J Comput Aided Mol Des (2006) 20:209–225
DOI 10.1007/s10822-006-9046-2

ORIGINAL PAPER

Hierarchical clustering analysis of flexible GBR 12909 dialkyl piperazine and piperidine analogs

Kathleen M. Gilbert · Carol A. Venanzi

Received: 2 January 2006 / Accepted: 20 April 2006 / Published online: 20 July 2006

Abstract Pharmacophore modeling of large, drug-like molecules, such as the dopamine reuptake inhibitor GBR 12909, is complicated by their flexibility. A comprehensive hierarchical clustering study of two GBR 12909 analogs was performed to identify representative conformers for input to three-dimensional quantitative structure–activity relationship studies of closely-related analogs. Two data sets of more than 700 conformers each produced by random search conformational analysis of a piperazine and a piperidine GBR 12909 analog were studied. Several clustering studies were carried out based on different feature sets that include the important pharmacophore elements. The distance maps, the plot of the effective number of clusters versus actual number of clusters, and the novel derived clustering statistic, percentage change in the effective number of clusters, were shown to be useful in determining the appropriate clustering level. Six clusters were chosen for each analog, each representing a different region of the torsional angle space that determines the relative orientation of the pharmacophore elements. Conformers of each cluster that are representative of these regions were identified and compared for each analog. This study illustrates the utility of using hierarchical clustering for the classification of conformers of highly flexible molecules in terms of the three-dimensional spatial orientation of key pharmacophore elements.

K. M. Gilbert · C. A. Venanzi (✉)
Department of Chemistry and Environmental Science, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA
e-mail: venanzi@adm.njit.edu

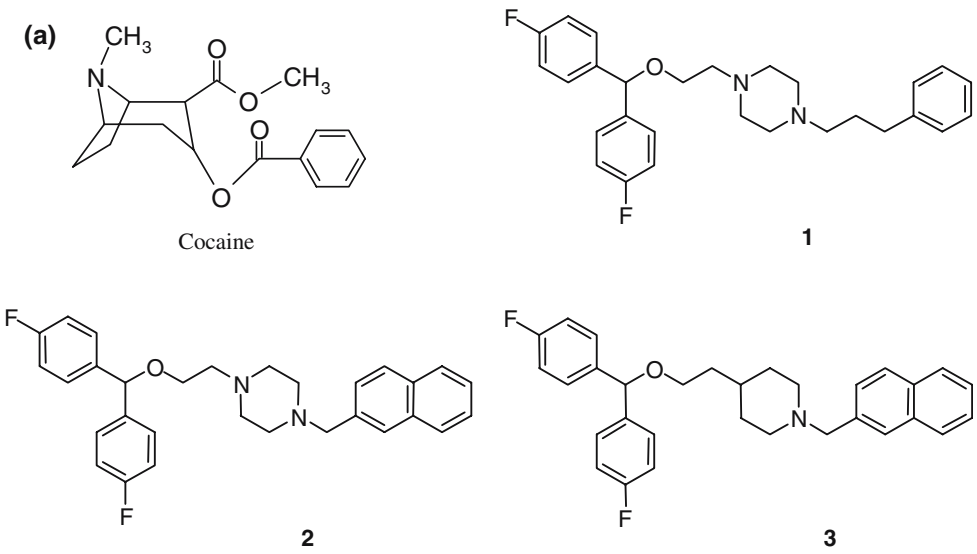
Keywords Clustering · Cocaine · Conformational analysis · Dopamine reuptake inhibitor · GBR 12909 · Hierarchical · Molecular mechanics · Pharmacophore modeling · XCluster

Abbreviations

%ΔEff	percentage change in the effective number of clusters
3D-QSAR	three-dimensional quantitative structure–activity relationship
CoMFA	Comparative Molecular Field Analysis
DA	dopamine
DAT	dopamine transporter
FRC	fuzzy relational clustering
QSAR	quantitative structure-activity relationship
RMSD	root mean square deviation
RTB	number of rotatable bonds
SAR	structure–activity relationship
SERT	serotonin transporter

