Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 1113–1124
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Pharmacophore design and database searching for selective monoamine neurotransmitter transporter ligands

Iain J.A. MacDougall, Renate Griffith *

School of Environmental and Life Sciences, The University of Newcastle, Australia

Received 26 June 2007; received in revised form 24 September 2007; accepted 3 October 2007
Available online 9 October 2007

Abstract

Neuronal monoamine transporters (MATs) are involved in the pathophysiology and treatment of mental health conditions such as depression, attention deficit hyperactivity disorder, substance abuse and neurodegenerative disorders including Alzheimer's disease and Parkinson's disease. Various structural classes of compounds have been synthesized and tested *in vitro* for activity against transporters of three monoamine signaling molecules: noradrenaline (NET); serotonin (SERT) and dopamine (DAT). We have developed and validated a number of pharmacophore models describing the interaction of two classes of compounds with each of these three MATs. These pharmacophores explain the selectivity of binding to the MATs for various compound classes and have been used to search *in silico* databases for novel, potentially selective ligands. These ligands, after confirmation of their activities, will provide tools for investigating the function of MATs as well as the potential for new therapeutic agents in mental health applications. The database searches also retrieved close analogues of known MAT ligands, further validating the approach.

Keywords: Noradrenaline transporter; Dopamine transporter; Serotonin transporter; Pharmacophores; Database searching

1. Introduction

The three closely related monoamine neurotransmitters dopamine, serotonin and noradrenaline mediate signal transduction between neurons *via* interaction with specific receptors. Attenuation of the signal is caused by the reuptake of neurotransmitters by specific transporter proteins. The balance of interaction between neurotransmitter, receptor and transporter is important for a number of disease states including depression, attention deficit hyperactivity disorder, drug dependence and neurodegenerative conditions such as Alzheimer's and Parkinson's diseases. Therefore, compounds which selectively modulate the effect of the monoamine transporter proteins (MATs) are important therapeutic and research tools.

While there are a number of approaches used for investigating these interactions, the focus of this paper is the development of computer-generated pharmacophores for ligands which interact with dopamine transporter (DAT), serotonin transporter (SERT) and noradrenaline transporter (NET). Computational approaches to investigating the ligand-transporter interaction can be either transporter-based or ligand-based. As there are no crystal structures for MATs and very few crystal structures available for similar transporters generally [1,2] and these are of low resolution (>3 Å), accurate structural data is hard to come by for most transporters. Homology models of the three MATs based on the distantly related crystal structures have been published [3,4], however the low level of relationship has rendered transporter-based investigation difficult to date.

The ligand-based approach as used thus far can be divided into 3D-QSAR modeling and pharmacophore modeling. The 3D-QSAR CoMFA technique has been used successfully for DAT and SERT [5–8], in particular with good correlation data for the DAT. However this technique is limited to structurally related compounds. The pharmacophore approach has previously been used in particular for the DAT, with excellent outcomes from *in silico* database screening [9–11]. Three distinct structural classes of DAT modulators were discovered using this technique. These pharmacophores previously used for searching consisted of atomic constituents rather than chemical features, which may result in an even greater diversity

* Corresponding author at: Biology Building, Callaghan, NSW 2308, Australia. Tel.: +61 249216990; fax: +61 249216923.
E-mail address: Renate.Griffith@newcastle.edu.au (R. Griffith).

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.10.003

of retrieved hits. A pharmacophore based on derivatives of 1-aryl-3-[4-arylpiperazin-1-yl]-1-propane for interaction at the SERT has been published [12]. However, no similar attempts have been made for the NET, or to distinguish differences in binding between the three MATs.

Most of the synthetic chemistry and drug development involving MATs has been centred on DAT and SERT with a large number of selective modulators being tested, especially with the possibility of therapeutic benefit in treatment of depression [13,14]. There are a wide variety of structures that have been tested, including rigid tropane-based cocaine analogues [6,15,16], mazindol analogues [5] and lobeline analogues [17], as well as more flexible structures such as GBR-12909 piperidine derivatives [18]. Much less effort has been spent on the NET with selective modulators mainly restricted to the rigid tropane and tricyclic decane derivatives, with a few exceptions. A series of papers describes the synthesis of GBR compounds with high affinities for all three MATs [19–23] and is of particular interest to this study.

In this work we have constructed two sets of pharmacophores for each of the MATs, the first set based on GBR-12909 derivatives and the second set based on tropane derivatives. We chose to use single training sets that incorporated ligands selective for each of the MATs, rather than individual training sets for each MAT. Individual training sets would mean six separate training sets and we believe there is sufficient data available to be incorporated into just two training sets. This also means that variability of data between different laboratories' assays as well as too much structural diversity was avoided. We constructed the two training sets based on two base structures, GBR-12909 derivatives and tropane derivatives. This approach was used, rather than covering all known MAT modulator structural classes, in order to improve the statistical validity of the pharmacophores. Database searching with the resultant pharmacophores then allows for structural diversity to be covered. By using more than one hypothetical answer (interaction pattern) for each MAT, a stepwise protocol can be used for *in silico* database searching.

2. Methods

2.1. General methodology

Pharmacophore analysis was conducted with the Catalyst® program, version 4.11 (Accelrys Inc., San Diego, CA, USA), run on a Silicon Graphics O2 workstation. Database searching was performed on the NCI2000 chemical database provided with the installation of Catalyst.

2.2. Training set selection

Training sets were compiled from published data describing the inhibition of MATs. K_i measurements were taken from the literature for the interactions with each individual MAT. K_i measurements were determined *via* the displacement of a radioactively labeled ligand for each MAT: GBR-12935 or WIN 35428 for DAT, paroxetine or citalopram for SERT and nisoxetine for NET. From the large amount of data available, a sample representative of a spread of inhibition values and structural variety was selected. Where possible the spread of inhibition values exceeded 3.5 orders of magnitude, as recommended to ensure a subtractive phase during HypoGen [24], for each MAT. This is because Catalyst defines the compound with the lowest K_i value (most active) as being the most important and penalizes against those compounds whose K_i value is more than 3.5 orders of magnitude higher than the most active compound. For NET and SERT hypothesis generation using GBR-12909 derivatives the subtractive phase was modified to consider all compounds with activities more than 2.5 orders of magnitude below the top compound. For DAT hypothesis generation using GBR-12909 derivatives, the two compounds with the lowest affinities had their value modified to artificially create a subtractive phase, as the activity spread was less than 2.5. Structural variety was ensured so that there was as little redundancy within the training set as possible. SinOne training set was constructed for the tropane-like compounds and one for the GBR-12909 derivatives, both with 32 compounds.

Conformers were generated for each compound in Catalyst using a 20 kcal/mol energy range, as recommended [25], using the ''Best'' search option. For compounds with unknown stereochemistry, conformers for both enantiomers were generated. Where stereochemistry was known, conformers were only generated for the enantiomer responsible for the highest activity. Catalyst uses the poling algorithm to sample the conformational space effectively [26–28]. For each compound the number of conformers was less than the maximum number of 255, indicating that the conformational space had been effectively sampled within the energy range.

2.3. Hypothesis generation

Pharmacophore hypotheses were generated with the HypoGen or HypoRefine algorithms within Catalyst. The HypoGen algorithm includes three phases: (1) ''Constructive Phase'', where hypotheses are generated; (2) ''Subtractive Phase'', where inactive compounds are penalized against and (3) ''Optimization Phase'', where simulated annealing is used to improve the fits of the hypotheses [24]. The HypoRefine algorithm allows the addition of excluded volume features to penalize against steric interactions causing a reduction in activity [25]. Hypotheses were generated with the possibility of two feature combinations: (1) H-bond acceptor (HBA), H-bond donor (HBD), hydrophobic (aromatic) (Har), hydrophobic (aliphatic) (Hal) and positive charge/ionisable (PC/PI) or (2) HBA, HBD, ring aromatic (RA), hydrophobic (HY) and PC/PI. These combinations of features give a good coverage of the potential interactions of the training set as well as incorporating expected interactions such as with the basic nitrogen. The features HBA, HBD and RA are vectored features, in that there are two spheres per feature, representing both sides of the interaction and the direction between them. For each feature combination the possibility of excluded volumes was explored, as were variable spatial tolerances and variable weights of

individual features in order to fully explore the possibilities of the training set. As not all compounds had a reported experimental error with which to calculate the uncertainty value for Catalyst, uncertainty values of 2 or 3 were used in separate hypothesis generation runs. The use of different uncertainty values affects the Constructive Phase and gives the potential for different hypotheses to be constructed.

