donor and acceptor properties.[12–14]

Several variations in the alignment schemes by superimposing the structurally similar pharmacophoric features are considered. C_1, C_2, C_3, C_4, C_{4a}, C_{6a}, C_7, C_{10}, C_{10a}, C_{10b} and the oxygen atoms in the template conformers of **12** are selected for the structural superimposition processes.[1] The alignment of the molecules was based on an atom-by-atom superimposition of selected atoms, which are common in all compounds in the training set. The criteria applied for the selection were: (i) overlap of the putative biologically relevant pharmacophore groups (with minimum RMS) and (ii) form of statistically significant 3D QSAR/CoMSIA models.[1] Figure 3 shows the superimpositions of CB analogues used as a training set to construct 3D QSAR/CoMSIA models based on the conformers **a** and **b** of template ligand **12**, respectively.

Table 1. Molecular structures and binding affinity K_i values for CB analogues used in analyses of CoMSIA models[15–19]
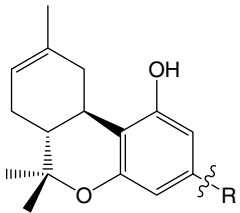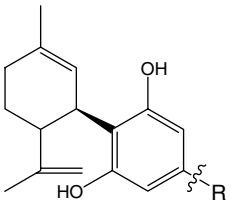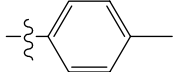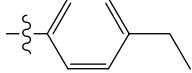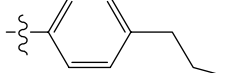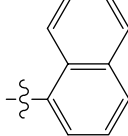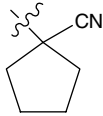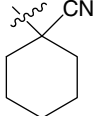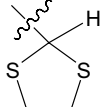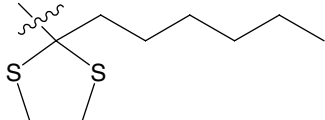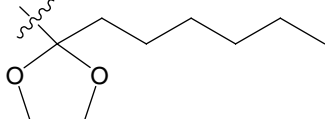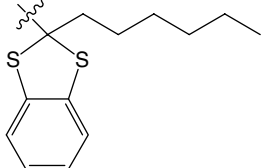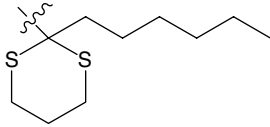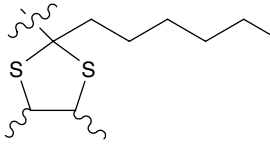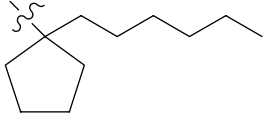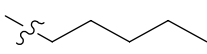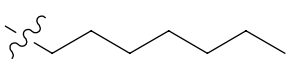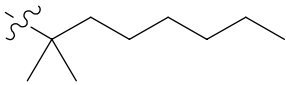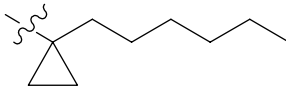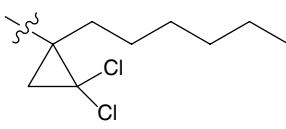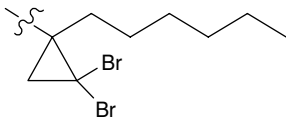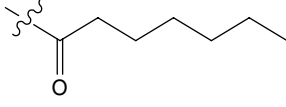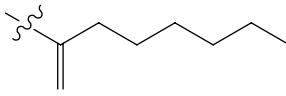

Compound	R	K_i for CB1 receptor (nM)	Compound	K_i for CB1 receptor (nM)
1		95.49	**2**	638.1
3		119.0		
4		57.77		
5		11.73	**6**	753.5
7		27.90	**8**	255.0
9		8.26	**10**	319.0
11		168.0		
12		0.32	**13**	136.0
14		0.52		
15		56.90		

Table 1 (*continued*)