Introduction

Cocaine addiction is a world-wide problem. The dopamine hypothesis [1] suggests that the biochemical side effects of cocaine are due to the inhibition of dopamine reuptake by the binding of cocaine (Fig. 1a) to the dopamine transporter (DAT), resulting in an excess of dopamine in the synaptic cleft. Chronic use of cocaine leads to depletion of dopamine in the reward centers of the brain, contributing to the potential for relapse from cocaine withdrawal. A treatment agent is sought which would inhibit the binding of cocaine to



Compound	DAT K_i (nM)	SERT K_i (nM)	-log [DAT K_i /SERT K_i]	Reference
1	3.7	126	1.53	[15]
2	8	312	1.59	[12]
3	0.71	229	2.51	[15]

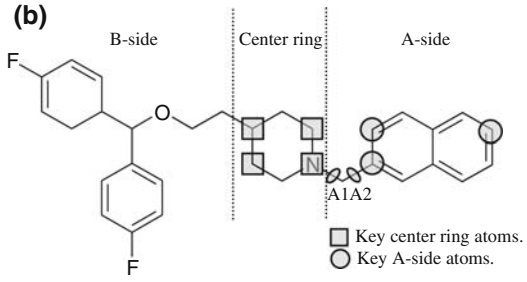


Fig. 1 (**a**) Structures of cocaine, GBR 12909, and analogs: GBR 12909 **1**, the piperazine **2**, and the piperidine **3**; DAT and SERT binding and selectivity data. (**b**) Diagram of **3** and its features. Identified are pharmacophore elements, key torsional angles and key atoms

the DAT, yet allow some reuptake of dopamine. Several different classes of compounds based on a variety of structural frameworks (tropane, benztropine, methylphenidate, mazindol), have been investigated as dopamine reuptake inhibitors potentially useful in the treatment of cocaine abuse [2]. Most dopamine reuptake inhibitors contain the common pharmacophore features of a basic nitrogen and aromatic ring. Homology modeling of the DAT, a protein with 12 transmembrane α-helices, followed by cocaine–DAT docking studies identified several specific interactions. For example, the ester group of cocaine forms hydrogen bonds with two tyrosine residues, the benzene ring stacks with a phenylalanine residue, the tropane ring forms hydrophobic interactions with two other phenylalanine residues, and the protonated nitrogen interacts with an aspartate residue [3]. In addition, mutational studies of the residues in the second transmembrane segment of the DAT indicate that this region is not part of the cocaine-binding site, but rather contributes indirectly to the binding of cocaine through a conformational change [4].

GBR 12909 (**1**, shown in Fig. 1a) is a promising agent for use in agonist substitution therapy for cocaine addiction, having demonstrated a decrease in cocaine-maintained responding without affecting food-maintained responding in behavioral studies of rhesus monkeys [5]. In addition, prior use of cocaine caused cross-sensitization for several dopamine (DA) reuptake inhibitors, but not for **1** [6]. Compound **1** and related dialkyl piperazines and piperidines have also been shown to bind selectively to the cocaine binding site on the DAT and cause minimal DA reuptake inhibition, without causing other of cocaine's deleterious side effects. GBR 12909 also has been found to reduce the ability of cocaine to increase extracellular DA levels [7]. Phase I clinical studies on the safety of **1** have been completed [8], with further studies underway [9].

These positive attributes have led to structure–activity relationship (SAR) studies to improve the DAT binding affinity of analogs of **1**, as well as their selectivity for the DAT compared to the serotonin transporter (SERT), and are summarized in a recent review [9]. Even more rotationally restricted analogs of **1** have been synthesized recently, although they are less closely-related to **1** than the more flexible analogs: one set includes an exocyclic nitrogen and the other a bicyclic center ring structure [10, 11].

Flexible analogs of **1** consist of three major parts, indicated by the piperidine model in Fig. 1b:

- Heterocyclic center ring, usually a piperazine or piperidine ring
- Alkyl/single aromatic group (A-side)
- Ether–bisphenyl group (B-side).