The ''Best'' two resultant hypotheses from all generated hypotheses for each MAT were selected as representative of the pharmacophore. It should be noted that the top-ranked hypothesis according to Catalyst is not necessarily the best. In some cases, where there is poor data, the null hypothesis is the top-ranked hypothesis. The best hypothesis was judged from a combination of, among others, statistical merit; whether the entire configurational space has been sampled; and the ability of the hypothesis to predict activities of compounds within a test set. Fischer randomization was performed on the selected hypotheses to generate negative control hypotheses. Nineteen hypotheses with randomly reassigned activity values were created for each selected hypothesis generation run, giving the statistical significance of the selected hypothesis at a 95% confidence level.

2.4. Database searching

The NCI2000 database, which contains 238,819 compounds, was searched within Catalyst using the Best searching algorithm. For each subtype and compound class, the database was searched with the best hypothesis. Where the resultant hitlist was large, the subsequent hitlists were searched with the second best hypothesis. Within each compound class, the number of unique hits per MAT was calculated using the Boolean operators provided by Catalyst.

3. Results and discussion

3.1. Quality assessment of pharmacophore hypotheses

Statistical cost of a hypothesis is determined with a number of parameters generated by Catalyst. The correlation between predicted activities and experimentally determined activities of the whole training set is described by the *R* measurement, where a value of 1.0 is perfect correlation. The correlation range for selected hypotheses based on tropanes is 0.69–0.84, for hypotheses based on GBR-12909 derivatives the range is 0.81–0.88. The configuration cost represents the entropy of the conformational space sampled during hypothesis generation. If the value is less than 17 then the entire space has been sampled, if it is greater than 17 then not all potential hypotheses have been explored and more favourable hypotheses may have been missed. All our hypotheses had configuration values less than 17. The difference between the fixed cost and the null hypothesis of the HypoGen run should exceed 60 in order for 90% probability that there is a real correlation with biological activity. The fixed cost is calculated from the combined error of estimation from predicting experimental activities, deviation of features from the ideal feature weight of 2.0, and the configuration cost. The null hypothesis contains no features and is simply the error of estimation of the training set, assuming that all compounds have the same activity and there is no structure–activity relationship. If the difference between the fixed cost and the null hypothesis is less than 40, then there is a less than 75% chance of there being a true correlation within the data. Our hypotheses have differences which range from 43 to 202.

The merit of hypotheses is further determined by various other analyses. To determine the subtype-selectivity of a pharmacophore, cross-correlations between training sets are performed. To do this, a hypothesis constructed, for example, for NET, is regressed against the same training set's activity data for the interaction with SERT. If the correlation is low then the hypothesis could be described as selective, if it is high, then it is most likely not going to be useful for retrieving NET-selective ligands from a database search. To determine the ability of a pharmacophore to accurately predict the activities of compounds outside the training set, a test set is constructed. If a pharmacophore is unable to accurately predict activities of compounds structurally similar to the training set, then it should be discarded. Ideally, the test set contains compounds from a variety of structural classes in order to ascertain the broadness

Table 1
Summary of pharmacophore statistics and composition

Pharmacophore	RMS	*R*	Feature composition	Statistical significance (%)
NET1a	0.91	0.88	HBA, HBD, 2 × Har, PC	90
NET1b	0.92	0.87	3 × RA, PC	95
NET2a	1.37	0.75	HY, 2 × RA, XVOL	70
NET2b	1.40	0.74	HBD, HY, 2 × RA, XVOL	70
SERT1a	1.23	0.81	HBA, 3 × Har, PC	95
SERT1b	1.00	0.88	HY, 2 × RA, PC	95
SERT2a	2.00	0.84	HBA, Hal, 2 × Har, PI, XVOL	95
SERT2b	2.04	0.83	HBA, 2 × HY, PI, RA, XVOL	95
DAT1a	1.45	0.86	HBA, 2 × Har, PC, XVOL	90
DAT1b	1.40	0.87	HY, 2 × RA, PC, XVOL	80
DAT2a	1.46	0.69	HBA, Hal, 2 × Har, XVOL	85
DAT2b	1.47	0.69	HBA, 3 × HY, XVOL	30

RMS = root mean square, *R* = correlation, statistical significance is at 95%. Pharmacophores constructed from GBR-12909 derivatives are referred to as 1a and 1b, pharmacophores constructed from tropane derivatives are referred to as 2a and 2b.

of the pharmacophore's predictability. The results of our cross-correlations and test set analysis are discussed in detail below.

Fischer randomization is used to reassign the activities within a training set to further analyse the biological significance of a hypothesis. By creating new spreadsheets containing the randomly reassigned activity data, the statistical significance of each original hypothesis can be determined. We used a 95% confidence level for determining the statistical significance of our pharmacophores. The statistical significance can be seen in Table 1, for pharmacophores generated from GBR-12909 derivatives the significance levels are 90% or more for all but the DAT1b pharmacophore. For tropane pharmacophores the significance levels are lower, with the range being 70–95% for the first hypothesis and 30–95% for the second hypothesis. SERT pharmacophores exhibited significance levels of 90–95%, NET pharmacophores 70–95% and DAT pharmacophores 30–90%. It was found that the pharmacophores exhibiting statistical significance below 85% contained the HY, RA, HBA, HBD and PC feature combination as well as fewer of the more stringent vectored features, which may be an indication of the usefulness of more stringent features, at least for conferring statistical merit.

3.2. Hypothesis analysis

3.2.1. Comparison of GBR-12909-based pharmacophores

Fig. 1 shows the training set of GBR-12909 derivatives used to create pharmacophores for each of the three MATs. The table displays the K_i values used for the three transporters. All three top hypotheses, DAT1a, SERT1a and NET1a consist of the same basic combination of features: 2 hydrophobic aromatic (Har) features, 1 positive charge (PC) feature and a hydrogen bond acceptor (HBA). Fig. 2 shows compound **20** mapped to the three top hypotheses in similar orientations. All three hypotheses have a similar arrangement, with the directionality of the HBA and inter-feature distances differing slightly. When mapped to **20**, the PC and HBA feature map to the central part of the compound (Fig. 2). SERT1a contains an additional Har feature which maps to the indole ring. Interestingly, DAT1a has an excluded volume (XVOL) in the location corresponding to the indole ring in the mapping of the SERT1a model. This result indicates that the same chemical substituents that increase the ability of ligands to interact favourably with SERT diminish the interaction with DAT. In the same approximate region as the third Har feature in SERT1a, NET1a exhibits a hydrogen bond donor (HBD) feature.

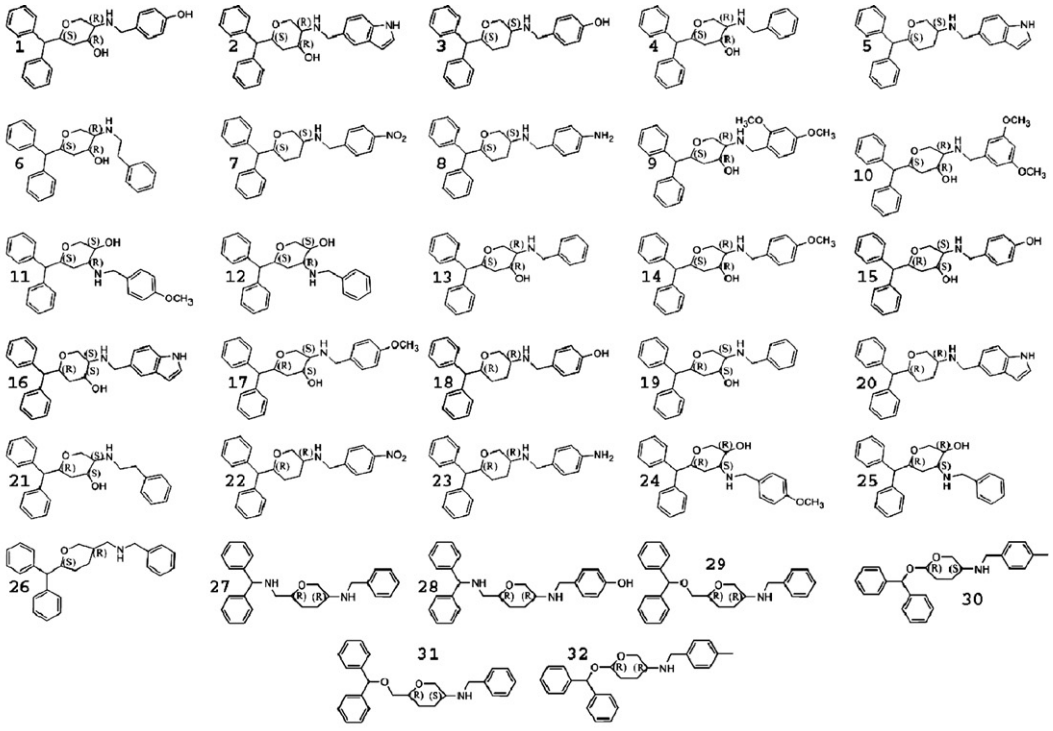
Zhang et al. [23] suggested a structure–activity relationship (SAR) for their GBR-12909 derivatives with respect to each of the MATs. They proposed that the common pharmacophore for the interaction with all MATs was the two adjacent aromatic groups which correspond to the two conserved Har features in each of our pharmacophores. Our common pharmacophore also has a PC feature corresponding to the positively ionisable nitrogen and an HBA feature mapping to the oxygen of the central pyranyl ring. Zhang et al. [23] also suggested that a hydroxyl substituent on the ring is important for interaction with NET and SERT but not DAT, however our pharmacophores do not reflect this. They did however agree with the suggestion that a third Har feature is important for SERT and NET interaction, with a further H-bond forming substituent of the aromatic ring important for interaction with NET. Of the top 10 hypotheses returned for NET, all contained the HBD feature. Of the top 10 hypotheses returned for SERT, three contained the third Har feature and the remaining seven contained a HBD replacement, similar to the best NET pharmacophore. This similarity in the different suggested possibilities for the SAR of ligands at SERT and NET is reflected in the cross-correlation analysis of the training sets, as discussed below.

One method of judging a pharmacophore's merit is its ability to predict the activities of individual compounds within the training set. For DAT1a the top five compounds have errors of prediction of less than 2 which is a very acceptable result. The error of prediction is a ratio between the experimentally determined K_i and the K_i estimated by Catalyst. A good model should be able to predict the training set compounds to within a factor of 10. The only compounds with an error of prediction greater than 10 are **12** and **24**, both of which had their activity values modified to allow a subtractive phase. For SERT1a the top three compounds have errors of less than 2.5, the fourth top compound, **8**, has an error of 12. Two other compounds have errors greater than 10, however neither are top compounds. For NET1a the largest error of 10 occurs for a compound with moderate activity. The second top compound, **13**, has an error of 6.4 due to it lacking a HBD in the appropriate position to map to the pharmacophore.

3.2.2. Comparison of tropane-based pharmacophores

Fig. 3 shows the compounds used to create the tropane-based pharmacophores for all three MATs. The same training set was used for each MAT, with the appropriate K_i value as shown in the table. The best DAT hypothesis DAT2a consists of 2 Har features, 1 hydrophobic aliphatic (Hal) feature, 1 HBA and an XVOL. The HBA is situated adjacent to the Hal feature, between it and one of the Har features. The second Har feature is situated near the first Har feature and further away from the HBA feature. The XVOL is located adjacent to the second Har feature. Fig. 4 shows the top compound mapped onto the MAT hypotheses.

The best SERT hypothesis SERT2a has a similar arrangement of hydrophobic and HBA features as DAT2a. However the spacing between the features is different and the directionality of the HBA feature is different. SERT2a also contains a positive ionisable (PI) feature which is unique among the 2a hypotheses. SERT2a also contains an XVOL in a different location to the XVOL in DAT2a. The best NET hypothesis is quite different to DAT2a and SERT2a in that the statistically best combination of features involved only hydrophobic (HY) and ring aromatic (RA) features rather than the more stringent Hal and Har with HBA combination. NET2a consists of two RA features, a HY feature and an XVOL. The arrangement of the HY and RA features is similar to the arrangement of the Hal and Har features in the DAT and SERT hypotheses, however the XVOL is in a position similar to the location of the HBA in each of the other two hypotheses.



Compound	K_i DAT(nM)	K_i SERT(nM)	K_i NET(nM)
1[23]	172	237	10.4
2[23]	120	15.3	2.1
3[23]	85	37.7	5.1
4[21]	125	607	12.2
5[23]	214	14.5	6.6
6[21]	317	3210	406
7[23]	34.6	19.9	54.3
8[23]	62.4	16.1	12.6
9[21]	528	20.8	95
10[21]	100	19.7	101
11[21]	611	793	686
12[21]	(300000)	9560	296
13[23]	446	707	4.9
14[23]	115	25.9	15.8
15[23]	91.5	2540	94.4
16[23]	184	223	13.2
17[21]	110	669	412
18[23]	67.4	187	40.9
19[21]	327	5000	1050
20[23]	142	11.8	15.6
21[21]	162	2650	71
22[23]	59.9	64.6	81.6
23[23]	43.9	45.7	25.6
24[21]	(300000)	3420	1700
25[21]	716	3970	2890
26[22]	159.7	3956.5	3108.7
27[22]	124.8	3210.1	1348.7
28[22]	177.5	1328.1	485.5
29[22]	77.2	8992	1202.9
30[22]	158.2	1492.7	1529.9
31[22]	122.5	8992	1299.8
32[22]	229.5	2676.9	262.7

Fig. 1. Training set with activities at MATs for GBR-12909-based hypothesis generation. Numbers in brackets are the assigned activities rather than the actual reported activities [21–23].

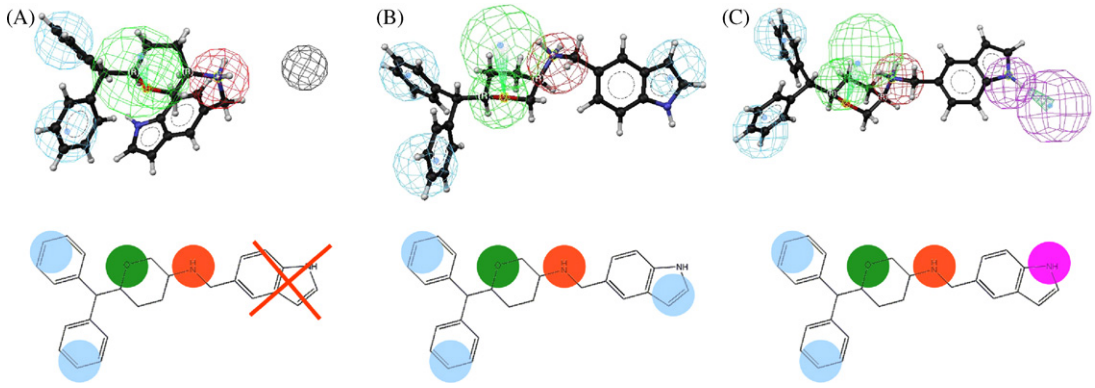


Fig. 2. Compound **20** mapped onto MAT pharmacophores: (A) DAT1a, (B) SERT1a, (C) NET1a. Chemical features are represented by coloured spheres which describe the location constraint for each feature: Har—light blue; Hal—light blue; HBA—green; PC—red; HBD—purple; XVOL—grey. Top row: 3D representation of mapping. Atoms are colour coded (C—black, N—blue, O—red, H—white). Bottom row: 2D representation of mapping. The red X indicates where the excluded volume would map to if compound **20** was in a similar conformation to the mappings for SERT1a and NET1a.

In regard to the ability of the pharmacophores to predict their own training set, the tropane-based pharmacophores are worse than the GBR-12909-based pharmacophores. DAT2a has 5 compounds with errors greater than 10, including the second top compound, **42**, which has an error of 12. The poor estimation of compound **42** is surprising given its similarity to the top compound **60**. It may be due to the presence of the methyl group attached to the protonated nitrogen, which is lacking in compound **60**. The position of the methyl group in all generated conformers of **42** is situated in such a way as to make it sterically difficult for both the HBA and Hal features to be mapped simultaneously. For NET2a the top five compounds have an error of less than 5, with 4 compounds having an error of prediction greater than 10. SERT2a has 8 compounds with large errors, including the top compound **38** with an error of 15. Despite the large error, the order of the compounds within the training set is correct with regard to the fit score *i.e.* the top compound is estimated to have the best activity, the second top compound the second best activity, *etc*. In general, the pharmacophores with lower statistical significance perform worse at predicting their own training set than those with higher statistical significance, as can be seen from the *R* values and statistical significances in Table 1.

3.2.3. Comparison of tropane- and GBR-12909-based pharmacophores

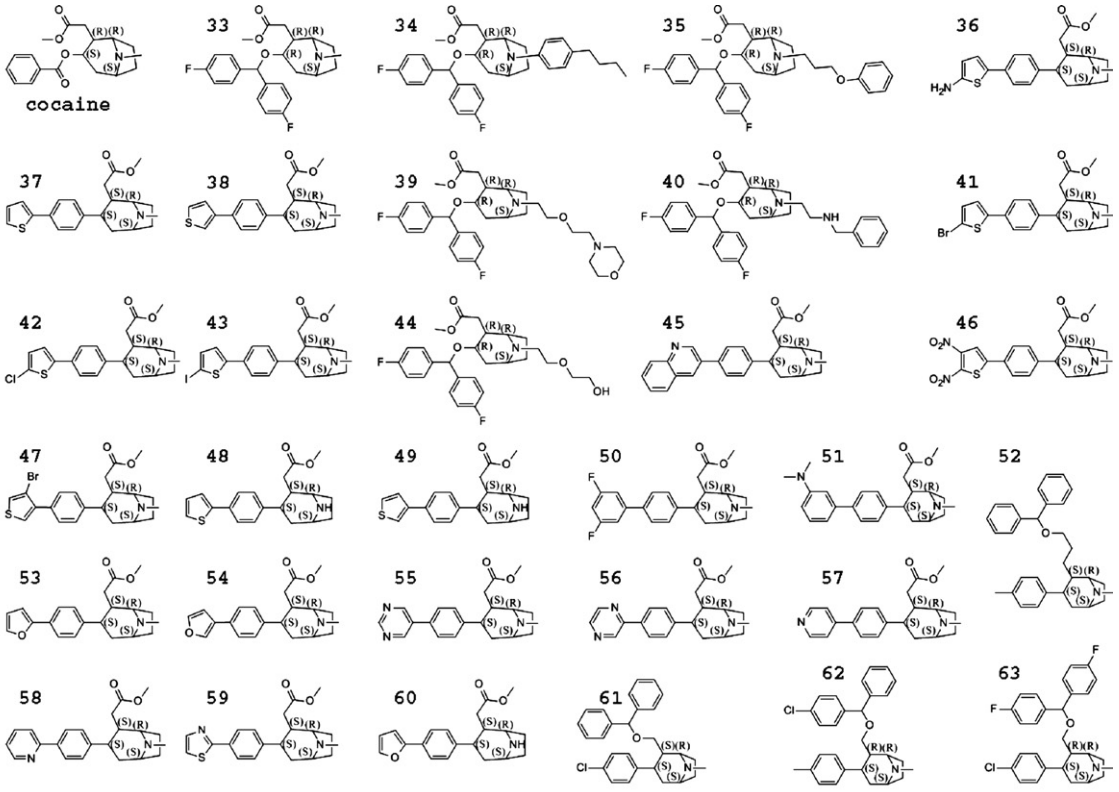
As the two sets of pharmacophores were constructed from disparate sets of chemical structures, one set consisting of fairly flexible GBR-12909 derivatives with a central pyranyl ring, one with a conformationally constrained tropane ring system, it is interesting to compare the resultant pharmacophores. Fig. 5 shows the overlay of DAT (5A) and SERT (5B) pharmacophores, NET pharmacophores could not be overlaid due to the difference in constituent features. Fig. 5A shows that DAT1a and DAT2a have two Har groups and an HBA group in common, situated in a similar orientation. The positive feature of DAT1a is missing in DAT2a and DAT2a has a Hal feature that DAT1a does not have. Fig. 5B shows the overlay of SERT1a and SERT2a. The similarity between SERT1a and SERT2a is much the same as between DAT1a and DAT2a in that there is an alignment of two Har features and a HBA feature. Although both SERT pharmacophores have a positive feature, it cannot be aligned while maintaining the alignment of the highest number of features possible. SERT1a has a further Hal feature while SERT2a has an XVOL located near the positive feature and one of the Har features.

3.2.4. Fitness/statistical analyses

Cross-correlation analysis was performed on the two training sets individually. For the GBR-12909-based training set, DAT1a was regressed against the activity values used to create SERT1a, and then regressed against the activity values used to create NET1a. This was then repeated for SERT1a and NET1a, with regression against the appropriate activity data. In this way Table 2 representing the cross-correlation was constructed. It can be seen that the highest correlation values are for each pharmacophore with its own training set. Interestingly, both SERT1a regressed against the NET training set and NET1a regressed against the SERT training set revealed an *R* value of over 0.7, suggesting relatedness between the two transporters' affinities for compounds within the training set. This relationship was not found with the DAT. It may be expected that as NET1a and SERT1a have the same basic features, a correlation will exist. However DAT1a contains the same basic features as the other two hypotheses, but a much lower correlation exists. We have previously found that despite two pharmacophores being similar, a correlation between the training sets did not exist [29].

The cross-correlations for the tropane-based pharmacophores did not indicate the same relatedness as for the GBR-12909-based pharmacophores. The self-correlations for NET2a and DAT2a were lower than the corresponding GBR-12909-based pharmacophores, while SERT2a had a reasonably high *R* value of 0.84. The cross-correlations were also lower than in the analysis of the GBR-12909-based pharmacophores and there was no obvious pattern indicating relatedness between activities at two or more transporters.

To determine the predictivity of pharmacophores for compounds outside their own training set, a test set containing a diverse range of over 100 compounds with activity at two or



Compound	K_i DAT(nM)	K_i SERT(nM)	K_i NET(nM)
Cocaine[31]	89	1050	3320
33[6]	11.8	3260	610
34[6]	8.51	376	2210
35[6]	703	7730	12100
36[31]	10000	64.7	30000
37[31]	52	0.15	158
38[31]	12.1	0.017	189
39[6]	1620	23000	28100
40[6]	36	1090	1130
41[31]	6.43	0.38	324
42[31]	4.42	0.64	311
43[31]	22.1	4.56	1137
44[6]	42.6	14300	3650
45[32]	59.8	20	30000
46[31]	30000	5000	10000
47[31]	183	4.02	10000
48[31]	12.2	0.11	75.3
49[31]	6.4	0.23	39
50[32]	31.7	0.41	109
51[32]	545	1000	30000
52[15]	139	61	207
53[32]	7.14	1.13	1400
54[32]	14.2	0.35	127
55[32]	344	33.1	10000
56[32]	121	11.3	213
57[32]	1566	1000	30000
58[32]	44	161	10000
59[32]	520	1.67	10000
60[32]	3.76	0.15	12.7
61[15]	52	147	1190
62[15]	937	1001	22500
63[15]	250	724	52300

Fig. 3. Training set with activities at MATs for tropane-based hypothesis generation [6,15,31,32].