Compound	R	K_i for CB1 receptor (nM)	Compound	K_i for CB1 receptor (nM)
16		1.80		
17		32.30		
18		0.45		
19		47.60	**20**	1265.0
21		22.00		
22		0.83		
23		0.44	**24**	58.68
25		1.27	**26**	666.4
27		0.71	**28**	189.0
29		21.70		
30		2.17		

In order to build 3D QSAR/CoMSIA models for the binding affinity (K_i) at the CB1 receptor, a set of 30 Δ^8-THC and cannabidiol (CBD) analogues[15–19] (Table 1) were analyzed using the QSAR option of SYBYL.[11] The logarithmic values of $1/K_i$ (pK_i) were used in the 3D QSAR correlations, as they are related to changes in the free energy of binding. Cross-validated partial least-squares (PLS) analyses were applied for both models (using the template ligand as conformers **a**

and **b** of **12**). Steric and electrostatic field columns of CoMSIA are created automatically by SYBYL. The same CoMSIA settings, PLS analyses, and validations have been applied as in a previously reported study.[1] A very high correlation was observed for both models as it is demonstrated by the high values of r^2 (Table 2). Additionally, the credibility of the models is proved by the high values of cross-validated $r^2(r_{cv}^2)$ (Table 2).
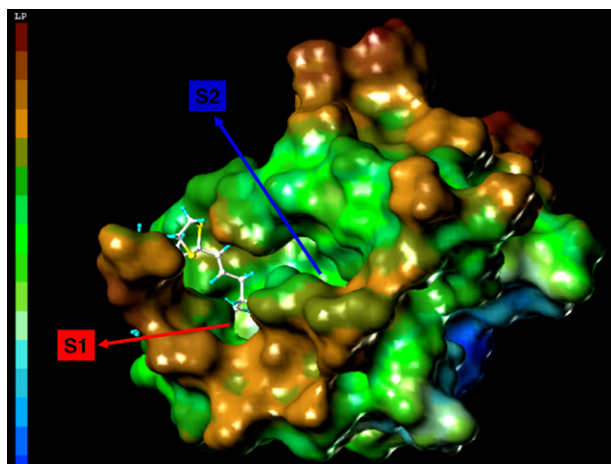
Figure 2. Two cavities S1 and S2 are observed at the active site of the CB1 receptor: (i) S1 site: contains a cavity with a ~7 Å depth and accommodates alkyl chain segment of CB ligand. (ii) S2 site: contains a cavity with a ~10 Å depth and accommodates alkyl chain segment of CB. (LP, lipophilicity map; lipophilicity decreases from top to bottom.)

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.[1,20] The statistical calculation is performed on each of these bootstrap samplings, with new values being calculated for each of the parameters to be estimated.[1,20] The results obtained support the reliabi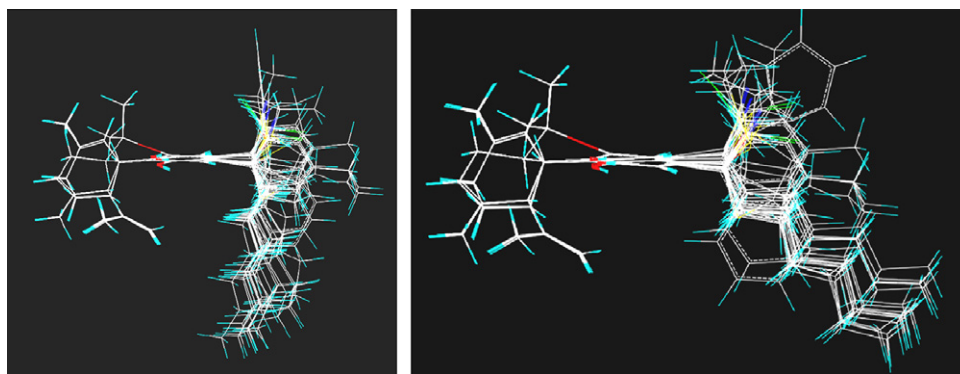lity of the models (Table 2). Statistical validation tests are compared for the models; the model constructed on the conformer **b** of **12** has better statistical values than the model constructed on the conformer **a** of **12**.