The A-side and its neighboring nitrogen contain molecular features that are key pharmacophore elements found in most classes of DA reuptake inhibitors: a basic nitrogen and aromatic ring. The B-side functional groups (ether oxygen and bisphenyl moiety) are not common to all DA reuptake inhibitors; rather, they are found only in analogs of **1** and benztropine. The B-side appears to have a role in the binding of these analogs to the DAT [12], but its role has not been determined [9]. SAR studies have addressed changes to each of the three areas. Of the center ring alterations, most studies have focused on either piperazine or piperidine analogs of **1**. These two subtypes appear to have somewhat different SAR [13–15]. Considerably more SAR studies have targeted the A-side, compared to the other two regions. For this reason, and because the present work is a precursor companion paper to a modeling study of a set of **1** analogs with only A-side changes, the present clustering study uses feature sets defined by the key pharmacophore elements of the A-side and its proximal nitrogen. A piperidine and a piperazine analog of **1** were selected from a series developed by the Rice group at the National Institutes of Health [9]. The piperazine **2** [12] and the piperidine **3** [15], both shown in Fig. 1a, are identical except for their center rings and are less flexible than **1**, yet demonstrate better DAT/SERT selectivity than **1** (see selectivity data in Fig. 1). Analog **3** has a methyne group in the central ring to which the B-side is connected instead of the nitrogen found in **1** and **2**. These analogs are interesting because they have fewer rotatable bonds than **1** and no chiral centers.

In the absence of an X-ray crystal structure of the DAT, ligand-based three-dimensional quantitative structure-activity relationship (3D-QSAR) techniques such as Comparative Molecular Field Analysis (CoMFA) [16] may be useful in identifying molecular features that improve DAT binding affinity and DAT/SERT selectivity. However, these techniques require the use of a template conformer upon which each 3D-QSAR model is based. For rigid molecules that can adopt only a limited number of conformers, selection of the template conformer is relatively straightforward. However, large, flexible molecules such as **1**, which can take on a continuum of low energy conformations, represent a challenge to the application of 3D-QSAR techniques. As a result, molecular modeling studies to date have been carried out on fairly rigid classes of dopamine reuptake inhibitors, such as tropanes [17–28], piperidine-based cocaine analogs [29], benztropine [30–33], 1-[1-(2-benzo[*b*]thienyl)cyclohexyl]piperidine (BTCP) [34], bupropion [35], mazindol [36], methylphenidate [37–39], novel piperadinols [40], and more rigid bicyclic **1** analogs [10]. In contrast, the present work uses hierarchical clustering to select a variety of representative conformers of the flexible GBR 12909 analogs **2** and **3** to be used as templates for 3D-QSAR studies of closely-related analogs of **2** and **3**.

Although there is considerable evidence that many ligands do not bind to proteins in their vacuum phase global energy minimum (GEM) conformation [41–45], many 3D-QSAR models of DA reuptake inhibitors have been based on the GEM structure of the ligand, or on a few structures very close in energy to the GEM [21, 36, 46]. Other work has shown that considering conformations other than the GEM is important in pharmacophore modeling [47–53]. For example, Venanzi and coworkers carried out a conformational analysis study of methylphenidate [39] and found that several local energy minimum conformers (with energy within several kcal/mol of the GEM) closely match those of a rigid methylphenidate analog that has the same DAT binding affinity as methylphenidate. Since the rigid analog is assumed to contain the important pharmacophore elements in their binding (or bioactive) orientation, this supports the idea that the bioactive conformer need not be the GEM conformer.

Since CoMFA results are very sensitive to the conformation used as the template, it is important to explore the conformational potential energy surface of **2** and **3** and use clustering techniques to select conformers that are representative of the three-dimensional spatial orientation of the key pharmacophore elements (basic nitrogen and aromatic ring). As a first step in this direction, a random search conformational analysis was carried out on both **2** and **3** and identified over 700 conformers of each analog. Then, a comprehensive