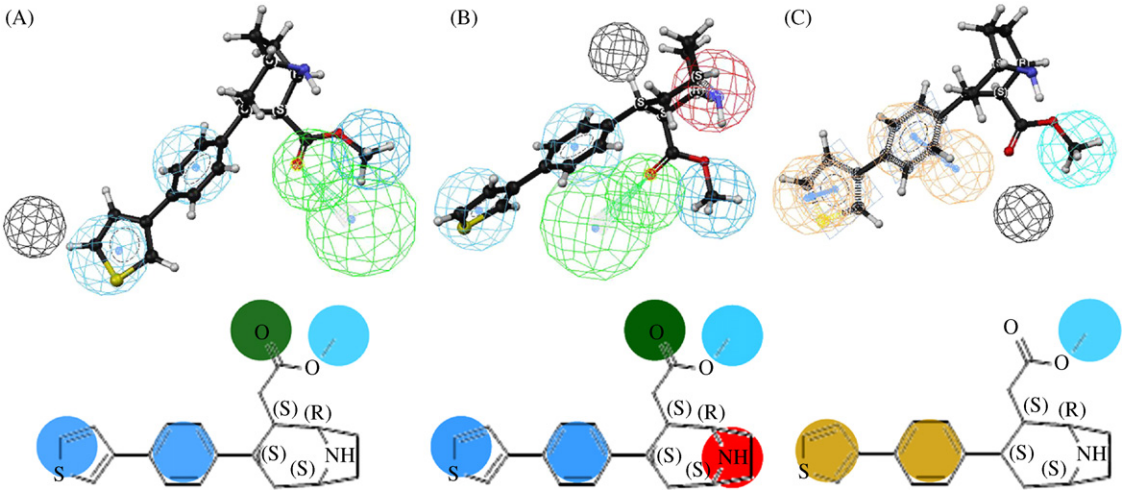


Fig. 4. Compound **49** mapped onto MAT pharmacophores: (A) DAT2a, (B) SERT2a, (C) NET2a. See Fig. 2 for explanation of colours. Additional features: HY—light blue; RA—brown.

more of the MATs was constructed from literature data. Each compound of the test set was fitted to the pharmacophores and the error of estimation was calculated. Fig. 6 shows the structures for representative compounds of the test set and errors of estimation from the prediction of activity by each of the pharmacophores. It can be seen that the activities of some compounds are predicted very well on the whole, *e.g.* **65**, **73**, while some are predicted quite poorly *e.g.* β-CIT, **69**. A benefit of having more than one pharmacophore for each MAT is evidenced by the prediction of compounds **79** and **80**. While the SERT1a and NET1a pharmacophores mis-predict the activities by factors of >10,000 and >100, the prediction of activities by the tropane-based pharmacophores is near perfect for SERT and good for NET. This indicates that when trying to describe the wide variety of structures with activity at the MATs, it is better to have more than one description of their SAR. This may also prove useful in the reduction of false negatives in database searching, given the experience with compounds **79** and **80**.

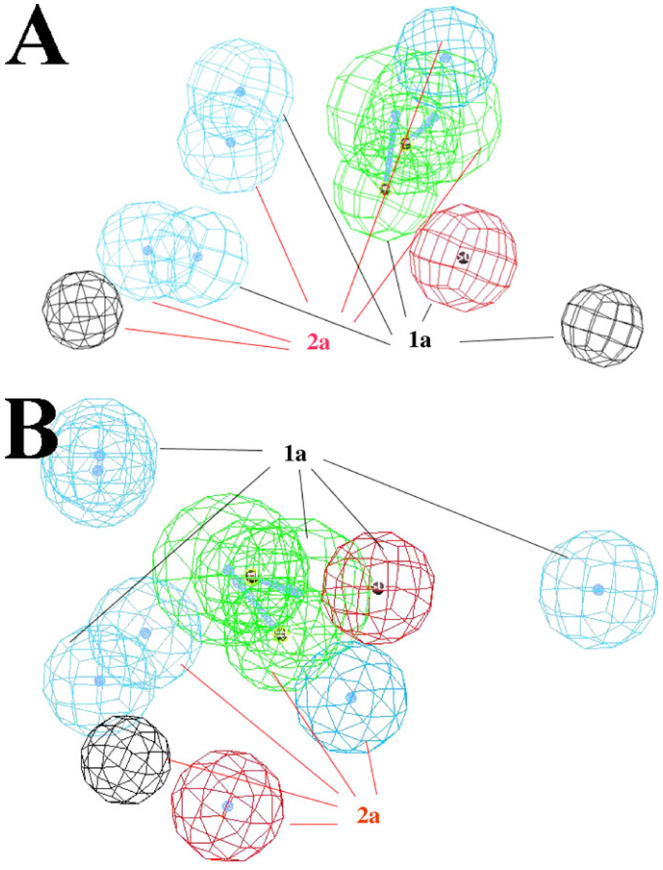


Fig. 5. Overlay of (A) DAT1a and DAT2a, (B) SERT1a and SERT2a. Chemical features are represented by coloured spheres which describe the location constraint for each feature: Har—light blue; Hal—dark blue; HBA—green; PC—red; XVOL—grey.

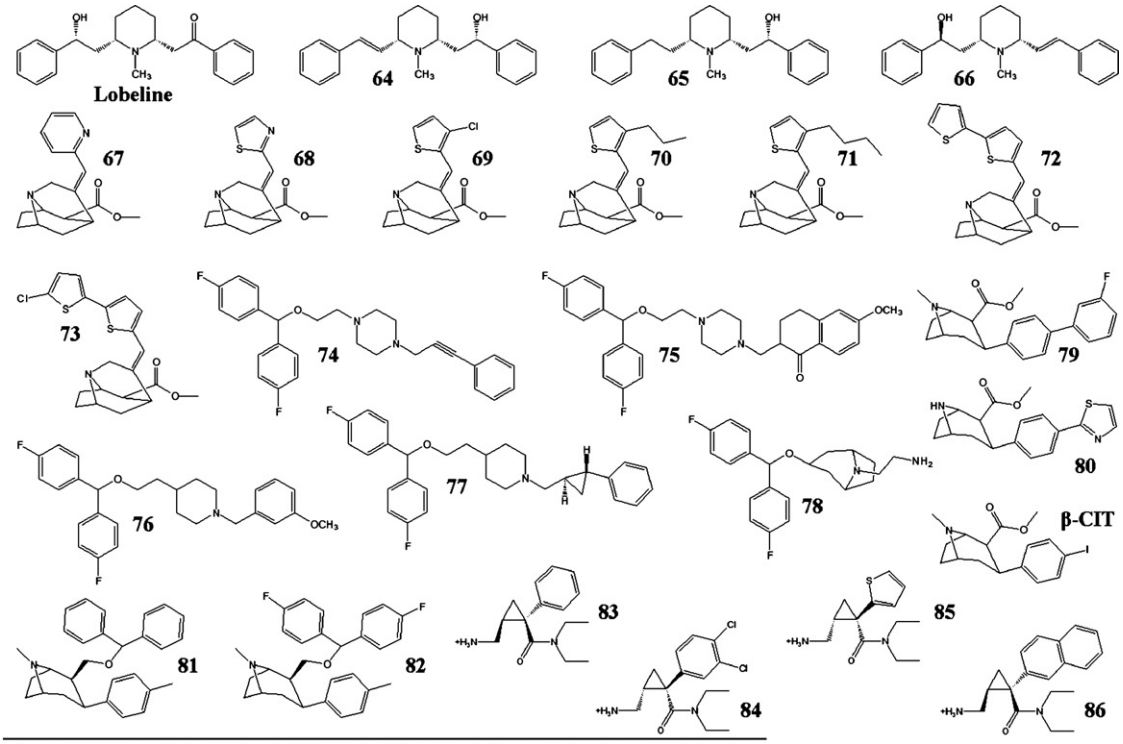
3.3. Database searching

Our pharmacophores were used to search the NCI2000 compound database within the Catalyst program. Each pharmacophore was used to initially screen the 200,000+ compound library. Table 3 summarises the number of returned hits from the initial searches. Searches with the GBR-12909-based pharmacophores returned manageable numbers of hits immediately, with the return ranging from 14 to 580. Searches with the tropane-based pharmacophores returned far too many hits to be easily dealt with, hitlists ranged from 2552 to 49,027 hits. In these instances a second screen was performed on the

Table 2
Cross-correlations of pharmacophores with training sets for other MATs

Pharmacophore	Training set		
	NET	SERT	DAT
NET1a	0.88	0.72	0.38
SERT1a	0.75	0.88	0.38
DAT1a	0.35	0.37	0.86
NET2a	0.75	0.63	0.62
SERT2a	0.45	0.84	0.42
DAT2a	0.57	0.44	0.69

Separate training sets were used for 1a and 2a pharmacophore generations. Numbers are *R* correlation values.



Name	DAT1a	SERT1a	NET1a	DAT2a	SERT2a	NET2a
Lobeline	-35.0	-22.0		-45	-9.3	
64[17]	1.3	-1.1		-2.3	1.1	
65[17]	1.1	1.2		1	1.3	
66[17]	19.0	-4.0		5.6	-5.7	
67[33]	1.2	-4.9	1.1	1.0	-3.6	2.1
68[33]	-2.9	-16.0	-4.2	-3.7	-11.0	-3.4
69[33]	19.0	12.0	17.0	12.0	9.4	25.0
70[33]	1.5	2.4	-4.7	1.2	3.9	-2.6
71[33]	2.2	94.0	7.3	-1.6	100	7.0
72[33]	-1.4	4.3	2.0	3.3	1.8	-4.6
73[33]	-1.1	3.4	1.8	3.3	1.4	-1.5
74[34]	-1.5	-2.2		1.7	-1.2	
75[35]	-5.0	8.4		-3.4	26.0	
76[36]	-1.3	-6.1		-2.1	-1.7	
77[37]	14.0	10.0		11.0	9.9	
78[6]	4.1	-1.7	-1.6	4.8	-14.0	1.1
79[32]	4.1	950000	670	1.8	-1.1	5.1
80[32]	-1.1	33000	210	-2.1	-1.2	2.0
81[15]	1.6	36.0	2.6	3.9	2.1	2.2
82[15]	-1.1	13.0	-1.8	-1.2	1.1	-1.2
83[38]	-1.9	-1.9	-4.6	-1.5	-1.1	-3.7
84[38]	5.8	1.1	2.9	2.4	-1.5	2.1
85[38]	5.2	2.6	8.8	35	2.6	9.9
86[38]	1.1	-1.4	-1.9	-1.8	-1.1	1.9
β-CIT	140	1000000	60000	190	650	1200

Fig. 6. Summary of the predictive power of top pharmacophores for a test set of compounds with various structures and activities. The values shown are the error of prediction given as a ratio between experimentally derived activities and activities predicted by Catalyst. A negative value indicates an over-prediction of activity *i.e.* the predicted activity is better than the actual activity. A positive value is an under-prediction of activity [6,15,17,32–38].

returned hitlist with the second best pharmacophore *i.e.* 2b for each MAT. Interestingly, for the SERT and DAT searches, there was no significant decrease in numbers of hits, meaning that the top two hypotheses fit almost all of the returned compounds despite the difference in feature composition. Following the pharmacophore screens the hitlists were cut down according to Lipinski's Rule of Five [30] (Screen 3). This resulted in a further refinement of the hitlists, in the most extreme case the NET1 hitlist was reduced to just a single compound. The final hitlists of potential lead compounds contained between

Table 3
Search result statistics for NCI2000 database in Catalyst

Pharmacophore	Screen 1		Screen 2		Screen 3		Unique	
	Hits	(%)	Hits	(%)	Hits	(%)	Hits	(%)
DAT1	580	0.24	–	–	351	0.15	336	0.14
SERT1	70	0.03	–	–	14	0.01	0	0
NET1	14	0.01	–	–	1	0.01	0	0
DAT2	21996	9.21	21783	9.12	16621	6.96	12570	5.26
SERT2	2552	1.07	2439	1.02	1718	0.72	240	0.10
NET2	49027	20.53	10644	4.46	6348	2.66	3250	1.36

See text for details of screens.

<0.01% and 6.96% of the original 238,819 compounds in the database.

In order to determine unique hitters from the database searches Boolean operators were used within Catalyst to discard compounds which were found in more than one hitlist. Following refinement of the hitlists derived with the GBR-12909 derivative-based pharmacophores with Lipinski's Rule of Five [30], there were 336 unique DAT hits, but no unique hits for SERT or NET. For the hitlists retrieved with the tropane-based pharmacophores, there were 12,570 unique DAT hits, 3250 unique NET hits and 240 unique SERT hits.

Previous searches of the NCI2000 database by another group have also turned up unique DAT modulators [9–11]. Encouragingly, several analogues of the compounds selected by this group and found to be DAT modulators (**87**, **88**) were found by our search using DAT1a (Fig. 7), despite the pharmacophore being based on GBR-12909 derivatives rather than cocaine and its analogues [11]. In particular an analogue of **87**, which was found to have nanomolar K_i at DAT [9], was found by our search (NCI0078685). Table 4 shows the predicted and actual affinities of four previously tested compounds and three of our retrieved analogues fitted to the DAT2a pharmacophore. DAT2a gave smaller errors of prediction for the previously tested compounds than DAT1a.

Table 4
Actual affinity at DAT (μM) and predicted affinities (μM) at MATs for compounds retrieved from database searches

Compound	Actual	Predicted		NET
	DAT	DAT	SERT	
87	2.3	1.6	3900	0.96
88	0.2	19	1.5×10^6	9.5
89	>1.0	2.1	630	8.5
90	0.61	0.015	0.023	5.0
NCI0078685	–	1.6	1.5	0.68
NCI0340038	–	0.16	2.1	9.4
NCI0400726	–	0.07	1.0	5.0

The analogues retrieved *via* database searching were fitted to the pharmacophore and were found to have at least equal predicted activity, in one case a predicted activity 5-fold better than the tested compound.

Each compound was also fitted to the SERT2a and NET2a pharmacophores to predict their activities at these MATs, summarised in Table 4. Compounds **89** and **90** are predicted to be at least slightly DAT selective by the 2a pharmacophores, while **87** and **88** are predicted to have an affinity order NET > DAT > SERT, with both highly favouring NET and DAT over SERT. Of the retrieved compounds analysed here,

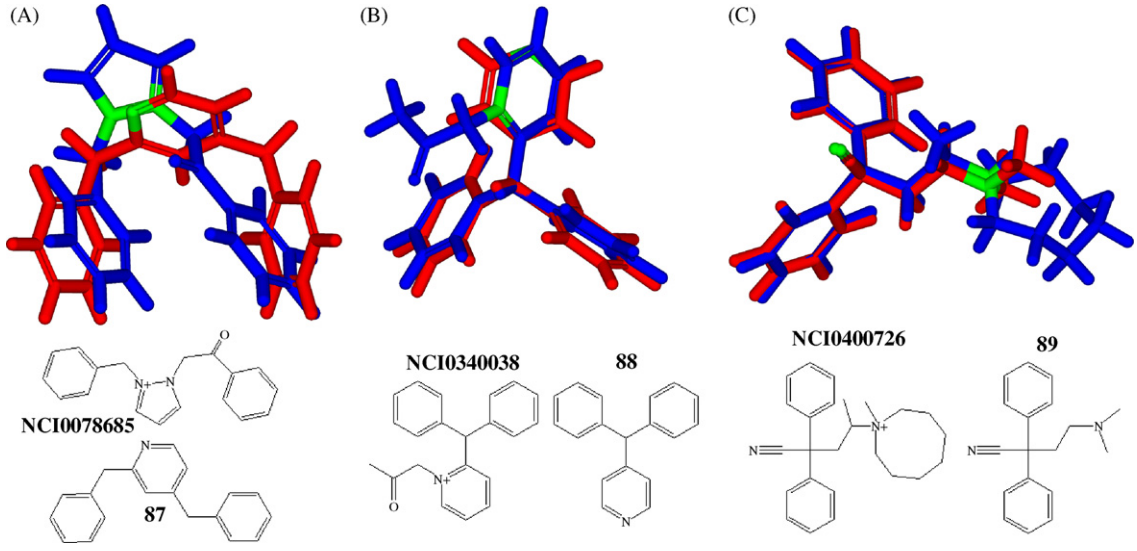


Fig. 7. (A) Overlay of NCI0078685 retrieved from NCI2000 database with the DAT1a pharmacophore (blue) and DAT inhibitor tested by Enyedy et al. (red). (B) Overlay of retrieved NCI0340038 (blue) and DAT inhibitor tested by Enyedy et al. (red). (C) Overlay of retrieved NCI0400726 (blue) and DAT inhibitor tested by Enyedy et al. (red). Nitrogens are shown in green.

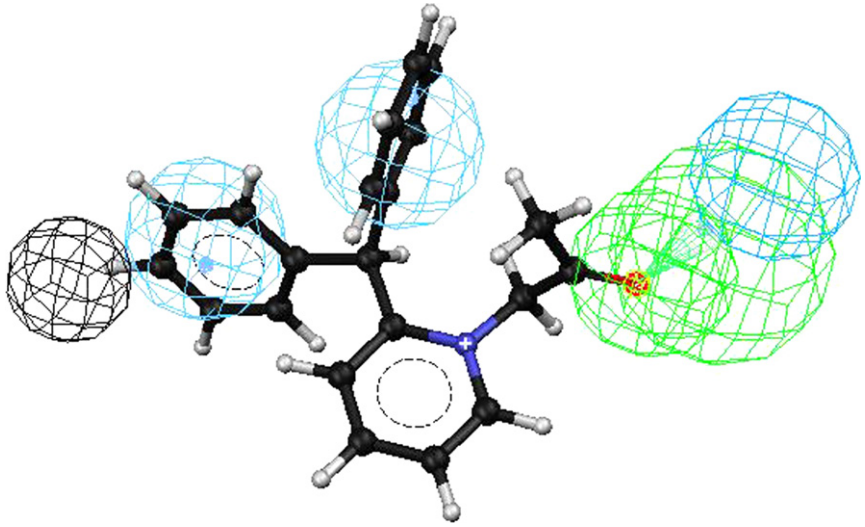


Fig. 8. NCI0340038 fitted to the DAT2a pharmacophore, predicted activity 1.6 μM.

NCI0078685 is predicted to be fairly unselective between MATs while NCI0340038 and NCI0400726 are both predicted to be 100-fold selective for DAT. These retrieved analogues may well be worth testing to see if their activities are indeed better than those of the compounds tested by Enyedy et al., or used in direct synthesis. Fig. 8 shows the mapping of our retrieved unique, potentially selective DAT ligand NCI0340038, a substituted pyridine, onto the DAT2a pharmacophore. It can be seen that the compound satisfies the two Har features and the HBA. The analogue retrieved by Enyedy et al., fits one Har feature and the HBA in the mapping with the best predicted activity.

4. Conclusions

We have constructed a number of pharmacophores describing the interaction of compounds at the dopamine, serotonin and noradrenaline transporters. We developed the pharmacophores from derivatives of two base structures: GBR-12909 and tropane. Using different feature and weight combinations we were able to explore more of the potential pharmacophoric description of interaction at these sites than would otherwise be possible. By constructing pharmacophores for the three transporters concurrently we have been able to establish differences and similarities between binding modes.

The GBR-12909 derivative pharmacophores provided a similar description of the SAR as that provided by Zhang et al. [22], at all three MATs. They also indicated a connection between affinity towards SERT and NET with similar orientations and combinations of features describing the interactions at these two sites. The DAT pharmacophore was similar to the other two in the basic set up of 4 features, however the presence of an excluded volume where a 5th feature is positioned in the NET and SERT pharmacophores indicates that the same substituents that enhance affinity at the SERT and NET diminish affinity at the DAT. Of the pharmacophores constructed from tropane derivatives, the SERT and DAT hypotheses have similar arrangements of hydrophobic and HBA features, with the directionality of the HBA differing. The SERT pharmacophore also contains a positive ionisable feature. The best NET hypothesis has fewer, and a different combination of, features which map to the same functional groups of the compounds.

The pharmacophores were used to search the NCI2000 database in Catalyst. The GBR-12909 derivative-based pharmacophores provided for quite a stringent search of the database, with very manageable numbers of hits returned, while the tropane-based searches required a multiple step procedure to reduce the numbers of hits. Of the resultant hits that were analysed, quite a few compounds were of the same structural classes as previously reported MAT modulators. The returned hits, including some previously unreported structural classes of potential MAT modulators, should be investigated further with biological testing. This would reveal whether the retrieved compounds do indeed exhibit the expected selectivity profiles as determined by application of our pharmacophores.

One of the advantages of constructing multiple pharmacophores, using different structural classes in the training sets, is demonstrated by the application of test compounds. If a database containing the compounds **79** and **80** had been searched only with the MAT1 pharmacophores they would have been discarded as potential modulators. However, by using the MAT2 pharmacophores, the affinities of **79** and **80** were correctly estimated, thereby reducing the numbers of false negatives that would be missed in a database search.

References

[1] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, et al., Structure and mechanism of the lactose permease of *Escherichia coli*, Science 301 (2003) 610–615.

[2] Y. Huang, M.J. Lemieux, J. Song, M. Auer, D.-N. Wang, Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*, Science 301 (2003) 616–620.

[3] A.W. Ravna, I. Sylte, S.G. Dahl, Molecular model of the neural dopamine transporter, J. Comput. Aided Mol. Des. V17 (2003) 367–382.

[4] A.W. Ravna, I. Sylte, S.G. Dahl, Molecular mechanism of citalopram and cocaine interactions with neurotransmitter transporters, J. Pharmacol. Exp. Ther. 307 (2003) 34–41.
[5] S.S. Kulkarni, A.H. Newman, W.J. Houlihan, Three-dimensional quantitative structure–activity relationships of mazindol analogues at the dopamine transporter, J. Med. Chem. 45 (2002) 4119–4127.
[6] S.S. Kulkarni, P. Grundt, T. Kopajtic, J.L. Katz, A.H. Newman, Structure–activity relationships at monoamine transporters for a series of *N*-substituted 3α-(bis[4-fluorophenyl]methoxy)tropanes: comparative molecular field analysis, synthesis, and pharmacological evaluation, J. Med. Chem. 47 (2004) 3388–3398.
[7] H. Yuan, A.P. Kozikowski, P.A. Petukhov, CoMFA study of piperidine analogues of cocaine at the dopamine transporter: exploring the binding mode of the 3α-substituent of the piperidine ring using pharmacophore-based flexible alignment, J. Med. Chem. 47 (2004) 6137–6143.
[8] D.L. Roman, C.C. Walline, G.J. Rodriguez, E.L. Barker, Interactions of antidepressants with the serotonin transporter: a contemporary molecular analysis, Eur. J. Pharmacol. 479 (2003) 53–63.
[9] I.J. Enyedy, S. Sakamuri, W.A. Zaman, K.M. Johnson, S. Wang, Pharmacophore-based discovery of substituted pyridines as novel dopamine transporter inhibitors, Bioorg. Med. Chem. Lett. 13 (2003) 513–517.
[10] I.J. Enyedy, W.A. Zaman, S. Sakamuri, A.P. Kozikowski, K.M. Johnson, et al., Pharmacophore-based discovery of 3,4-disubstituted pyrrolidines as a novel class of monoamine transporter inhibitors, Bioorg. Med. Chem. Lett. 11 (2001) 1113–1118.
[11] S. Wang, S. Sakamuri, I.J. Enyedy, A.P. Kozikowski, O. Deschaux, et al., Discovery of a novel dopamine transporter inhibitor, 4-hydroxy-1-methyl-4-(4-methylphenyl)-3-piperidyl 4-methylphenyl ketone, as a potential cocaine antagonist through 3D-database pharmacophore searching. molecular modeling, structure–activity relationships, and behavioral pharmacological studies, J. Med. Chem. 43 (2000) 351–360.
[12] L. Orus, S. Perez-Silanes, A.M. Oficialdegui, J. Martinez-Esparza, J.C. Del Castillo, et al., Synthesis and molecular modeling of new 1-aryl-3-[4-arylpiperazin-1-yl]-1-propane derivatives with high affinity at the serotonin transporter and at 5-HT1A receptors, J. Med. Chem. 45 (2002) 4128–4139.
[13] D. Spinks, G. Spinks, Serotonin reuptake inhibition: an update on current research strategies, Curr. Med. Chem. 9 (2002) 799–810.
[14] M.J. Millan, The role of monoamines in the action of established and ''novel'' antidepressant agents: a critical review, Eur. J. Pharmacol. 500 (2004) 371–384.
[15] L. Xu, S.S. Kulkarni, S. Izenwasser, J.L. Katz, T. Kopajtic, et al., Synthesis and monoamine transporter binding of 2-(diarylmethoxymethyl)-3β-aryltropane derivatives, J. Med. Chem. 47 (2004) 1676–1682.
[16] F. Bois, R.M. Baldwin, N.S. Kula, R.J. Baldessarini, M. Al Tikriti, et al., Synthesis and monoamine transporter affinity of 2β-carbomethoxy-3β-(4′-*p*-substituted phenyl)-piperidine analogs of cocaine, Bioorg. Med. Chem. Lett. 16 (2006) 5222–5225.
[17] G. Zheng, D.B. Horton, A.G. Deaciuc, L.P. Dwoskin, P.A. Crooks, Desketo lobeline analogs with increased potency and selectivity at dopamine and serotonin transporters, Bioorg. Med. Chem. Lett. 16 (2006) 5018–5021.
[18] T. Prisinzano, E. Greiner, E.M. Johnson, C.M. Dersch, J. Marcus, et al., Piperidine analogues of 1-[2-[bis(4-fluorophenyl)methoxy]ethyl]-4-(3-phenylpropyl)piperazine (GBR 12909): high affinity ligands for the dopamine transporter, J. Med. Chem. 45 (2002) 4371–4374.
[19] S. Zhang, M.E.A. Reith, A.K. Dutta, Design, synthesis, and activity of novel *cis*- and *trans*-3,6-disubstituted pyran biomimetics of 3,6-disubstituted piperidine as potential ligands for the dopamine transporter, Bioorg. Med. Chem. Lett. 13 (2003) 1591–1595.
[20] S. Zhang, J. Zhen, M.E.A. Reith, A.K. Dutta, Structural requirements for 2,4- and 3,6-disubstituted pyran biomimetics of *cis*-(6-benzhydryl-piperidin-3-yl)-benzylamine compounds to interact with monoamine transporters, Bioorg. Med. Chem. 12 (2004) 6301–6315.
[21] S. Zhang, J. Zhen, M.E.A. Reith, A.K. Dutta, Discovery of novel trisubstituted asymmetric derivatives of (2*S*,4*R*,5*R*)-2-benzhydryl-5-benzylaminotetrahydropyran-4-ol, exhibiting high affinity for serotonin and norepinephrine transporters in a stereospecific manner, J. Med. Chem. 48 (2005) 4962–4971.
[22] S. Zhang, J. Zhen, M.E.A. Reith, A.K. Dutta, Design, synthesis, and preliminary SAR study of 3- and 6-side-chain-extended tetrahydro-pyran analogues of *cis*- and *trans*-(6-benzhydryl-tetrahydropyran-3-yl)-benzylamine, Bioorg. Med. Chem. 14 (2006) 3953–3966.
[23] S. Zhang, F. Fernandez, S. Hazeldine, J. Deschamps, J. Zhen, et al., Further structural exploration of trisubstituted asymmetric pyran derivatives (2*S*,4*R*,5*R*)-2-benzhydryl-5-benzylamino-tetrahydropyran-4-ol and their corresponding disubstituted (3s, 6s) pyran derivatives: a proposed pharmacophore model for high-affinity interaction with the dopamine, serotonin, and norepinephrine transporters, J. Med. Chem. 49 (2006) 4239–4247.
[24] H. Li, J. Sutter, R. Hoffmann, HypoGen: an automated system for generating 3D predictive pharmacophore models, in: Pharmacophore Perception, Development, and Use in Drug Design (Book), International University Line, La Jolla, CA, 2000, pp. 171–187.
[25] Catalyst 4.11 User Guide, Accelrys Inc., San Diego, CA, USA, 2006.
[26] A. Smellie, S.D. Kahn, S.L. Teig, Analysis of conformational coverage. 1. Validation and estimation of coverage, J. Chem. Inf. Comp. Sci. 35 (1995) 285–294.
[27] A. Smellie, S.D. Kahn, S.L. Teig, Analysis of conformational coverage. 2. Applications of conformational Models, J. Chem. Inf. Comp. Sci. 35 (1995) 295–304.
[28] A. Smellie, S.L. Teig, P. Towbin, Poling: promoting conformational variation, J. Comput. Chem. 16 (1995) 171–187.
[29] I.J.A. MacDougall, R. Griffith, Selective pharmacophore design for α1-adrenoceptor subtypes, J. Mol. Graph. Model. 25 (2006) 146–157.
[30] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 23 (1997) 3–25.
[31] G. Tamagnan, D. Alagille, X. Fu, N.S. Kula, R.J. Baldessarini, et al., Synthesis and monoamine transporter affinity of new 2β-carbomethoxy-3β-[4-(substituted thiophenyl)]phenyltropanes: discovery of a selective SERT antagonist with picomolar potency, Bioorg. Med. Chem. Lett. 15 (2005) 1131–1133.
[32] G. Tamagnan, D. Alagille, X. Fu, N.S. Kula, R.J. Baldessarini, et al., Synthesis and monoamine transporter affinity of new 2β-carbomethoxy-3β-[aryl or heteroaryl]phenyltropanes, Bioorg. Med. Chem. Lett. 16 (2006) 217–220.
[33] J. Zhou, T. Klaβ, A. Zhang, K.M. Johnson, C.Z. Wang, et al., Synthesis and pharmacological evaluation of (*Z*)-9-(Heteroarylmethylene)-7-azatricyclo[4.3.1.03,7]decanes: thiophene analogues as potent norepinephrine transporter inhibitors, Bioorg. Med. Chem. Lett. 13 (2003) 3565–3569.
[34] L.-W. Hsin, T. Prisinzano, C.R. Wilkerson, C.M. Dersch, R. Horel, et al., Synthesis and dopamine transporter affinity of chiral 1-[2-[bis(4-fluorophenyl)methoxy]ethyl]-4-(2-hydroxypropyl)piperazines as potential cocaine abuse therapeutic agents, Bioorg. Med. Chem. Lett. 13 (2003) 553–556.
[35] D. Lewis, Y. Zhang, T. Prisinzano, C.M. Dersch, R.B. Rothman, et al., Further exploration of 1-{2-[bis-(4-fluorophenyl)methoxy]ethyl}piperazine (GBR 12909): role of *N*-aromatic, *N*-heteroaromatic, and 3-oxygenated *N*-phenylpropyl substituents on affinity for the dopamine and serotonin transporter, Bioorg. Med. Chem. Lett. 13 (2003) 1385–1389.
[36] E. Greiner, T. Prisinzano, E.M. Johnson, C.M. Dersch, J. Marcus, et al., Structure–activity relationship studies of highly selective inhibitors of the dopamine transporter: *N*-benzylpiperidine analogues of 1-[2-[bis(4-fluorophenyl)methoxy]ethyl]-4-(3-phenylpropyl)piperazine, J. Med. Chem. 46 (2003) 1465–1469.
[37] E. Greiner, T.L. Boos, T.E. Prisinzano, M.G. DeMartino, B. Zeglis, et al., Design and synthesis of promiscuous high-affinity monoamine transporter ligands: unraveling transporter selectivity, J. Med. Chem. 49 (2006) 1766–1772.
[38] H. Roggen, J. Kehler, T.B. Stensbol, T. Hansen, Synthesis of enantiomerically pure milnacipran analogs and inhibition of dopamine, serotonin, and norepinephrine transporters, Bioorg. Med. Chem. Lett. 17 (2007) 2834–2837.