Table 3 summarizes the experimental (observed) and CoMSIA-predicted pK_i results for the binding affinities at the CB1 receptor. Figure 4 shows the relationship between the 3D QSAR/CoMSIA predicted and experimental pK_i values of the noncross-validated analyses for the constructed models based on conformers **a** (left) and **b** (right) of **12**, respectively. The linearity of the plot concerning conformer **b** was better than the linearity of the plot concerning conformer **a**. Both plots showed good correlations for the constructed models.

Figure 5 shows the steric–electrostatic CoMSIA contour maps of **12** (on the left) and its corresponding CBD analogue **13** (on the right) for the CB1 receptor using as template ligand the conformer **a** of **12**, whereas Figure 6 shows the steric–electrostatic CoMSIA contour maps of **12** (on the left) and its corresponding CBD analogue **13** (on the right) for the CB1 receptor using as template

In order to obtain the confidence limit and test the stability of the obtained PLS models, for every conventional CoMSIA PLS run, bootstrapping was also performed (100 runs, column filtering: 2.00 kcal/mol). 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.[1,20]

Figure 3. Structural alignments of the compounds in the training set for constructing 3D QSAR/CoMSIA models based on conformers **a** (on the left) and **b** (on the right) of the template ligand **12**, respectively.

Table 2. PLS analyses for the CB1 receptor using the CoMSIA models based on compound **12** as template

	CB1 model-previous study[1] (template ligand **12**-conformer **a**)	CB1 model (template ligand **12**-conformer **b**)
Number of compounds in the training set	30	30
r_{cv}^2	0.746	0.764
r^2	0.944	0.953
Standard error of estimate	0.296	0.272
F	65.031	77.600
Relative contributions of steric/electrostatic fields	0.890:0.110	0.890:0.110
$r_{bootstrapping}^2$	0.971	0.974
Number of optimal components	6	6

Table 3. Summary of experimental (observed) and CoMSIA-predicted pK_i results of training set for the binding affinity at the CB1 receptor

Compound	pK_i (observed)	CB1 CoMSIA model	
		pK_i (predicted) (template ligand **12**-conformer **a**)	pK_i (predicted) (template ligand **12**-conformer **b**)
1	7.02	7.19	7.38
2	6.20	6.12	5.93
3	6.92	7.03	7.16
4	7.24	6.98	7.11
5	7.93	7.54	7.78
6	6.12	6.40	6.15
7	7.55	7.63	7.66
8	6.59	6.51	6.41
9	8.08	7.83	7.93
10	6.50	6.51	6.33
11	6.77	6.91	6.83
12	9.49	9.00	9.45
13	6.87	6.89	6.86
14	9.28	9.20	8.99
15	7.24	7.15	7.03
16	8.74	8.45	8.80
17	7.49	8.03	7.63
18	9.35	9.82	9.61
19	7.32	7.33	7.25
20	5.90	5.98	5.88
21	7.66	7.95	7.95
22	9.08	9.05	9.20
23	9.36	9.13	8.79
24	7.23	6.74	7.03
25	8.90	9.19	9.06
26	6.18	6.65	6.85
27	9.15	9.10	9.05
28	6.72	6.61	6.81
29	7.66	7.82	7.89
30	8.66	8.55	8.39

groups near the yellow colored contours; more positive charge near the blue colored contours, and more negative charge near the red colored contours.

Three general conclusions could be drawn from the characteristics of derived 3D contour maps of CoMSIA models using both conformations of template ligand **12**:

1. Steric effects determine the binding affinity. The relative contributions of steric fields are larger than those of the electrostatic fields.
2. The orientation of the C3-alkyl chain plays a crucial role in determining the biological activity. The green colored contours along the left side of the end of the alkyl chain (corresponding to shown snapshot contour plots, Figs. 5 and 6) show that bulky groups enhance the binding affinity, whereas bulky groups in the right sides of the C3-alkyl chain of analogues lead to decreased binding affinity.
3. Because of the structural differences of Δ^8-THC and CBD derivatives at the cyclic ring segment, these groups have different pharmacophoric requirements for their receptors in these regions. While sterically unfavorable areas are located on the methyl or propenyl groups of CBD analogues, these unfavorable regions are located at the vicinity of the tricyclic segment of Δ^8-THC analogues (Figs. 5 and 6). Therefore, Δ^8-THC analogues have higher binding affinities than their corresponding CBD analogues.

The conformers **a** and **b** of **12** used as a template compound in CoMSIA show similarities and differences in contour maps. Their similarities are reflected in the same regions that contour levels of identical color cover. However, close observation reveals significant differences in their shape and extent of covering of the contour regions. The conformational differences of conformers **a** and **b** are localized in the alkyl chain. The early SAR studies have been reviewed comprehensively by Thakur et al.,[21] Khanolkar et al.,[22] Razdan,[23] and Makriyannis et al.[24] Our results confirm the earlier literature reports that the lipophilic alkyl chain plays crucial role in determining cannabimimetic activity for the CB1 receptor. Thus, the differences of contour maps
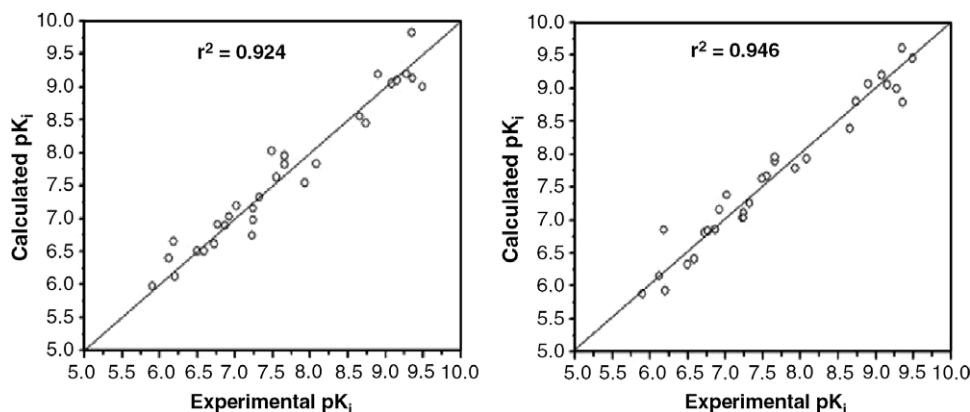
ligand the conformer **b** of **12**. The individual contributions from the steric and electrostatic favored and disfavored levels are fixed at 80% and 20%, respectively. The CoMSIA contours of the steric maps are shown in yellow and green colors, and those of the electrostatic contour maps are shown in red and blue colors. Greater values of 'bioactive measurement' are collected with: bulky groups near the green colored contours; not bulky


Figure 4. Plots of corresponding 3D QSAR/CoMSIA predicted and experimental values of binding affinity (given as pK_i) of CB analogues in the training set at the CB1 receptor for the constructed models based on conformers **a** (on the left) and **b** (on the right) of **12**, respectively.
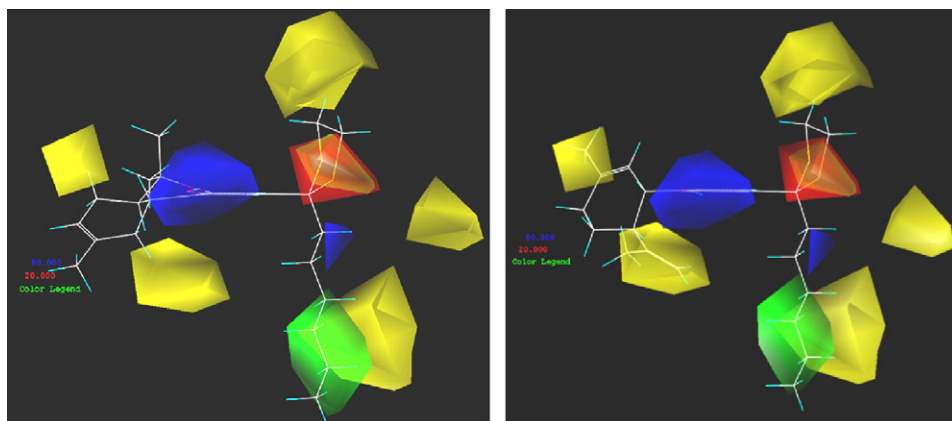
Figure 5. CoMSIA contour maps of **12** (on the left) and its corresponding CBD analogue **13** (on the right). Conformer **a** is used as template. Sterically favored areas are shown in green color (contribution level of 80%). Sterically unfavored areas are shown in yellow color (contribution level of 20%). Positive potential favored areas are shown in blue color (contribution level of 80%). Positive potential unfavored areas are shown in red color (contribution level of 20%).
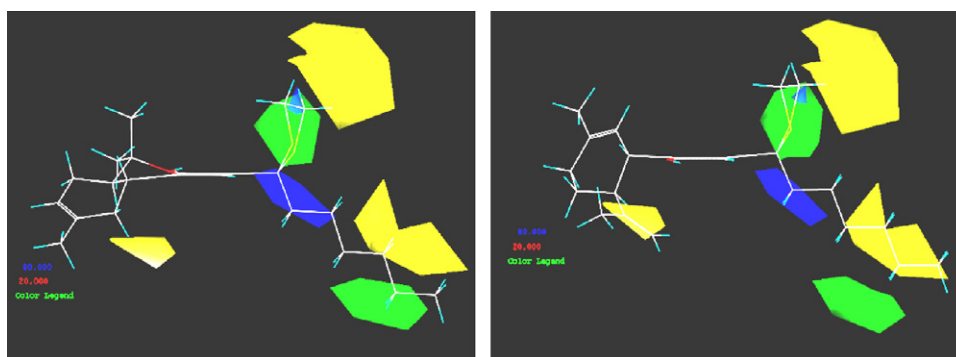
Figure 6. CoMSIA contour maps of **12** (on the left) and its corresponding CBD analogue **13** (on the right). Conformer **b** is used as template. Sterically favored areas are shown in green color (contribution level of 80%). Sterically unfavored areas are shown in yellow color (contribution level of 20%). Positive potential favored areas are shown in blue color (contribution level of 80%). Positive potential unfavored areas are shown in red color (contribution level of 20%).

at alkyl chain are important for the interpretation of pharmacophore groups that affect the binding affinity. When conformer **a** is used as a template, both THC and CBD analogues have green colored contour (depicts sterically favorable groups) at the tail of alkyl chain (Fig. 5). However, if conformer **b** is used as a template compound, then at the tail of alkyl chain only THC fits green colored contour (left in Fig. 6). CBD analogues do not fit green colored contours but they fit yellow colored contours (depicts sterically unfavorable groups) (right in Fig. 6). These important observations are obtained only by the model that was constructed on conformer **b** of **12**. The contour plots at the tail of alkyl chain derived by the model that was constructed on conformer **b** of **12** demonstrate the better binding affinity of THC analogues than the corresponding CBD analogues.

In addition, to validate the higher predictive ability of conformer **b** of the template ligand **12**, 10 other Δ^8-THC analogues (Table 4) have been added to the training set and CoMSIA models have been reconstructed (binding affinities have been taken from reported values

in the literature[21,22]). The same CoMSIA settings and PLS analyses have been performed for the reconstructed CoMSIA models. The same atoms in the template conformers of **12** have been selected for the structural superimposition processes. Results did not significantly modify the initially obtained models. Reconstructed 3D QSAR/CoMSIA models for the binding affinities to the CB1 receptor have a very good cross-validated correlation. Reconstructed models validate the initially obtained results: the model based on conformer **b** of **12** shows better statistical results than the model based on conformer **a** of **12** (Table 5).

Conformer **a** of **12** fits the S1 site of the receptor, whereas conformer **b** of **12** fits the S2 site of the receptor. More clearly seen in Figure 2, the S1 site has more lipophilic character than the S2 site. Unsaturation of the alkyl chain leads its orientation toward the S2 site. For this reason an analogue of **12** was designed possessing four unsaturated bonds which were directed specifically to the S2 cavity. The proposed molecule will be synthesized and tested for its biological activity in order to validate our rational design. Depending on the observed

S. Durdagi et al. / Bioorg. Med. Chem. Lett. 17 (2007) 6754–6763

6761

Table 4. Molecular structures and binding affinity K_i values of CB analogues that were added to training set to construct new CoMSIA models for validating the higher predictive ability of the conformer **b** of template ligand **12** (binding affinities have been taken from reported values in the literature[21,22])

| Compound	R	K_i for CB1 receptor (nM)
31		10.90
32		3.90
33		2.70
34		141.0
35		19.00
36		0.43
37		1.75
38		0.20
39		1.50
40		4.50

activity we will be able to differentiate if optimum activity is induced by S1 or S2 sites. These observations may help open new avenues to synthetic chemists for synthesizing novel compounds.

In conclusion, we applied a novel approach to generate new structures aiding in the rational drug design. This approach is based on the combination of theoretical calculations, molecular docking, and 3D QSAR studies.

Table 5. PLS analyses for the CB1 receptor using the re-obtained CoMSIA models based on compound **12** as template

	Re-obtained CB1 model (template ligand **12**-conformer **a**)	Re-obtained CB1 model (template ligand **12**-conformer **b**)
Number of compounds in the training set	40	40
r^2_{cv}	0.700	0.745
r^2	0.924	0.946
Standard error of estimate	0.331	0.279
F	66.544	95.977
Relative contributions of steric/electrostatic fields	0.917: 0.083	0.923: 0.077
$r^2_{bootstrapping}$	0.957	0.963
Number of optimal components	6	6
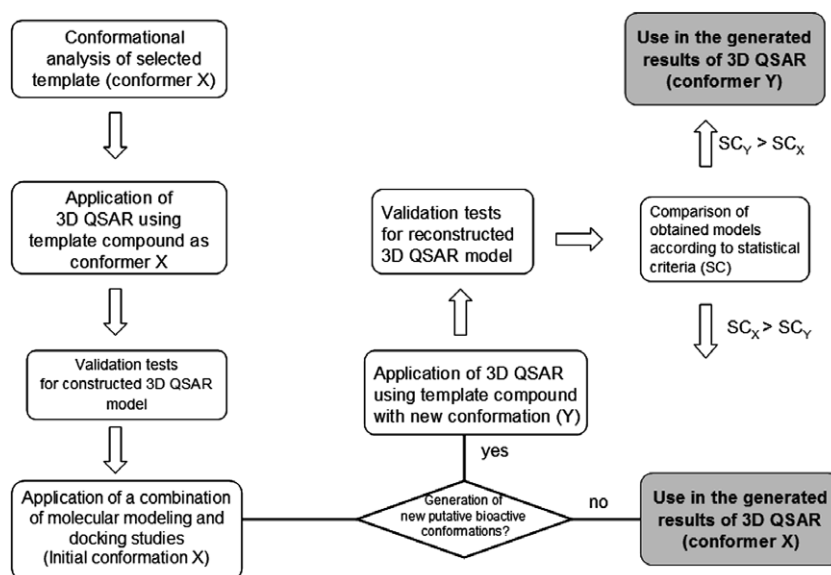
Figure 7. Flow chart showing the comparative steps used in the analysis of data derived through a combination of molecular modeling and NMR experiments with molecular modeling and docking results.

Such an approach appears to be superior to the 3D QSAR results, previously reported by our group, using a combination of theoretical calculations and NMR spectroscopy. Figure 7 describes the overall procedure using a flowchart. It is therefore advised, when the crystal structure of a 3D model of a receptor is known, to use the conformation of the template derived from the combination of theoretical calculations and ranking of docking scores, as well as population analysis of docked conformers. It is well known that the knowledge of the receptor structure is not a prerequisite for 3D QSAR analysis, however, this study clearly shows that the availability of a crystal structure or 3D model for a receptor facilitates the structure alignment and provides a model marginally more reliable statistically.

References and notes

1. Durdagi, S.; Kapou, A.; Kourouli, T.; Andreou, T.; Nikas, S. P.; Nahmias, V. R.; Papahatjis, D. P.; Papadopoulos, M. G.; Mavromoustakos, T. *J. Med. Chem.* **2007**, *50*, 2875.
2. Mavromoustakos, T.; Theodoropoulou, E.; Zervou, M.; Kourouli, T.; Papahatjis, D. *J. Pharm. Biomed. Anal.* **1999**, *18*, 947.
3. Mavromoustakos, T.; Kapou, A.; Benetis, N. P.; Zervou, M. *Drug Des. Rev. Online* **2004**, *1*, 235.
4. Zoumpoulakis, P.; Mavromoustakos, T. *Drug Des. Rev. Online* **2005**, *2*, 537.
5. Shim, J.-Y.; Welsh, W. J.; Howlett, A. C. *Biopolymers* **2003**, *71*, 169.
6. Tuccinardi, T.; Ferrarini, P. L.; Manera, C.; Ortore, G.; Saccomanni, G.; Martinelli, A. *J. Med. Chem.* **2006**, *49*, 984.

7. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739.

8. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

9. Quanta, Molecular Simulations Inc., San Diego, CA, 1997.

10. Reggio, P. H. *Curr. Pharm. Des.* **2003**, *9*, 1607.

11. *Sybyl Molecular Modeling Software Package*, ver. 6.8., Tripos Inc., St. Louis, MO 63144, 2001.

12. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.

13. Chen, J.-Z.; Han, X.-W.; Liu, Q.; Makriyannis, A.; Wang, J.; Xie, X.-Q. *J. Med. Chem.* **2006**, *49*, 625.

14. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.

15. Papahatjis, D. P.; Kourouli, T.; Abadji, V.; Goutopoulos, A.; Makriyannis, A. *J. Med. Chem.* **1998**, *41*, 1195.

16. Papahatjis, D. P.; Nikas, S. P.; Kourouli, T.; Chari, R.; Xu, W.; Pertwee, R. G.; Makriyannis, A. *J. Med. Chem.* **2003**, *46*, 3221.

17. Papahatjis, D. P.; Nahmias, V. R.; Andreou, T.; Fan, P.; Makriyannis, A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1616.

18. Papahatjis, D. P.; Nikas, S. P.; Andreou, T.; Makriyannis, A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3583.

19. Gareau, Y.; Dufrense, C.; Gallant, M.; Rochette, C.; Sawyer, N.; Slipetz, D. M.; Tremblay, N.; Weech, P. K.; Metters, K. M.; Labelle, M. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 189.

20. Tripos Bookshelf, v. 6.8, Tripos Inc., St. Louis, MO 63144, 2001.

21. Thakur, G. A.; Duclos, R. I., Jr.; Makriyannis, A. *Life Sci.* **2005**, *78*, 454.

22. Khanolkar, A. D.; Palmer, S. L.; Makriyannis, A. *Chem. Phys. Lipids* **2000**, *108*, 37.

23. Razdan, R. K. *Pharmacol. Rev.* **1986**, *38*, 75.

24. Makriyannis, A.; Rapaka, R. S. *Life Sci.* **1990**, *47*, 2173.