hierarchical clustering study of the conformers of **2** and **3** was performed to classify them into a manageable number of large clusters based on the similarity of the relative spatial orientation of their key pharmacophore elements. Hierarchical clustering was carried out using the XCluster program [54]. XCluster handles both the symmetry of chemical structures and the circularity of torsional angles, which makes it ideally suited to the analysis of molecular conformations. Two different feature sets, consisting of only torsional angles or only heavy atoms, and three different alignment options were used to cluster the analogs. Inspection of the distance maps and analysis of clustering statistics identified the most appropriate feature set/alignment option and the optimal clustering level. Representative conformers for each cluster were determined for use in a companion 3D-QSAR study of analogs of **2** and **3** with A-side substitutions. Although the bioactive conformer *may* be one of the representative conformers, computations alone cannot prove that point. Instead, confirmation of the bioactive conformer is frequently derived from pharmacological testing of rigid analogs, none of which are presently available for this SAR series. The goal of the present work is simply to find representative conformers that can be used in future work to develop a stable and predictive model for the biological activity of these analogs.

Methods

All calculations were performed on a SGI Origin 2000 with 20 300-MHz processors, 20 GB of memory, using the IRIX® 6.5.18m UNIX®-based operating system. Molecular modeling was carried out with SYBYL Version 6.9 (available from Tripos, Inc., St, Louis, MO) and Macromodel 8 and 9 (available from Schrödinger, Inc., Portland, OR).

Hierarchical clustering using versions 8 and 9 of XCluster [54], a Macromodel module, was performed to determine representative conformers of **2** and **3** from the large data sets produced by random search conformational analysis. Various feature set/alignment options were examined and the resulting distance maps and clustering statistics were analyzed to determine the most appropriate option and clustering level. A representative conformer was selected from each of the major clusters found at the selected clustering levels for use in 3D-QSAR studies. The results of the extensive 3D-QSAR studies will be given in a separate publication.

Random search conformational analysis

The random search procedure was carried out on the eight non-ring torsional angles of **2** and **3** using SYBYL and the Tripos force field [55] with Gasteiger–Hückel charges. The energy-minimized structures of **2** and **3** were used as input to the random search procedure. Both molecules were modeled as protonated species, with each center ring kept as an aggregate in the chair conformation. Studies on DA binding have shown that it most likely binds to the DAT in the protonated state [56]. Cocaine is thought to exist as a 90/10 mixture of protonation states, favoring the protonated form, but having similar conformations in the neutral and protonated forms [20]. HF/6-31G* molecular orbital calculations performed on **1** showed that protonation of the A-side nitrogen is favored over the B-side nitrogen [57]. This agrees with binding studies of piperidine analogs of **1** that show that the binding affinity for the DAT is substantially higher when the A-side nitrogen rather than the B-side nitrogen is present [58]. The A-side nitrogens in **2** and **3** were therefore modeled with an associated proton.

The detailed protocol for the random search conformational analysis of **2** has been published elsewhere [59]. Briefly, the random search procedure randomly alters the eight torsional angles, then performs a geometry optimization (i.e. energy minimization) to produce a new conformer. The same protocol was repeated for **3**. A full analysis of the conformational profiles of **2** and **3** will be given in a future publication. The random search conformational analysis found 728 and 739 conformers for **2** and **3**, respectively.

Clustering options

Atoms and angles used in the feature set/alignment options are identified in Table 1. Distance matrices were constructed separately for **2** and **3** based on these options. Each entry in the distance matrix consists of the root mean square deviation (RMSD) of the atoms or torsional angles in the feature set calculated for pairs of conformers superimposed according to the chosen alignment option. The feature sets include key atoms or torsional angles that contain the DAT pharmacophore elements and are paired with different choices of atoms used for aligning the conformers by structural superposition.

Based on the selected feature set/alignment option, XCluster performs clustering by first calculating an intermolecular distance matrix of a set of superimposed molecules. For feature sets consisting of atoms only, XCluster allows the user to select a new super-

Table 1 Feature set/alignment options

Option	Feature set[a]	Alignment[b]
a	□ Center ring and ○ A-side key atoms	□ Center ring[c]
b	□ Center ring and ○ A-side key atoms	□ Center ring and ○ A-side key atoms
c	A1 and A2	Five heavy atoms defining A1 and A2[d]

[a] Feature set: The atoms or torsional angles used to calculate the intermolecular distances. Key atoms (identified by symbols noted) and torsional angles are shown in Fig. 1(b). RMSD calculations for the distance matrix were carried out on the feature set atoms in the corresponding alignment. Options *a* and *b* are atom-based clusterings; option *c* is a torsional angle-based clustering

[b] Alignment performed by XCluster unless otherwise noted

[c] Alignment performed in SYBYL

[d] As noted in the text, alignment of conformers for the torsional angle studies is not necessary to calculate RMSD values, but alignment was carried out to properly visualize clusters

position, or to use the molecular database's existing superposition. In the case of atom superpositions, each interconformational distance is a RMSD average between locations of the atoms in the feature set. Superpositions are not applicable to torsional angle-based clusterings; those interconformational distances are calculated based on torsional angle values, as opposed to atom locations. For torsional angles, the distance is the RMS of the differences between the torsional angle values, and is independent of atom locations. Issues such as circularity of torsional angle data and symmetry are handled automatically with user options available. Torsional angle periodicity and equivalency of certain carbons (such as *ortho* positions in symmetrical phenyl rings) were taken into account.

XCluster was used to calculate intermolecular atomic or torsional RMSD values. An XCluster command file was created for each study. Each distance value d_{ij} represents the similarity between conformers i and j by comparison of the items in the feature set. The critical distance d^* is incremented from 0 at clustering level 1 (where the number of clusters equals N, the number of conformers) to the minimum distance d_{ij} that is more than zero, to create clustering level 2 with $N - 1$ clusters. As the d^* value increases (in accordance with the increases in the actual d_{ij} in the intermolecular distance matrix), additional clusters are formed by combining two clusters found on the previous clustering level. Eventually, all of the conformers would be in the same cluster, the d^* value reaching its maximum value. XCluster is hierarchical; each cluster is an agglomeration of clusters formed on the previous level, and no cluster is disassembled at any succeeding level.

Selection of the optimal number of clusters

Two levels of review were necessary to select the optimal number of clusters. First, the distance maps for each feature set/alignment option were reviewed for both analogs, and those that showed small numbers of well-defined large clusters were selected for analysis. The percentage of conformers included in the large clusters was calculated and the maps that had the largest percentages were selected for calculation of clustering statistics. Then those selected distance maps along with various clustering statistics were used to determine the feature set/alignment option that gave the optimal number of large clusters. In addition, the radius of gyration was calculated to determine the coverage of conformational space by the selected clusters.

Distance maps

Distance maps are a color-coded display of the RMSD values for pairs of molecules in each analog's set of conformers. In the present case, the input order was not meaningful, so a generic ordering was used to display the distance maps. This procedure orders the conformers such that clusters form from adjoining conformers; multiple generic orderings are possible, depending on the first conformer selected. The distance maps shown in the present work are based on the generic ordering starting with conformer 1 on the left side.

The distance maps resulting from the three feature set/superposition options in Table 1 were reviewed for **2** and **3**. Distance maps which show optimal clustering are those which display several large, black or blue, well-defined squares of similar size *only* along the diagonal, indicating separation of the data set into well-defined clusters [54]. The number of large squares along the diagonal represents the number of major clusters. This type of distance map was selected for detailed analysis. In contrast, maps that show large, off-diagonal, black or blue elements indicate similarity between conformers in *different* clusters and represent poor clustering [54]. These, as well as maps with many

small squares along the diagonal, indicating many small clusters, were not analyzed further.

Clustering statistics

The XCluster run produces simple statistics by default; there are additional statistics which may be requested when the run is completed. Basic statistics are [54]:

- Clustering level, L: corresponding to $N - L + 1$ clusters, where N is the number of conformers.
- Critical distance, d^*_L: the critical threshold distance between the clustering level L and the previous clustering level, $L - 1$. The critical distances are derived directly from the distance matrix, and match the exact distance matrix entries. Thus, a sorted list of all distance matrix values would match a sorted list of critical distances for a particular clustering study.
- