J Comput Aided Mol Des (2008) 22:1–17
DOI 10.1007/s10822-007-9146-7

Three-dimensional quantitative structure-activity relationship (3D QSAR) and pharmacophore elucidation of tetrahydropyran derivatives as serotonin and norepinephrine transporter inhibitors

Prashant S. Kharkar · Maarten E. A. Reith · Aloke K. Dutta

Received: 11 June 2007 / Accepted: 28 October 2007 / Published online: 4 December 2007

Abstract Three-dimensional quantitative structure-activity relationship (3D QSAR) using comparative molecular field analysis (CoMFA) was performed on a series of substituted tetrahydropyran (THP) derivatives possessing serotonin (SERT) and norepinephrine (NET) transporter inhibitory activities. The study aimed to rationalize the potency of these inhibitors for SERT and NET as well as the observed selectivity differences for NET over SERT. The dataset consisted of 29 molecules, of which 23 molecules were used as the training set for deriving CoMFA models for SERT and NET uptake inhibitory activities. Superimpositions were performed using atom-based fitting and 3-point pharmacophore-based alignment. Two charge calculation methods, Gasteiger-Hückel and semiempirical PM3, were tried. Both alignment methods were analyzed in terms of their predictive abilities and produced comparable results with high internal and external predictivities. The models obtained using the 3-point pharmacophore-based alignment outperformed the models with atom-based fitting in terms of relevant statistics and interpretability of the generated contour maps. Steric fields dominated electrostatic fields in terms of contribution. The selectivity analysis (NET over SERT), though yielded models with good internal predictivity, showed very poor external test set predictions. The analysis was repeated with 24 molecules after systematically excluding so-called outliers (5 out of 29) from the model derivation process. The resulting CoMFA model using the atom-based fitting exhibited good statistics and was able to explain most of the selectivity (NET over SERT)-discriminating factors. The presence of −OH substituent on the THP ring was found to be one of the most important factors governing the NET selectivity over SERT. Thus, a 4-point NET-selective pharmacophore, after introducing this newly found H-bond donor/acceptor feature in addition to the initial 3-point pharmacophore, was proposed.

Keywords CoMFA · 3D QSAR · NET · Pharmacophore · SERT

P. S. Kharkar · A. K. Dutta
Department of Pharmaceutical Sciences, Wayne State University, Detroit, MI 48202, USA

M. E. A. Reith
Department of Psychiatry, New York University, New York, NY 10016, USA

A. K. Dutta (✉)
Department of Pharmaceutical Sciences, Applebaum College of Pharmacy & Health Sciences, Rm# 3128, Detroit, MI 48202, USA
e-mail: adutta@wayne.edu

Introduction

The function of monoamine transporters is to terminate the action of biogenic amine neurotransmitters such as dopamine (DA), serotonin (5-HT) and norepinephrine (NE) in the synaptic cleft by transporting it back into respective neurons, a process called as reuptake [1]. These neurotransmitters are vital for the regulation of cognitive functions and affective states and also in the development of many neuropsychiatric disorders such as depression, substance dependence and Parkinson's disease [2]. The monoamine transporter systems including DAT, SERT and NET, thus, play an important role in maintaining the concentration of these neurotransmitters in the central nervous system (CNS) as well as modulating the intensity and duration of monoamine signaling in the synapses. These transporters are actively involved in different pathological

processes and are implicated in several neurological disorders such as depression, drug abuse, anxiety, mood problems, attention deficit hyperactivity disorder (ADHD), Parkinson's disease, Alzheimer's disease, schizophrenia, etc. [3–8]. The selective monoamine transporter inhibitors have been developed to treat abovementioned and other psychological and neurological disorders [9–12]. Selective serotonin reuptake inhibitor (SSRI) fluoxetine is known to possess pharmacological activity against depression and is used extensively as an antidepressant agent. On the other hand, NET selective molecules are relatively rare compared to those for DAT and SERT. A highly selective NET blocker known as reboxetine (Fig. 1) has been advanced for its strong antidepressant effect [13–15]. Compounds which exhibit dual activities at both, SERT and NET, are known as SNRIs. Two well-known molecules belonging to the SNRI category are duloxetine and venlafaxine (Fig. 1) [16]. SNRIs such as venlafaxine exhibit antidepressant activity in clinical trials and produce somewhat greater response and remission rates compared to SSRIs [17].

However, in spite of these progresses, major issues concerning the selectivity and mechanisms of action of these compounds remain unsolved. Variation in the selectivity and potency profiles of the novel inhibitors at these monoamine transporters may be of value in unraveling the relevant underlying mechanisms and thus aid in the design and discovery of novel therapies tailor-made for a particular condition avoiding the unwanted CNS side effects.

In our recent effort to identify novel templates for monoamine transporter inhibitors, we embarked on the development of 3,6-disubstituted and 2,4,5-trisubstituted tetrahydropyran (THP) derivatives targeting monoamine transporter systems [18–21]. In one of our recent reports, we have demonstrated the development of asymmetric trisubstituted THP derivatives in different stereo- and regioselective isomers with respect to the location of hydroxyl and substituted-amino groups on the THP ring [20]. These (2S,4R,5R)-2,4,5-trisubstituted THP derivatives exhibited selective or nonselective interaction mainly with the SERT or NET [19, 20]. For trisubstituted THP derivatives, the selectivity for different transporters varied with the substitutions on the N-benzyl moiety. Further, addition of potential hydrogen-bonding functionalities such as hydroxyl-, amino- or 5-indolyl at the 4-position of the N-benzyl moiety significantly improved potencies for the SERT and NET, in particular for NET [21]. Similar modifications with the asymmetric (3S,6S)-disubstituted compounds resulted in low nanomolar potencies for the SERT and NET. Thus, our SAR study led to the development of asymmetric THP derivatives, which can broadly be categorized in three different classes : SNRIs, NRIs and triple-uptake blockers. The compounds D-168, D-165 and D-161 in Fig. 1 represent example of each class, respectively.

To further explore the structural requirements of the di- and trisubstituted THP derivatives for the SERT and NET inhibitory potencies along with the factors responsible for the observed selectivity for one transporter over the other, we herein report the three-dimensional quantitative structure-activity relationship (3D QSAR) study using comparative molecular field analysis (CoMFA) method. The goal of our present study is to extract important information regarding the potency and selectivity of the novel THP derivatives for the SERT and NET inhibitory activities and aid in the design and development of potent and selective inhibitors tailor-made for a particular therapeutic indication. Furthermore, in order to add more diversity to our compound database to increase the reliability of the proposed model, we designed and synthesized three new compounds, **4–6**, as shown in Scheme 1.

The 3D QSAR analyses the quantitative relationship between the biological activity of a set of compounds and their 3D properties using statistical correlation methods. The CoMFA method of 3D QSAR was introduced by Cramer, et al., in 1988, in which an assumption is made that the interaction between an inhibitor and its molecular target is primarily noncovalent in nature and shape-dependent [22]. Therefore, a QSAR may be derived by sampling the steric and electrostatic fields surrounding a set of ligands and correlating the differences in these fields to observed variations in biological activity. In this paper, we

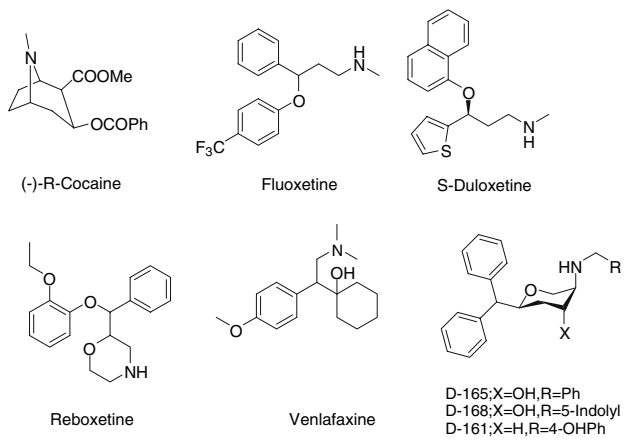
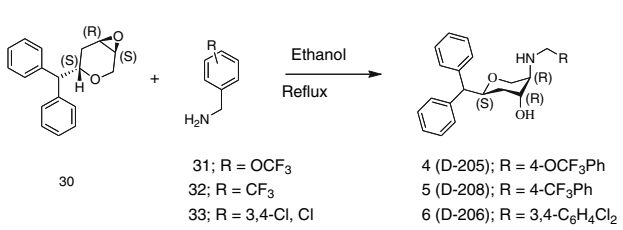


Fig. 1 Structures of selected DAT, SERT and NET inhibitors



Scheme 1

present 3D QSAR using CoMFA of the novel di- and tri-substituted THP derivatives as SERT and NET inhibitors.

Materials and methods

Reagents and solvents were obtained from commercial suppliers and used as received unless otherwise indicated. Dry solvent was obtained according to the standard procedure [23]. All reactions were performed under inert atmosphere (N_2) unless otherwise noted. Analytical silica gel-coated TLC plates (Si250F) were purchased from Baker, Inc. and were visualized with UV light or by treatment with phosphomolybdic acid (PMA). Flash chromatography was carried out on Baker Silica Gel 40 μm. ^{1}H NMR spectra were routinely obtained at Varian 400 MHz FT NMR. The NMR solvent used was $CDCl_3$ as indicated. TMS was used as an internal standard. Elemental analyses were performed by Atlantic Microlab, Inc. and were within ±0.4% of the theoretical value.

[^{3}H]Dopamine (59.4 Ci/mmol), [^{3}H]serotonin (30.0 Ci/mmol), and [^{3}H]norepinephrine (52.0 Ci/mmol) were obtained from Dupont-New England Nuclear (Boston, MA, U.S.A). (−)-Cocaine hydrochloride, WIN 35,428 naphthalene sulfonate, GBR 12909 dihydrochloride (1-[2-[bis(4-fluorophenyl)methoxy]ethyl]-4-[3-phenylpropyl]piperazine) were purchased from SIGMA-ALDRICH (St. Louis, MO).

Synthesis of trisubstituted THP derivatives

Synthesis of (−) (2S, 4R, 5R)-2-Benzhydryl-5-(4-(trifluoromethoxy)benzylamino)tetrahydropyran-4-ol (4)

A mixture of (−)-trans-(1*S*, 4*S*, 6*R*)-4-benzhydryl-3,7-dioxabicyclo-[4.1.0]heptane, 30 [20], (67 mg, 0.25 mmol) and 4-trifluoromethoxybenzylamine (0.10 mL, 0.13 g, 0.64 mmol) in ethanol (2 mL) was refluxed under N_2 overnight. The solvent was removed and the residue was purified by flash chromatography on silica gel (1:1 hexane–ethyl acetate + 5% Et_3N) to give a white solid that was recrystallized from ethyl acetate–hexane to white crystals (120 mg, 100%): mp 129–130 °C. ($[\alpha]_D = -68.4$, $c = 1.0$, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.44 (bdt, 1H, H-3eq), 1.64 (bs, 2H, OH, NH), 1.70 (ddd, $J = 14.4, 10.4, 2.8$ Hz, 1H, H-3ax), 2.43 (bd, $J = 2.4$ Hz 1H, H-5eq), 3.72 (d, $J = 13.6$ Hz, Ar-CH_2), 3.79 (d, $J = 10.8$ Hz, 1H, H-4eq), 3.86–4.00 (m, 4H, H-6, Ph-CH_2), 4.50 (dt, $J = 9.2$, 2.4 Hz, 1H, H-2ax), 7.13–7.36 (m, 14H, aromatic H). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 33.7, 50.8, 56.7, 57.0, 65.1, 67.7, 73.8, 121.2, 126.6, 126.8, 128.6, 128.7, 128.9, 129.5, 139.3, 142.2, 142.3. ^{19}F NMR ($CDCl_3$, 376 MHz) δ −58.3. The free base was converted into the oxalate salt: mp 218–219 °C (dec.). Anal. [$C_{26}H_{26}F_3NO_3 \cdot (COOH)_2$] Calc; C, 61.42, H, 5.15, N, 2.56. Found; C, 61.31, H, 5.31, N, 2.50.

Synthesis of (−) (2S, 4R, 5R)-2-Benzhydryl-5-(3,4-(dichloro)benzylamino)tetrahydropyran-4-ol (5)

A mixture of (−) trans-(1*S*, 4*S*, 6*R*)-4-benzhydryl-3,7-dioxabicyclo[4.1.0]-heptane, 30 [20], (67 mg, 0.25 mmol) and 3,4-dichlorobenzylamine (0.09 mL, 0.11 g, 0.62 mmol) in ethanol (2 mL) was refluxed under N_2 overnight. The solvent was removed and the residue was purified by flash chromatography on silica gel (1:1 hexane–ethyl acetate + 5% Et_3N) to give a white solid that was recrystallized from ethyl acetate–hexane to white crystals (90.5 mg, 82%): mp 144–146 °C. ($[\alpha]_D = -82.6$, $c = 1.0$, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.40 (bdt, 1H, H-3eq), 1.70 (ddd, $J = 14.4, 10.4, 2.8$ Hz, 1H, H-3ax), 1.77 (bs, 2H, OH, NH), 2.36 (bd, $J = 2.4$ Hz 1H, H-5eq), 3.64 (d, $J = 14.0$ Hz, Ph-CH_2), 3.75 (d, $J = 12.0$ Hz, 1H, H-4eq), 3.81 (d, $J = 14.0$ Hz, 1H, H-6ax), 3.87–3.94 (m, 4H, H-6eq, Ar-CH_2), 4.48 (dt, $J = 9.2$, 2.4 Hz, 1H, H-2ax), 7.10–7.44 (m, 13H, aromatic H). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 33.7, 50.5, 56.7, 57.0, 65.1, 67.6, 73.8, 126.6, 126.8, 127.5, 128.6 (2C), 128.9, 130.1, 130.5, 131.0, 132.6, 141.0, 142.1, 142.3. The free base was converted into the oxalate salt: mp 219–220 °C (dec.). Anal. [$C_{25}H_{25}Cl_2NO_2 \cdot (COOH)_2$] Calc; C, 60.91, H, 5.11, N, 2.63. Found; C, 60.92, H, 5.00, N, 2.61.

Synthesis of (−) (2S, 4R, 5R)-2-Benzhydryl-5-(4-(trifluoromethyl)benzylamino)tetrahydropyran-4-ol (6)

A mixture of (−)-trans-(1*S*, 4*S*, 6*R*)-4-benzhydryl-3,7-dioxabicyclo-[4.1.0]heptane**,** 30 [20]**,** (67 mg, 0.25 mmol) and 4-trifluoromethylbenzylamine (0.09 mL, 0.11 g, 0.62 mmol) in ethanol (2 mL) was refluxed under N_2 overnight. The solvent was removed and the residue was purified by flash chromatography on silica gel (1:1 hexane–ethyl acetate + 5% Et_3N) to give a white solid that was recrystallized from ethyl acetate–hexane to white crystals (94.9 mg, 86%): mp 127–128 °C. ($[\alpha]_D = -81.9$, $c = 1.0$, MeOH). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.44 (bdt, 1H, H-3eq), 1.66 (bs, 2H, OH, NH), 1.73 (ddd, $J = 14.4, 10.4, 2.8$ Hz, 1H, H-3ax), 2.41 (bd, $J = 2.4$ Hz 1H, H-5eq), 3.78 (d, $J = 13.6$ Hz, Ar-CH_2), 3.79 (d, $J = 10.8$ Hz, 1H, H-4eq), 3.90–4.00 (m, 4H, H-6, Ph–CH_2), 4.50 (dt, $J = 9.2$, 2.4 Hz, 1H, H-2ax), 7.13–7.36 (m, 10H, aromatic H), 7.43 (d, $J = 8.4$ Hz, 2H, aromatic H-2), 7.56 (d, $J = 8.4$ Hz, 2H, aromatic H-3). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 33.7, 51.1, 56.7, 57.0, 65.1, 67.7, 73.8, 125.5 (q, $J = 3.8$ Hz), 126.6, 126.8, 128.4, 128.6 (2C), 128.7, 128.9, 142.2, 142.3, 144.7. ^{19}F NMR ($CDCl_3$, 376 MHz) δ −62.8. The free

base was converted into the oxalate salt: mp 207–208 °C (dec.). Anal. [$C_{26}H_{26}F_3NO_2 \cdot (COOH)_2$] Calc; C, 63.27, H, 5.31, N, 2.64. Found; C, 63.21, H, 5.25, N, 2.60.

Hardware and Software

All the molecular modeling studies including CoMFA reported herein were performed on a Hewlett-Packard xw4300 computer workstation with main memory of 2 GB and Intel® Pentium® 4 CPU of 3.4 GHz running under the operating system Linux Red Hat 3. The molecular modeling software package Sybyl 7.1 from Tripos Inc. was employed for the present work [24].

Biological data

The given biological data had been determined for the DAT, SERT and NET by competitive radioligand uptake assays. For DAT uptake assays, uptake of [^{3}H]DA into rat striatal synaptosomes was measured exactly as described by us previously [20]. The same general protocol was used for the measurement of uptake of [^{3}H]serotonin and [^{3}H]NE into synaptosomes prepared from rat cerebral cortex. The IC_{50} value, thus estimated, was converted to K_i with Cheng–Prusoff equation. The DAT activity distribution for the present set of molecules is negatively skewed and the spread of DAT activity values is 1.76 log units (Table 1) as opposed to 3.13 and 2.91 log units for the NET and SERT activities, respectively. Thus, for the derivation of 3D QSAR models, only the biological activities at the SERT and NET were taken into account. The negative logarithms (pK_i) of the respective equilibrium inhibition constants (K_i) for SERT and NET uptake were used as dependent variables in the subsequent CoMFA analyses. For selectivity analysis between SERT and NET inhibitors, the differences between the pK_is for each inhibitor at the SERT and the NET were applied as dependent variable in the CoMFA study. This method has been successfully applied previously [25–27].

A total of 29 molecules were divided into training (23) and test sets (6). Selection of the training and test sets was done by considering the fact that the test set molecules represent range of biological activity similar to that of the training set. The primary criterion for the selection of the test sets was the uptake inhibitory activities of NET as evident from Fig. 2a, b. Further, the structural diversity was applied as secondary criterion wherein (+)/(−)-tri-substituted regioisomers as represented by test sets 1 and 2 and 3 and 4, respectively, and disubstituted compounds as represented by test set 5, were tried. Similar criteria were applied for the SERT uptake inhibitory activities for the selection of training and test sets. Thus, the test set was true representative of the training set. The structures and the respective biological activities of the molecules used in the present study at the DAT, SERT and NET are shown in Tables 1 and 2.

CoMFA analyses

All the compounds in the present investigation were either constructed from X-ray crystal structure or using fragments in the Sybyl's library. All molecules used for 3D QSAR derivation were generated in their N-protonated form. Each structure was fully geometry-optimized using Tripos force field [28] with a distance-dependent dielectric function until a root mean square (rms) deviation of 0.001 kcal/mol Å was achieved. Partial atomic charges required for electrostatic interaction were computed by two methods (a) Gasteiger–Hückel and (b) semiempirical molecular orbital method using PM3 hamiltonian in MOPAC [29] program.

The conformational search was performed using systematic search protocol. The rotatable bonds in all molecules were searched from 0 to 359^0 in 10^0 increments. The conformations within ±5 kcal/mol from the lowest energy conformation were selected for further analysis. Some of the low energy conformations for few molecules exhibited cation-π interaction between one of the phenyl rings of diphenylmethyl substituent and the cationic nitrogen. There are no reports of the existence of such conformation; hence such potentially false low energy conformations were removed. The minimum energy conformations, thus obtained, were minimized using Tripos force field and subsequently used in the analyses.

Alignment

Atom-based alignment rules were used in the present study. For atom-based fitting, atoms indicated by asterisk (*) in Fig. 3a were used for rms fitting onto the corresponding atoms of the template structures (D-168 for NET uptake, D-187 for SERT uptake and D-165 for selectivity SERT/NET). In case of 3-point pharmacophore-based alignment, ring centroids of both phenyl rings of the diphenylmethyl substituent (hydrophobic features) and cationic nitrogen (positive ionizable feature) were used as shown in Fig. 3a. The Atom-based fitting of compounds representing each of the 6 chiral classes (compounds 4, 13, 14, 20, 25 and 29, Table 1) for NET uptake inhibitory activity with compound 4 as the template and 3-point pharmacophore-based alignment of the same compounds is shown in Fig. 3b and c, respectively.

Table 1 Uptake inhibitory activities of THP derivatives at DAT, SERT and NET in rat brain

Compound			K_i (nM)			Selectivity (SERT/NET)
No.	Chirality	R	DAT Uptake, [^{3}H]DA[a]	NET Uptake, [^{3}H]NE[a]	SERT Uptake, [^{3}H]5-HT[a]	
1 (D-165)[b]	2S,4R,5R	Ph	446 ± 59	4.92 ± 0.92	707 ± 40	143.69
2 (D-142)	2S,4R,5R	4-MeOPh	115 ± 12	15.8 ± 5.3	25.9 ± 3.6	1.64
3 (D-152)	2S,4R,5R	4-OHPh	172 ± 20	10.4 ± 1.0	237 ± 12	22.79
4 (D-168)	2S,4R,5R	5-Indolyl	120 ± 8	2.13 ± 0.41	15.3 ± 1.3	7.18
5 (D-205)	2S,4R,5R	4- CF_3OPh	130 ± 26	859 ± 258	198 ± 46	0.23
6 (D-208)	2S,4R,5R	3,4-Cl_2Ph	924 ± 208	34 ± 11.8	132 ± 40	3.88
7 (D-206)	2S,4R,5R	4-CF_3Ph	123 ± 21	391 ± 68	370 ± 105	0.95
8 (D-153)	2S,4R,5R	4-FPh	125 ± 20	12.2 ± 1.0	607 ± 30	49.75
9 (D-180)	2S,4R,5R	3,4-di-MeOPh	528 ± 95	95.0 ± 9.7	20.8 ± 0.2	0.22
10 (D-179)	2S,4R,5R	3,5-di-MeOPh	100 ± 2	101 ± 12	19.7 ± 3.0	0.20
11 (D-154)	2S,4R,5R		317 ± 50	406 ± 13	3210 ± 130	7.91
12 (D-166)	2S,4R,5S	Ph	1940 ± 100	296 ± 29	9560 ± 1250	32.30
13 (D-141)	2S,4R,5S	4-MeOPh	611 ± 93	686 ± 30	793 ± 84	1.16
14 (D-161)	3S,6S	4-OHPh	85 ± 5.9	5.09 ± 0.92	37.7 ± 2.6	7.41
15 (D-187)	3S,6S	5-Indolyl	214 ± 11	6.56 ± 1.01	14.5 ± 2.7	2.21
16 (D-199)	3S,6S	4-NO_2Ph	34.6 ± 6.5	54.3 ± 9.8	19.9 ± 1.5	0.37
17 (D-185)	3S,6S	4-NH_2Ph	62.4 ± 5.6	12.6 ± 3.7	16.1 ± 1.6	1.28
18 (D-140)	2R,4S,5S	Ph	253 ± 32	691 ± 29	5500 ± 650	7.96
19 (D-151)	2R,4S,5S	4-MeOPh	110 ± 16	412 ± 36	669 ± 26	1.62
20 (D-158)	2R,4S,5S	4-OHPh	91.5 ± 7.4	94.4 ± 18.5	2540 ± 430	26.91
21 (D-167)	2R,4S,5S	5-Indolyl	184 ± 16	13.2 ± 4.0	223 ± 45	16.89
22 (D-159)	2R,4S,5S	4-FPh	327 ± 22	1050 ± 50	5000 ± 570	4.76
23 (D-157)	2R,4S,5S		162 ± 15	71 ± 2.7	2650 ± 340	37.32
24 (D-143)	2R,4S,5R	Ph	716 ± 72	2890 ± 510	3970 ± 370	1.37
25 (D-160)	2R,4S,5R	4-MeOPh	1980 ± 190	1700 ± 50	3420 ± 410	2.01
26 (D-155)	3R,6R	4-OHPh	67.4 ± 4	40.9 ± 6	187 ± 25	4.57
27 (D-186)	3R,6R	5-Indolyl	142 ± 43	15.6 ± 2.3	11.8 ± 1.6	0.76
28 (D-200)	3R,6R	4-NO_2Ph	59.9 ± 14.2	81.6 ± 9.7	64.6 ± 11.8	0.79
29 (D-184)	3R,6R	4-NH_2Ph	43.9 ± 5.2	25.6 ± 3.6	45.7 ± 17.7	1.79

[a] For uptake by DAT, SERT and NET, [^{3}H]DA, [^{3}H]5-HT and [^{3}H]NE accumulation were measured. Results are average ± SEM of three to eight independent experiments assayed in triplicate

[b] Uptake affinity of all the listed compounds except D-205, D-206 and D-208, are taken from Reference # 20 and # 21

CoMFA steric and electrostatic interaction fields were calculated at each lattice interaction points of a regularly spaced grid of 2.0 Å. An sp^3 carbon atom with van der Waals radius of 1.52 Å and +1.0 charge was used as a probe to calculate steric and electrostatic fields. Values of both the fields were truncated at +30 kcal/mol. The electrostatic fields were ignored at the lattice points with maximal steric interactions.

PLS analyses

To obtain a 3D QSAR model, PLS [30] fitting was used. The PLS method has been utilized successfully in numerous QSAR studies aiming to rationalize those structural features affecting biological activity. PLS regression seeks a relationship between Y and X, where vector Y is the response or dependent variable and X represents the descriptor data.

PLS analyses were performed following CoMFA standard implementation in Sybyl 7.1. Different descriptor blocks have been scaled to each other using the CoMFA standard scaling option. To check the statistical significance of the 3D QSAR models, cross-validations were done by means of leave-one-out (LOO) procedure. The results from cross-validation analyses were expressed as cross-validated r^2 (r^2_{cv}). Optimum number of components was determined by selecting the smallest s_{press} value. Usually the smallest s_{press} value corresponds to the highest r^2_{cv} value. The optimal number of components was subsequently used to derive the

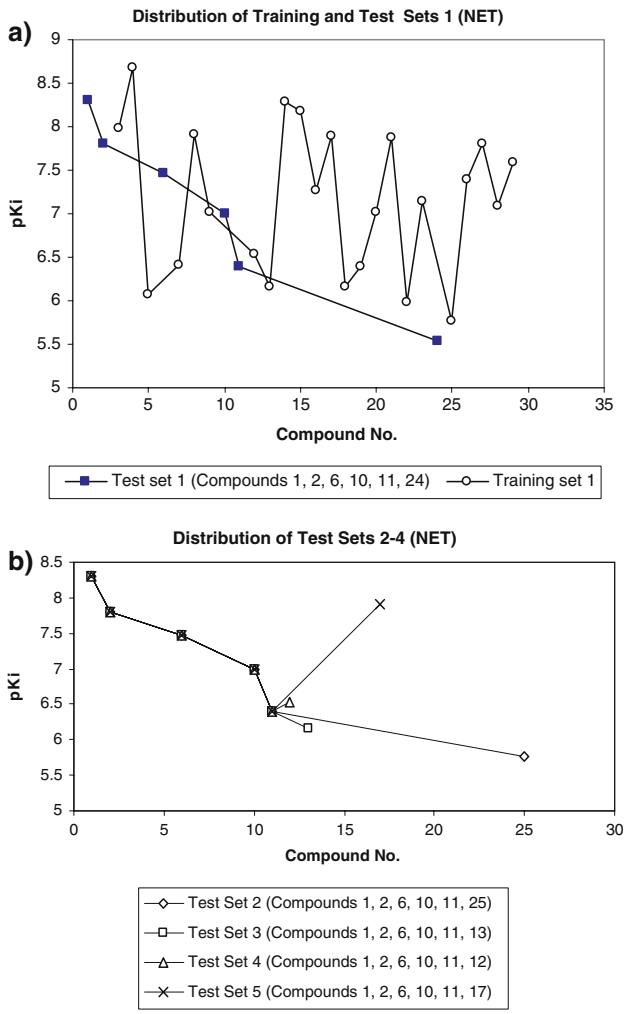


Fig. 2 (**a**) Distribution of training and test sets 1 and (**b**) distribution of test sets 2–4 for NET uptake inhibitory activity

final QSAR models. For all conventional analyses (no cross-validation), minimum σ standard deviation threshold was checked from 0.0 to 3.0 kcal/mol. The r^2_{cv}, s_{press}, r^2_{conv} and SE values were computed as defined in SYBYL. SE is the standard error of estimate. It is a measure of target property uncertainty still unexplained after the QSAR has been derived.

In Table 3, $Pr^2 = 0$ means the probability of obtaining the observed F ratio value by chance alone, if the target and the explanatory variables themselves are truly uncorrelated. When $Pr^2 = 0$, then the results are not by chance and are significant. Additionally, to perform an even more rigorous statistical test, several runs of cross-validations using five groups were done in which each target property value is predicted by a model based on about 4/5 or 80% of the available data. To further access the robustness and statistical confidence of the derived models, bootstrapping analysis (100 runs) was performed.

A common test to check the consistency of the models is to scramble the biological data and repeat model derivation process, allowing detection of possible chance correlations. After randomization of our data set, we observed very low or negative r^2_{cv} values in the PLS analyses (data not shown).

Predictive r^2 value

The predictive r^2 was computed using the molecules not included in the training set (test set). Like $r^2_{cv,}$ predictive r^2 can assume a negative value reflecting a complete lack of predictive ability of the training set for the molecules included in the test set [31].

Pharmacophore elucidation

The initial investigation of the 3-point monoamine inhibitory (SERT and NET, in particular) pharmacophoric (two aromatic/hydrophobic and one cationic features) distances using the monoamine transporter inhibitory drugs such as SSRIs (fluoxetine and paroxetine), SNRIs (venlafaxine and duloxetine) and NRIs (atomoxetine and reboxetine) resulted in the distance ranges as 4.37–6.03 Å (C_1-C_2), 5.2–7.61 Å (C_1-N^+) and 3.87–6.27 Å (C_2-N^+). These ranges overlap to some extent the distance ranges described for the SERT ligands [32] [4.6–5.2 Å (C_1-C_2), 7.4–8.3 Å (C_1-site point (sp)) and 6.2–6.9 Å (C_2-sp)]. These 3-point pharmacophoric features, shown in Fig. 3d, were used for 3-point pharmacophore-based alignment in the present study.

For the generation of the proposed 4-point pharmacophore, potent and selective NET-inhibitors belonging to the trisubstituted THP derivatives [(+)/(−)] such as compounds **1**, **3**, **8**, **20** and **23** (in the descending order of their NET selectivity) were selected. The ring centroids C_1 and C_2 along with the cationic feature (protonated N) of compounds **3**, **8**, **20** and **23** were aligned onto the corresponding 3-point pharmacophoric features of compound **1** using atom-based fitting. The 4-point pharmacophore along with the inter-feature distances are shown in Fig. 9.

Results and discussion

The present dataset consists of 29 molecules belonging to 3,6-disubstituted (8 out of 29) and 2,4,5-trisubstituted (21 out of 29) THP derivatives as shown in Chart 1. Compounds **1–17** are (–) (levorotatory) isomers whereas compounds **18–29** are (+) (dextrorotatory) isomers. The absolute configurations displayed by these compounds are as follows: (2S, 4R, 5R) by compounds **1–11**, (2S, 4R, 5S)

Table 2 Data set of substituted THP derivatives used in the training and test sets

Compound	pK_i[a]				Selectivity (SERT/NET) ΔpK_i[b]	
	Experimental		Fitted/Predicted		Experimental	Fitted
	NET	SERT	NET	SERT		
1	8.3080	6.1506	7.8108	6.5489	2.1574	1.9380
2	7.8013	7.5867	7.3161	7.5229	0.2146	0.4000
3	7.9834	6.6253	8.0552	6.5212	1.3581	1.5410
4	8.6716	7.8153	8.3950	7.3695	0.8563	0.8160
5	6.0660	6.6517	6.2410	6.7447	−0.5857	0.8150[c]
6	7.4685	6.8794	7.9935	6.6599	0.5891	1.4640[c]
7	6.4078	6.4318	6.4623	6.5266	0.0239	−0.0050
8	7.9136	6.2168	7.6581	6.3855	1.6968	1.7100
9	7.0223	7.6819	7.1416	7.6773	−0.6596	−0.5300
10	6.9957	7.7055	7.8676	7.6125	−0.7098	0.9370[c]
11	6.3915	5.4935	6.9615	5.5001	0.8979	0.8760
12	6.5287	5.0195	6.6285	5.9841	1.5092	1.3010
13	6.1637	6.1007	6.2326	5.9845	0.0629	0.0880
14	8.2933	7.4237	8.1205	7.2431	0.8696	0.9950
15	8.1831	7.8386	8.4348	8.0721	0.3445	0.2130
16	7.2652	7.7011	7.2320	7.6890	−0.4359	−0.4980
17	7.8996	7.7932	7.9255	7.5865	0.1064	0.7310[c]
18	6.1605	5.2596	6.4090	5.3022	0.9008	1.2250
19	6.3851	6.1746	6.2424	6.2408	0.2105	−0.0540
20	7.0250	5.5945	6.8397	5.5590	1.4305	1.0680
21	7.8794	6.6515	7.6170	6.6151	1.2279	0.3170[c]
22	5.9788	5.3010	5.9667	5.2198	0.6778	0.8780
23	7.1487	5.5768	7.1046	5.5686	1.5719	1.6130
24	5.5391	5.4012	5.8874	5.4690	0.1379	0.2410
25	5.7696	5.4660	5.5766	5.4944	0.3036	0.3710
26	7.3883	6.7282	7.5353	6.8544	0.6601	0.6560
27	7.8069	7.9281	8.1362	7.9653	−0.1212	−0.1070
28	7.0883	7.1898	7.2347	7.1787	−0.1015	−0.1520
29	7.5918	7.3401	7.4315	7.2216	0.2517	0.1840

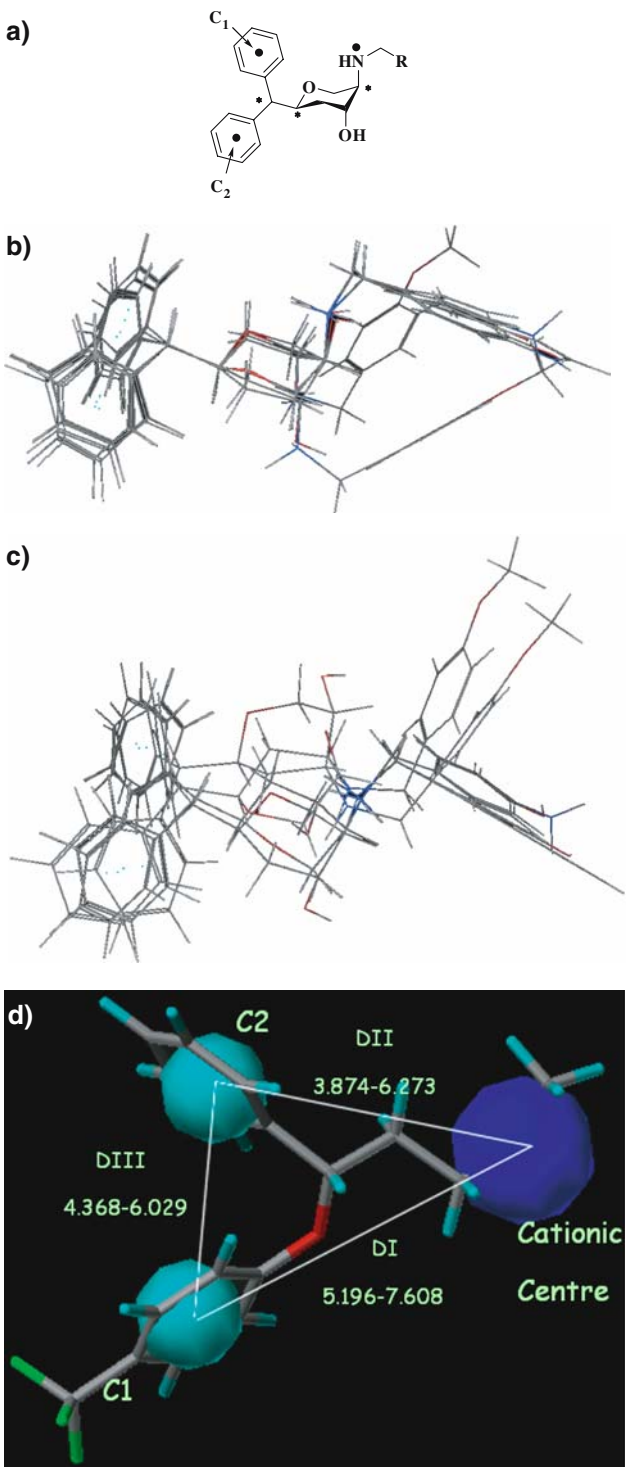
[a] pK_i is the negative logarithm of equilibrium inhibition constant

[b] $\Delta pK_i = pK_i$ NET $- pK_i$ SERT

[c] Outlier molecules

by **12–13**, (3S, 6S) by **14–17**, (2R, 4S, 5S) by **18–23**, (2R, 4S, 5R) by **24–25** and (3R, 6R) by **26–29**. Compounds **12–13** and **24–25** represent regioisomers in which the relative positions of –OH and substituted –NH groups on the THP ring are interchanged. Since the data set comprised of stereo- and regioisomers with varied substitutions and spectrum of biological activity (Table 1), we started our analysis by elucidating the common pharmacophoric features which could be used for superposition of these molecules for deriving 3D QSAR CoMFA models.

Based on the pharmacophoric features described in the literature for SERT ligands [32] and the derived 3-point pharmacophore based on the monoamine transporter inhibitory drugs, we selected the ring centroids of both phenyl rings of the diphenylmethyl substituent as hydrophobic features along with protonated N as the cationic center. The literature reports the use of a site point 2.8 Å from the basic nitrogen in the direction of the lone pair as one of the features [32]. The present set of molecules (stereo- and regioisomers) posed the obvious problems with use of such a site point for alignment. Instead, cationic nitrogen was utilized as pharmacophoric feature. The distance between the centroids C_1 and C_2 of the aromatic rings is confined to a narrow range of 4.75 to 4.93 Å. The distances C_1–N^+ and C_2–N^+ span ranges 5.46–6.69 Å and 5.65–6.7 Å, respectively. These pharmacophoric features



◀**Fig. 3** (**a**) Alignment points used for superposition[a] (**b**) Atom-based fitting of compounds representing each of the 6 chiral classes (compounds 4, 13, 14, 20, 25 and 29) for NET uptake inhibitory activity with compound 4 as the template (**c**) 3-point pharmacophore-based alignment of compounds 4, 13, 14, 20, 25 and 29 for NET uptake inhibitory activity with compound 4 as the template and (**d**) initial 3-point pharmacohore derived using a set of SSRIs, SNRIs and NRIs. R-Fluoxetine is shown for reference. *Note.* [a]The points indicated by asterisk (*) were used for atom-based fitting and the points indicated by solid dots (•) were used for 3-point pharmacophore-based alignment. C_1 and C_2 represent the ring centroids of aromatics rings of the diphenylmethyl substituent

were used for 3-point pharmacophore-based alignment (Fig. 3a). For comparison, atom-based fitting using the 'C' of diphenylmethyl group and atoms '2' and '5' of the THP ring were tried. The basic difference between these alignment procedures is the relative placement of cationic N and the substituted benzyl group attached to it for (−)- and (+)-stereoisomers.

For both alignments, compounds **4**, **15** and **1**, which represent the most active / one of the most active compounds for NET, SERT and selectivity (SERT/NET), respectively, were used as template structures. The minimum energy conformations obtained using the systematic search conformational analysis for each molecule, were used for alignment on the template structures. For most of the molecules, two conformational ensembles representing conformations differing in the orientation of the N-benzyl moiety were observed. The "extended" (N-benzyl moiety protruding away from the THP ring) conformations were lower in energy compared to their "folded" (N-benzyl group perpendicular to the THP ring) counterparts. These low-energy extended conformations were selected for superposition.

The binding affinities of these compounds for NET and SERT were measured in radioligand uptake inhibition assays in synaptosomes prepared from rat cerebral cortex using [^{3}H]-norepinephrine and [^{3}H]-serotonin as the radioligands. The ranges of inhibition constants for NET and SERT are 2.13 to 2890 nM and 11.8 to 9560 nM, respectively [20, 21]. The linear regression of the NET and SERT affinities exhibited r^2 value of 0.437 with standard error 0.6584. This moderate correlation may suggest similar yet distinct structural requirements for binding at both the transporters.

CoMFA analyses: NET uptake inhibitory activity

Using our carefully selected training sets of 23 molecules consisting of di- and trisubstituted stereo- and regioisomers, we obtained statistically significant 3D QSAR CoMFA models. The results of CoMFA analyses are summarized in Table 3. The best CoMFA model based on atom-based fitting using Gasteiger-Hückel charges yielded a correlation with r^2_{cv} of 0.573 (4 components) and conventional r^2 of 0.910 at $\sigma_{min} = 2.0$. This model showed good external predictivity with r^2_{pred} of 0.534. Increasing the minimum level of field variation (σ_{min}) for reducing the noise, the CoMFA models exhibited lowering of r^2_{cv} (data not shown). The CoMFA models based on atom-based fitting and PM3

Table 3 Summary of 3D QSAR CoMFA results

	Atom-based fitting			3-Point pharmacophore-based alignment	
	pK_i NET	pK_i SERT	Selectivity (SERT/NET)	pK_i NET	pK_i SERT
Test set	1, 6, 10,	4, 7, 8,	NA	1, 2, 6,	4, 7, 8,
molecules	11, 13, 17	14, 24, 28		10, 11, 24	12, 14, 28
r^2_{cv}	0.573	0.883	0.515	0.602	0.823
s_{press}	0.649	0.393	0.588	0.607	0.457
r^2_{conv}	0.910	0.984	0.952	0.955	0.980
SEE	0.293	0.145	0.186	0.204	0.154
Components	4	6	6	5	5
F values	45.694	165.883	55.650	72.260	166.258
$Pr^2 = 0$	0.00	0.00	0.00	0.00	0.00
Fractions					
Steric	0.587	0.677	0.783	0.632	0.709
Electrostatic	0.413	0.323	0.217	0.368	0.291
r^2_{pred}	0.534	0.820	NC	0.612	0.769
r^2_{bs}	NC	NC	NC	0.973	0.985
SD	NC	NC	NC	0.014	0.009
σ_{min}	2.0	1.0	0.0	1.0	1.0

charges resulted in poor internal as well as external predictivities (Model IV, Table 4).

The 3-point pharmacophore-based alignment yielded statistically significant models with both the charge calculation methods–PM3 and Gasteiger-Hückel (Tables 3 and 4). The models using PM3 charges were comparatively better in terms of the relevant statistics. The best model exhibited the r^2_{cv} value of 0.653 using 7 components and conventional r^2 of 0.974. The model, with 7 components, proved to be complex. In order to reduce the complexity of the model, we examined the changes in r^2_{cv} values at lower components. In accordance with the fact that an increase of r^2_{cv} values by less than 5 % is considered inappropriate due to the "parsimony-principle" [27], the CoMFA model with r^2_{cv} of 0.619 (5 components) and conventional r^2 of 0.946 ($\sigma_{min} = 0.0$) was chosen for further analysis. This model possessed good external predictivity ($r^2_{pred} = 0.540$) for a test set of 6 molecules (See Table 3). Further,

Chart 1 Substituted THP derivatives used in the training and test sets

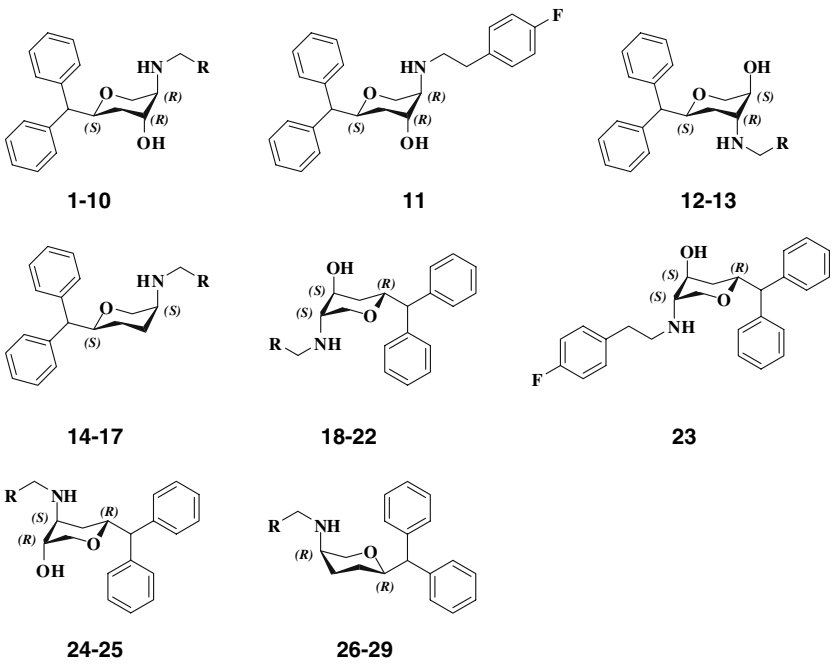


Table 4 Additional 3D QSAR CoMFA model statistics for NET uptake inhibitory activity

	I[a]	II[a]	III[a]	IV[b]	V[c]	VI[d]	VII[d]	VIII[d]
Test set	1, 2, 6,	1, 2, 6,	1, 2, 6,	1, 2, 6,	1, 2, 6	1, 2, 6	1, 2, 6	1, 2, 6
molecules	10, 11, 25	10, 11, 12	10, 11, 24	10, 11, 13	10, 11, 13	10, 11, 25	10, 11, 12	11, 13, 17
r^2_{cv}	0.461	0.614	0.473	0.404	0.617	0.584	0.668	0.642
s_{press}	0.701	0.603	0.675	0.757	0.641	0.633	0.574	0.584
r^2_{conv}	0.880	0.842	0.896	0.876	0.900	0.947	0.930	0.925
SEE	0.331	0.386	0.300	0.345	0.311	0.227	0.264	0.268
Components	4	3	4	4	4	5	4	4
F values	33.015	33.792	38.688	31.908	40.418	60.373	59.831	55.193
$Pr^2 = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fractions								
Steric	0.577	0.536	0.593	0.583	0.570	0.554	0.543	0.558
Electrostatic	0.423	0.464	0.407	0.417	0.430	0.446	0.457	0.442
r^2_{pred}	0.604	0.328	0.689	0.274	0.489	0.459	0.459	0.441
σ_{min}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a] Models based on atom-based fitting and Gasteiger-Hückel (GH) charges

[b] Model based on atom-based fitting and PM3 charges

[c] Models based on 3-point pharmacophore-based alignment and GH charges

[d] Models based on 3-point pharmacophore-based alignment and PM3 charges

reduction in noise by altering σ_{min} values yielded the final CoMFA model ($\sigma_{min} = 1.0$) for NET binding affinity with $r^2_{cv} = 0.602$ (5 components), conventional $r^2 = 0.955$ and $r^2_{pred} = 0.612$. Interestingly, the noise reduction in the 3D QSAR CoMFA model decreased the fraction of the steric contribution from 73.2% ($\sigma_{min} = 0.0$) to 63.2 % ($\sigma_{min} = 1.0$). The r^2_{pred} value also improved from 0.540 ($\sigma_{min} = 0.0$) to 0.612 ($\sigma_{min} = 1.0$). A high bootstrapped (100 runs) r^2 of 0.973 $\pm$ 0.014 adds a high confidence limit to this analysis.

The LOO cross-validation method might lead to high r^2_{cv} values, which do not necessarily reflect a general predictiveness of the models. Therefore, cross-validation using five groups was performed. In this method, a model based on about 80% of the available data predicts each target property value. The random formation of the cross-validation groups may have an effect on the results. Hence, cross-validation was performed 25 times for all of the analyses. The mean r^2_{cv} value was slightly lower (0.521) as compared to the value (0.602) obtained in the LOO method. In no case were r^2_{cv} values negative. The results obtained suggest that there is a good internal consistency in the underlying data set. Thus, this model was chosen for generating the CoMFA contour maps for graphical interpretation of the 3D QSAR model.

The real test for the model predictiveness is to predict activity of compounds not used in the model generation process. We used the test sets of 6 compounds belonging to different stereo- and regioisomers class from the di- and trisubstituted THP derivatives (Table 3). Since the SERT and NET activities are different, we used different test sets representing the spread in individual biological activities. Most of the models exhibited good predictiveness on these compounds (Table 4). Although it is generally observed that the selection of training and test sets have a strong influence on the outcome of the CoMFA models, we found that the statistics of the generated models with 5 different test sets (6 molecules each) were more dependent on alignment (atom-based fitting vs 3-point pharmacophore-based alignment) (Tables 3 and 4). Since 3-point pharmacophore-based models were better compared to the atom-based fitting, this gave the authors the confidence in the proposed pharmacophore. The authors strongly believe in the robustness of the generated 3D QSAR models which would further be validated and employed in future to design molecules tailor-made for a particular activity. The experimental and fitted NET binding affinity values are given in Table 2. The plots of fitted vs experimental activity values for training set molecules and predicted versus experimental values for the test set are shown in Figs. 4a and 5a, respectively.

The correlations obtained were not based on chance correlations as can be seen from the r^2_{cv} values (Table 3). All of these analyses were repeated 100 times with LOO cross-validation each time after randomly changing the NET binding affinities (random permutation testing) between the compounds. In all analyses, the mean r^2_{cv} values were negative. This indicates that greater than 99.9% likelihood that the relationship using the correct assignment of the biological activities did not arrive by chance.

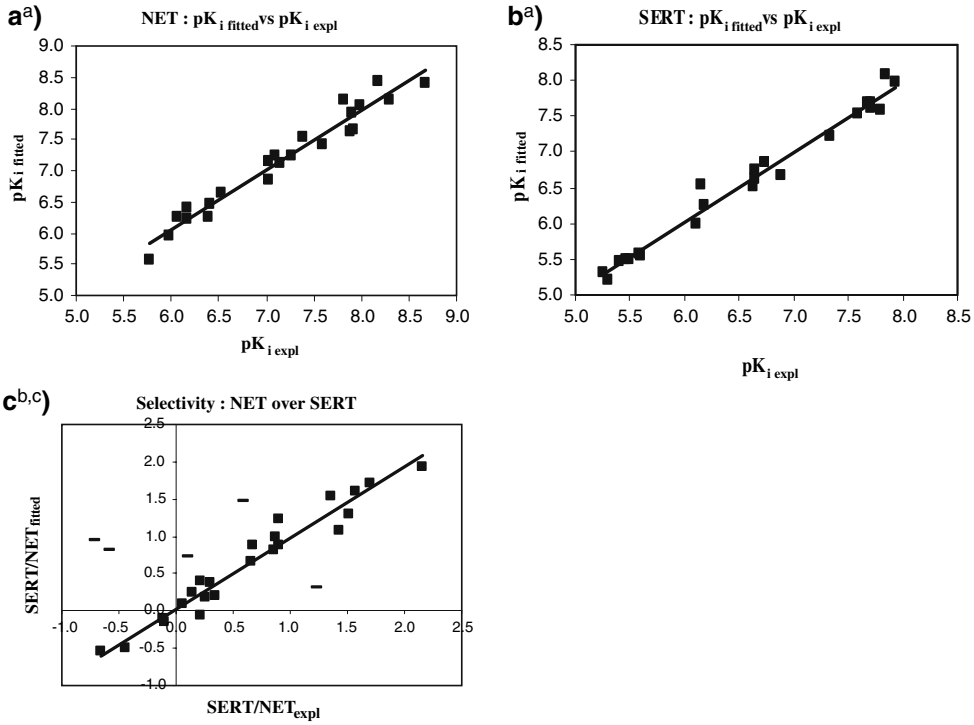


Fig. 4 Experimental versus fitted (training set) activity (**a**) NET (**b**) SERT and (**c**) Selectivity (SERT/NET) from the CoMFA analyses of the training sets. *Note.* [a]The results are from 3-point pharmacophore-based alignment. [b]The results are from atom-based fitting. [c] The training set molecules are shown as solid squares (■) and the outliers are shown as dash (−)

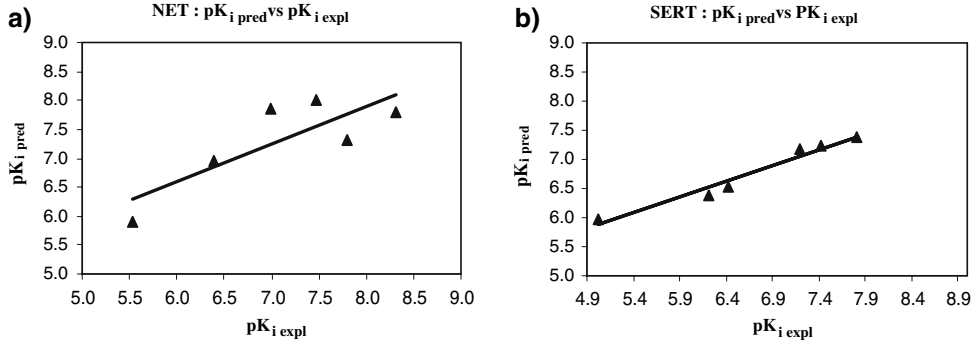


Fig. 5 Experimental vs predicted (test set) activity (**a**) NET and (**b**) SERT from the CoMFA analyses. Both the results are from 3-point pharmacophore-based alignment

CoMFA analyses: SERT uptake inhibitory activity

The CoMFA analyses of a set of 23 molecules (training set) resulted in statistically significant models for SERT binding affinity as shown in Tables 3 and 5. Both alignments resulted in models comparable in terms of the relevant statistical parameters. The best model using atom-based fitting and Gasteiger–Hückel charges exhibited cross-validated r^2 of 0.890 (7 components). This complex model showed only <5% in r^2_{cv} value using 6 components (0.883). For the increasing values of σ_{min}, the r^2_{cv} value increased from 0.883 ($\sigma_{min} = 0.0$) to 0.919 ($\sigma_{min} = 3.0$) with highest r^2_{pred} observed at 0.820 ($\sigma_{min} = 1.0$). The corresponding r^2_{cv} and conventional r^2 at $\sigma_{min} = 1.0$ were 0.892 and 0.984, respectively. The CoMFA contour maps generated using this model were complex and much more fragmented which made the interpretation of 3D QSAR difficult.

When the molecules were realigned using 3-point pharmacophore, we obtained better CoMFA models in terms of both the statistical significance and interpretability of the corresponding contour maps. The models with PM3 charges, although possessed higher internal predictivity, were poor in predicting the activity of the test set consisting of 6 molecules (Table 5). 3-Point pharmacophore-based alignemnt in combination with Gasteiger-Hückel charges resulted the best model with r^2_{cv} of 0.845 (6 components, $\sigma_{min} = 0.0$). The corresponding r^2_{cv} at 5 components was found to be 0.833 ($\sigma_{min} = 0.0$). The subsequent noise reduction by changing σ_{min} yielded the best model with $\sigma_{min} = 1.0$ with r^2_{cv}, r^2_{conv} and r^2_{pred} of 0.823, 0.980 and 0.769, respectively. This model was chosen for further analysis. The noise reduction also resulted in lowering of the steric contribution from 81.0% ($\sigma_{min} = 0.0$) to 70.9 % ($\sigma_{min} = 1.0$). The bootstrapped r^2 for 100 runs was found

Table 5 Additional 3D QSAR CoMFA model statistics for SERT uptake inhibitory activity

	IX[a]	X[a]	XI[b]	XII[b]	XIII[b]	XIV[c]	XV[d]	XVI[d]
Test set	4, 7, 8,	4, 7, 8,	4, 7, 8,	4, 7, 8,	4, 7, 8	4, 7, 8	4, 7, 8	4, 7, 8
molecules	12, 14, 28	13, 14, 28	12, 14, 28	13, 14, 28	14, 24, 28	13, 14, 28	12, 14, 28	13, 14, 28
r^2_{cv}	0.728	0.723	0.734	0.738	0.758	0.830	0.761	0.769
s_{press}	0.551	0.585	0.544	0.569	0.533	0.472	0.516	0.534
r^2_{conv}	0.966	0.957	0.965	0.981	0.964	0.985	0.973	0.981
SEE	0.196	0.231	0.198	0.153	0.205	0.139	0.174	0.152
Components	4	4	4	4	4	5	4	4
F values	126.017	99.577	123.051	233.006	121.145	226.393	161.213	235.67
$Pr^2 = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fractions								
Steric	0.654	0.664	0.651	0.654	0.641	0.670	0.610	0.627
Electrostatic	0.346	0.336	0.349	0.346	0.359	0.330	0.390	0.373
r^2_{pred}	0.788	0.652	0.671	0.507	0.576	0.379	0.408	0.274
σ_{min}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a] Models based on atom-based fitting and Gasteiger-Hückel (GH) charges

[b] Model based on atom-based fitting and PM3 charges

[c] Models based on 3-point pharmacophore-based alignment and GH charges

[d] Models based on 3-point pharmacophore-based alignment and PM3 charges

to be 0.985 ± 0.009. The experimental and fitted SERT binding affinity values are given in Table 2. The plots of fitted vs experimental activity values for training set molecules and predicted versus experimental values for the test set are shown in Figs. 4b and 5b, respectively.

Cross-validation using five groups resulted in mean r^2_{cv} value of 0.788 (25 runs). As observed with the best CoMFA model for NET binding affinity, none of the r^2_{cv} values were negative. Similar analysis using random permutation testing (100 runs) resulted in negative r^2_{cv} values. This again confirmed the absence of chance correlation in the correct assignment of biological activities in the data set used.

CoMFA analyses: selectivity for NET over SERT

To better understand the selectivity (NET over SERT)-discriminating criteria, a selectivity analysis was performed. The resulting models exhibited poor internal predictivity ($r^2_{cv} < 0.3$) (data not shown). Various combinations of training and test sets did not improve the statistics. Obviously, statistics are not expected as good as for the NET and SERT models alone, because the affinity differences are associated with a higher experimental uncertainty due to error propagation [27]. Careful examination of the residuals from the non-cross-validated PLS analysis of models using all compounds as training set led to the identification of compounds exhibiting higher residuals (outliers). Systematic exclusion of these outliers from the data set resulted in improvement of the statistics in few cases. The best model was obtained using atom-based fitting and PM3 charges. The corresponding values of the statistical parameters are shown in Table 3. This model resulted from the PLS analysis of 24 molecules (5 outliers) and exhibited r^2_{cv} of 0.515 (6 components) at $\sigma_{min} = 0.0$. The alterations in σ_{min} values led to decrease in r^2_{cv} and the contour maps derived from the corresponding CoMFA models were fragmented and difficult to interpret (data not shown). Thus, this model with $\sigma_{min} = 0.0$ was used for further analyses. The plot of fitted vs experimental selectivity values for data set of 24 molecules along with the respective values for the outliers (5) are shown in Fig. 4c.

The outlier molecules were found to be **5**, **6**, **10**, **17** and **21**. The molecules **5**, **6**, **10** and **21** belong to trisubstituted THP derivatives whereas compound **17** belong to the disubstituted category. The uniqueness of compounds **5**, **6** and **10** lies in the fact that these are (−)-trisubstituted compounds for which neither (+)-trisubstituted nor (−)/(+) disubstituted counterparts exist in the data set. Inclusion of such compounds in the model building process may improve the model statistics. No suitable explanation could be provided for the higher residuals obtained for molecules **17** and **21** in the selectivity analysis. Except for compounds **6** and **21**, the model correctly predicted NET-selective compounds [$\Delta pK_{i\ exp}$ ($pK_{i\ exp}$ NET − $pK_{i\ exp}$ SERT) > 0.5] as NET-selective with compound **1** (most selective compound) predicted as the most NET-selective. Potent and NET-selective compounds containing H-bond donor such as –OH (compounds **3** and **14**) were predicted to be more NET-selective compared to their experimental

selectivity. Dual SERT- and NET inhibitors ($\Delta pK_{i\ exp}$ −0.5 to 0.5) were correctly predicted with the exception of compounds **17** (predicted as NET-selective) and **10**. SERT-selective ($\Delta pK_{i\ exp} < -0.5$) compounds (**5**, **9**, **10**) were predicted as SERT-selective except compounds **5** and **10** (predicted as NET-selective). The criteria of defining NET-selective from SERT-selective was chosen arbitrarily. Overall, the model performed well for NET-selective and dual inhibitors.

Graphical interpretation of the CoMFA models

NET uptake inhibitory activity

To visualize the information content of the derived 3D QSAR models, CoMFA contour maps were generated by interpolating the products between 3D QSAR coefficients and their associated standard deviations. The 3D representations of the field contributions defined as "STDEV*COEFF" contour maps reveal where changes in molecular fields are able to explain the differences in experimental binding affinities. Figure 6 shows the stereoview of the CoMFA contour maps from the analysis based on 3-point pharmacophore-based alignment and PM3 charges. Figure 6a shows the steric and electrostatic contour plots with compound **4** (most active) inside the fields. The green and yellow contours present near the 4-position of the N-aryl moiety suggest the requirement of limited steric bulk at that position for high-affinity interactions. The pyrrole ring of 5-indolyl substituent overlaps the green contour indicating limited increase in bulk led to increased affinity at the NET. The corresponding compound **1** (K_i 4.92 nM) with 4-unsubstituted phenyl group is 2-fold less potent than compound **4** (K_i 2.13 nM). Similarly, in disubstituted derivatives compound **15** (5-indolyl) is more active than compound **17** (4-aminophenyl). Figure 6b depicts the contour plots with compound **5** (one of the least active molecules) inside the fields. The presence of yellow contours near the 4-OCF_3 substituent on the N-benzyl moiety suggest the detrimental effect of increased bulk near this position on the NET potency. A small sterically unfavorable yellow contour is observed near the 4-[in case of (+)-trisubstituted compounds] or 6-position [in case of (−)-trisubstituted compounds] of the THP ring indicating no steric bulk is allowed near this position. Compound **4** (K_i 2.13 nM), (−)-stereoisomer, lacks substitution at the 6-position on the THP ring as opposed to compound **21** (K_i 13.2 nM), (+)-stereoisomer, with a hydroxyl group on the THP ring overlapping the yellow contour.

In case of electrostatic contour plots, a small red contour is observed near the 4- (for (+)-stereoisomers) or 6-position (for (−)-stereoisomers) on the THP ring (Fig. 6a, b). This is an indication of the detrimental effect of the presence of negatively charged substituent and H-bond donor on the

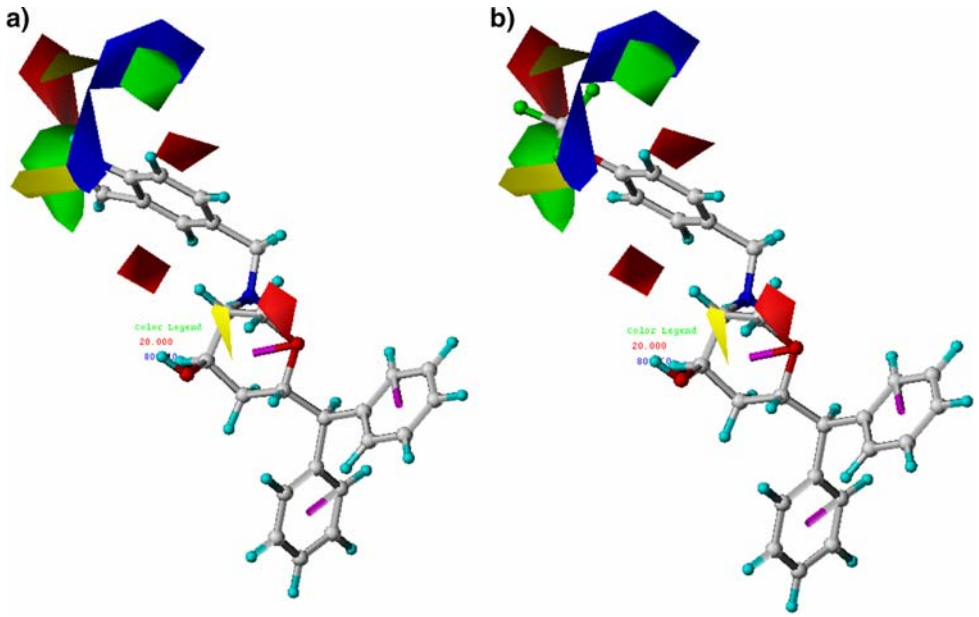


Fig. 6 CoMFA STDEV*COEFF contour plots showing steric and electrostatic features from analysis based on 3-point pharmacophore-based alignment and PM3 charges for NET potency. Sterically favored areas (contribution level of 80%) are represented by green polyhedra whereas sterically disfavored areas (contribution level of 20%) are represented as yellow polyhedra. Blue fields (contribution level of 80%) are regions of the molecule where more positive charge and H-bond donors are favored or negative charge and H-bond acceptors are disfavored for high-affinity interactions. Red fields (contribution level of 20%) indicate regions where negatively-charged substituents and H-bond acceptors are favored or more positive charge and H-bond donors are disfavored. Compound **4** is shown inside the fields in (**a**) while compound **5** is present inside the fields in (**b**)

activity. In case of (−)-stereoisomers, such as compounds **4** and **8**, the THP –OH group is located away from this red contour whereas in case of the corresponding (+)-stereoisomers, compounds **21** and **22**, respectively, the THP –OH group is located near the red contour. Thus, the location of –OH group on the THP ring is observed to be important for higher NET activity. Similar behavior was observed for other enantiomeric pairs, compounds **1** and **18** and **2** and **19**.

The prominent feature of these plots is the presence of two red contours above and below the phenyl ring of the N-benzyl moiety indicating necessity of increased electron density on the phenyl ring for high-affinity interactions. The increased electron density in the π-system may take part in π–π stacking or cation–π interactions with the NET residues surrounding the ligand-binding site. Thus, in general, electron-donating groups should increase and, electron-withdrawing groups should decrease activity for NET. Compound **4** (Fig. 6a) contains fused bicyclic π-system (indole ring) whereas compound **5** (Fig. 6b) has 4-OCF_3 substitution on the phenyl ring. The electron-withdrawing effect of the OCF_3 group reduces the electron density on the aromatic ring and may contribute, along with other steric and electrostatic factors, to reduced interactions with NET. Other (–)-trisubstituted compounds with electron-withdrawing substitution such as 3,4-dichloro (compound **6**) and 4-CF_3 (compound **7**) were less active than compounds with electron-donating groups such as 4–OCH_3 (compound **2**) and 4-OH (compound **3**). Referring to the differential activities of the enantiomeric pair, **8** and **22**, the high potency of compound **8** and the reduced potency of compound **22** for NET is the overall effect with predominant contribution from the correct placement of –OH group (compound **8**) and the possible stronger H-bond donor/acceptor interactions with the transporter.

In Fig. 6a, the indole –NH of compound **4** is directed towards the blue contour indicating presence of positively charged substituents and H-bond donors near the blue contour would increase the affinity. Small red contours were also observed near the blue contours which suggest that the presence of negatively charged substituents and H-bond acceptors would have a positive impact on the NET binding affinity. The compounds with H-bond donors were more active compared to their H-bond acceptor counterparts (compounds **3** and **2**, **20** and **19**) since some H-bond donors can also act as H-bond acceptors. Substitution with an H-bond donor group at the 4-position of the N-benzyl moiety has prominent effect on the NET affinity.

SERT uptake inhibitory activity

The steric and electrostatic contour plots generated using the CoMFA model based on 3-point pharmacophore-based alignment and Gasteiger–Hückel charges are shown in Fig. 7. Compound **15** (7a) and **18** (7b) are shown for reference. In Fig. 7a, a large sterically favorable green contour is observed near the 4-position of the N-aryl group (5-indolyl in case of compound **15**) suggesting the requirement of increased bulk near that position for high SERT binding affinity. The pyrrole ring of 5-indolyl moiety correctly overlaps onto the green contour. Similarly compounds with bulkier groups near 4-position of the N-benzyl substituent such as **2** (4-OCH_3), **9** (2,4-dimethoxy) and **10** (3,5-dimethoxy) exhibit potent affinity for SERT. Sterically unfavorable yellow contour(s) is not observed near the N-aryl substituents unlike those observed for the NET affinity (Fig. 6a, b). The presence of large and small yellow contours near the 4- and 6-positions (in case of (+)-stereoisomers) and 6- and 4-positions (in case of (−)-stereoisomers), respectively, on the THP moiety point to the fact that substitution at these positions have more or less deleterious effects on SERT potency. The effects are more

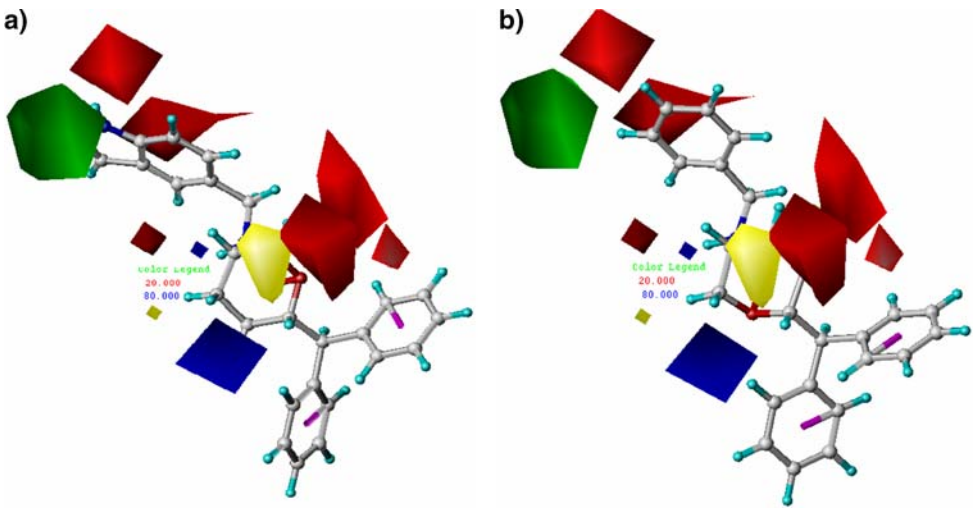


Fig. 7 CoMFA STDEV*COEFF contour plots showing steric and electrostatic features from analysis based on 3-point pharmacophore-based alignment and Gasteiger–Hückel charges for SERT potency. Compounds **15** (**a**) and **18** (**b**) are shown for reference

prominent for (+)-trisubstituted compounds in which the 4-OH substituent on the THP ring correctly overlaps onto the large yellow contour as opposed to (−)-trisubstituted compounds (Fig. 7a) as seen with compounds **18** ((+)-stereoisomer, SERT K_i 5500 nM) and **1** ((−)-stereoisomer, SERT K_i 707 nM). The disubstituted compounds [both (+)- and (−)-stereoisomers] exhibit potent SERT activity.

Significant number of large red and few blue contours are observed surrounding the molecules (Fig. 7a, b). The red contour seen near the 4-position of the N-benzyl substituent reveals the requirement of negatively-charged substituents and H-bond bond acceptors for high-affinity SERT interactions. It also indicates that more positive charge or H-bond donors are disfavored at these positions. This is evident from the higher potency of compound **2** (4-OCH_3, SERT K_i of 25.9 nM) compared to **3** (4-OH, SERT K_i of 237 nM) and compound **19** (4-OCH_3, SERT K_i of 669 nM) compared to **20** (4-OH, SERT K_i of 2540 nM). Another red contour below the plane of aromatic ring shows the positive impact of increased electron density on the N-substituted aromatic system for activity. Electron-withdrawing groups decrease the π-cloud over the aromatic ring and thus decrease the possible π–π stacking and/or cation-π interactions with SERT. In contrast, electron-donating groups have favorable effect on the SERT binding potency.

The large red contours present near the 4- or 6-position on the THP nucleus for the (+)- or (−)-stereoisomers, respectively, strongly disfavor the presence of H-bond donor group such as –OH for SERT affinity. This is evident from the lower affinity of compound **18** ((+)-stereoisomer, Fig. 7b) in which the 4-OH substituent on the THP ring is directed toward the large red contour. Similarly a blue contour seen near the 6- or 4-position of the THP ring for the (+)- or (−)-isomers, respectively, indicates that a H-bond donor like –OH group can form high-affinity interactions with SERT. In case of compound **1** ((−)-stereoisomer), 4-OH substituent on the THP ring is directed toward the large blue contour. In general, (+)-trisubstituted THP derivatives are the least active compounds in the data set by virtue of unfavorable presence of -OH group on the THP ring. Compounds **11** [(−)-trisubstituted] and **23** [(+)-trisubstituted], with 2-carbon spacer in between the protonated N and the aromatic ring, exhibited lower activity. This may probably be due to the altered orientation of the 4-fluorophenyl group away from the green contour seen in Fig. 7a and b, and the presence of –OH group on the THP ring.

Another large red contour is observed below the plane of the THP ring. This is indicative of the disfavoring of a H-bond donor or favoring of H-bond acceptor for SERT potency. In case of regioisomers **12–13** and **24–25**, the –OH group is present below the plane of the THP nucleus which is detrimental for activity. All these 4 compounds exhibited poor binding affinity at the SERT.

Selectivity for NET over SERT

The steric and electrostatic contour plots obtained from the CoMFA analysis based on atom-based fitting and PM3 charges are shown in Fig. 8. Compounds **1** (most selective) (a) and **9** (least selective) (b) are shown for reference. As seen with NET alone, limited steric bulk is allowed near the 4-position of the N-benzyl moiety as seen from small yellow (sterically unfavorable) and green (sterically favorable) contours surrounding that position (Fig. 8a). Compound **1**, possessing unsubstituted N-benzyl group, is the most selective. The selectivity is reduced with increasing steric bulk as observed with 4-F (compound **8**), 4-OH (compound **3**), 3,4-dichloro (compound **6**) and 4-OCH_3 (compound **2**). The 2,4-dimethoxy groups on

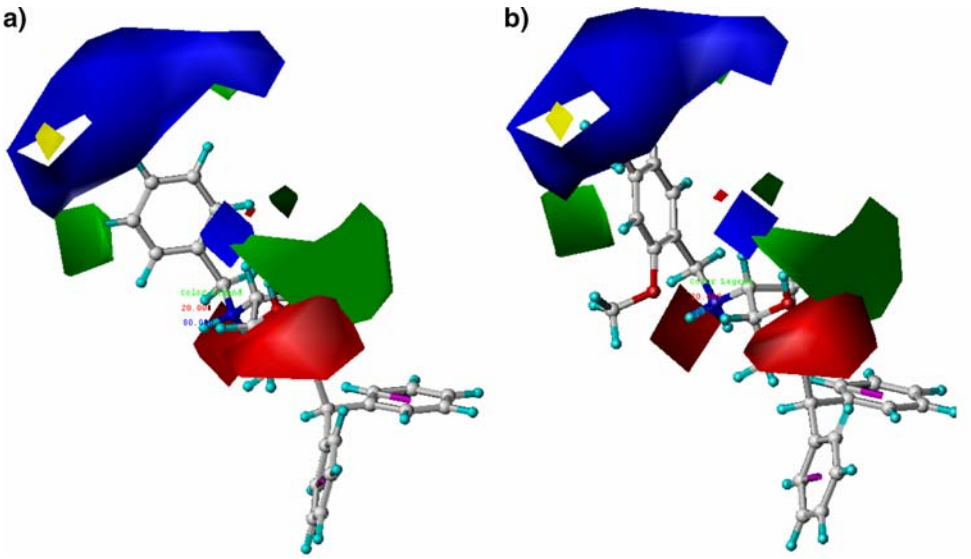


Fig. 8 CoMFA STDEV*COEFF contour plots showing steric and electrostatic features from analysis based on atom-based fitting and PM3 charges for NET selectivity. Compounds **1** (**a**) and **9** (**b**) are shown for reference

phenyl ring of the N-benzyl moiety (compound **9**) are found overlapping onto the yellow contours (Fig. 8b). Hence, the reduced selectivity was observed for NET. In addition, a large green contour is observed above the THP ring suggesting the favorable role of substitution on the THP core for the NET selectivity. This separates trisubstituted compounds from the disubstituted ones. Whereas the selectivity for NET of (−)-trisubstituted derivatives is more pronounced compared to their (+)-stereoisomeric counterparts, all the disubstituted compounds [(+)- and (−)-stereoisomers] exhibit less than 10-fold differences in the NET and SERT binding affinites ($\Delta pK_{i\ exp} < 1$, Table 2). The regioisomeric compounds **13**, **24**, **25** show poor selectivity due to lack of substitution near the 4- or 6-position on the THP ring (near the green contour).

A large blue and a small red contours observed near the 4-position of the N-benzyl group suggests the dominant role of more positive charge and H-bond donors over more negative charge and H-bond acceptors in determining NET selectivity (Fig. 8a, b). Compound **3** (4-OH) is more selective than compound **2** (4-OCH_3) and compound **20** (4-OH) is more selective than compound **19** (4-OCH_3). A similar combination of red and blue contours is observed near the 4-position (for (−)-stereoisomers)) of the THP ring indicating the importance of H-bond donor and acceptor group(s) for the activity. Since only –OH substituent (trisubstituted derivatives) is present on the THP ring, this delineates the importance of this group for the NET selectivity. Another large red and small blue contour below the THP ring is overlapped with the H-bond donor/acceptor –OH group found in the regioisomers **12** and **13**.

Overall, the –OH group on the THP moiety adds a dominant feature for NET recognition compared to SERT. This may be due to the fact that R-norepinephrine, substrate for NET, has –OH group in the side chain and thus, NET has some element of stereospecific recognition of the ligands containing suitably oriented –OH group. Hence, we add one more pharmacophoric feature to the earlier 3-point pharmacophore and propose a new NET-selective pharmacophore using trisubstituted THP derivatives as NET and SERT inhibitors. The 4-point pharmacophore along with the corresponding pharmacophoric distances is depicted in Fig. 9.

When the proposed NET-selective 4-point pharmacophoric distances were compared with the initial 3-point pharmacophore, we observed overlap in the C_1-C_2 distance (4.37–6.03 Å in case of 3-point pharmacophore versus 4.7–4.9 Å for the 4-point model) and C_1-N^+ (5.2–7.61 vs 6.3–6.6 Å). The C_2–N^+ range was found slightly higher (3.87–6.27 vs 6.3–6.6 Å). The distance ranges for the 3-point model were broad due to higher flexibility of the molecules used for generating the 3-point pharmacophore as opposed to the somewhat structurally rigid THP derivatives. The proposed 4-point model is a preliminary

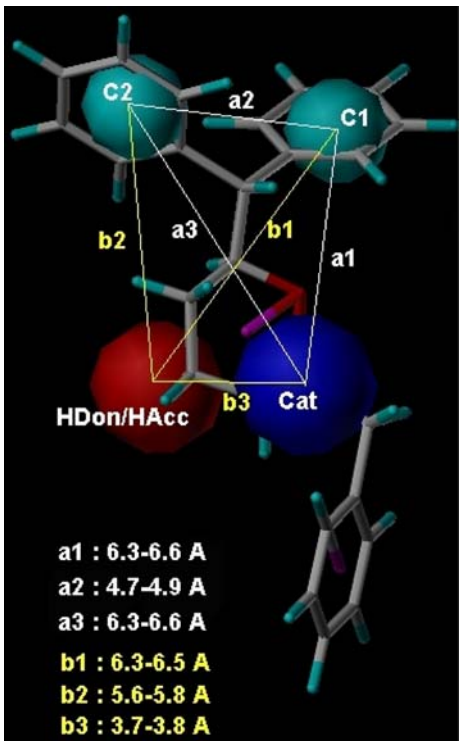


Fig. 9 Proposed 4-point NET-selective pharmacophore for the (−)-trisubstituted THP derivatives. The cyan spheres indicate the hydrophobic features (C1 and C2), the blue sphere indicates cationic feature and the red sphere indicates the H-bond donor/acceptor feature

model which does not take into consideration the contribution made by the N-benzyl/aryl substituent towards selectivity. Tri- and disubstituted THP derivatives such as compounds **3** and **14**, respectively, with same N-benzyl substituent (4-OHBz) exhibited 3-fold selectivity difference. The 4-point model can be an ideal starting point for the pharmacophore-based 3D searching of chemical databases to embark upon novel chemotypes for selective monoamine transporters blockers.

Conclusions

The 3D QSAR study using CoMFA for a series of di- and trisubstituted THP derivatives as SERT and NET inhibitors is presented. The study was conducted to rationalize the SERT and NET binding potencies along with the observed selectivity (NET over SERT). The 3-point pharmacophore-based alignment produced CoMFA models with significant statistics and readily interpretable contour maps for a dataset consisting of stereo- and regioselective isomers for SERT and NET binding affinities. A high bootstrapped r^2 value and small standard deviations indicate that a similar relationship exists in all compounds. The results show good

correlation between the steric and electrostatic fields and the SERT and NET inhibitory activities, with a dominant contribution made by the steric field over the electrostatic counterpart. The selectivity analysis using the same data set of 29 molecules did not yield models with significant internal and external predictivity. Systematic exclusion of the outliers improved the internal predictivity of the models. The best model using atom-based fitting could explain the observed selectivity for most of the compounds and led to identification of selectivity-discriminating criteria for NET over SERT. Based on this information, we identified new pharmacophoric feature signifying NET selectivity for these THP derivatives and proposed a new pharmacophore developed from the potent and NET-selective compounds incorporating this newly identified feature. The 4-point pharmacophore may be actively pursued for the rational design and development of potent and selective NET inhibitors using 3D pharmacophore-based database searching or de novo drug design approaches.

Acknowledgements This work was supported by the National Institute on Drug Abuse, Grant No. DA 12449 (AKD). We acknowledge Stuart Hazeldine for his help with the synthesis and Juan Zhen for her help with the biological characterization of compounds.

References

1. Amara SG, Kuhar MJ (1993) Annu Rev Neurosci 16:73
2. Laakso A, Hietala J (2000) Curr Pharm Des 6:1611
3. Hahn MK, Blakely RD (2002) Pharmacogenomics J 2:217
4. Heinz A, Mann K, Weinberger DR, Goldman D (2001) Alcohol Clin Exp Res 25:487
5. Klimek V, Stockmeier C, Overholser J, Meltzer HY, Kalka S, Dilley G, Ordway GA (1997) J Neurosci 17:8451
6. Robertson D, Flattem N, Tellioglu T, Carson R, Garland E, Shannon JR, Jordan J, Jacob G, Blakely RD, Biaggioni I (2001) Ann N Y Acad Sci 940:527
7. Miller GM, De La Garza R II, Novak MA, Madras BK (2001) Mol Psychiatry 6:50
8. Ozaki N, Goldman D, Kaye WH, Plotnicov K, Greenberg BD, Lappalainen J, Rudnick G, Murphy DL (2003) Mol Psychiatry 8:933
9. Nemeroff CB, Owens MJ (2002) Nat Neurosci 5:1068
10. Dutta AK, Zhang S, Kolhatkar R, Reith MEA (2003) Eur J Pharmacol 479:93
11. Olivier B, Soudijn W, van Wijngaarden I (2000) Prog Drug Res 54:59
12. Walter MW (2005) Drug Develop Res 65:97
13. Miller DK, Wong EHF, Chesnut MD, Dwoskin LP (2002) J Pharmacol Exp Ther 302:687
14. Scates AC, Doraiswamy PM (2000) Ann Pharmacother 34:1302
15. Hajos M, Fleishaker JC, Filipiak-Reisner JK, Brown MT, Wong EHF (2004) CNS Drug Rev 10:23
16. Koch S, Hemrick-Luecke SK, Thompson LK, Evans DC, Threlkeld PG, Nelson DL, Perry KW, Bymaster FP (2003) Neuropharmacology 45:935
17. Thase ME, Entsuah AR, Rudolph RL (2001) Br J Psychiatry 178:234
18. Zhang S, Reith MEA, Dutta AK (2003) Bioorg Med Chem Lett 13:591
19. Zhang S, Zhen J, Reith MEA, Dutta AK (2004) Bioorg Med Chem 12:6301
20. Zhang S, Zhen J, Reith MEA, Dutta AK (2005) J Med Chem 48:4962
21. Zhang S, Fernandez F, Hazeldine S, Deschamps J, Zhen J, Reith MEA, Dutta AK (2006) J Med Chem 49:4239
22. Cramer RD III, Patterson DE, Bunce JD (1988) J Am Chem Soc 110:5959
23. Vogel AI (1989) Vogel's textbook of practical organic chemistry, 5th edn. Longman, Harlow
24. SYBYL Molecular Modeling System, Version 7.1, Tripos Inc., St. Louis, MO 63144–2913
25. Wong G, Koehler KF, Skolnick P, Gu ZQ, Ananthan S Schonholzer P, Hunkeler W, Zhang W, Cook JM (1993) J Med Chem 36:1820
26. Matter H, Schwab W (1999) J Med Chem 42:4506
27. Böhm M, Stürzebecher J, Klebe G (1999) J Med Chem 42:458
28. Clark M, Cramer RD III, Van Opdenbosh N (1989) J Comput Chem 10:982
29. Stewart JJP (1990) J Comput-Aided Mol Des 4:1
30. Dunn WJ, Wold S, Edelund U, Helberg S (1984) Quant Struct-Act Relat 3:131
31. Cramer RD III, Bunce J, Patterson DE (1988) Quant Struct-Act Relat 7:18
32. van de Waterbeemd H, Testa B, Folkers G (eds) (1997) Computer-Assisted Lead Finding and Optimization, Verlag Helvetica Chimica Acta (VHCA). Basel and Wiley–VHC, Weinheim, pp 445–459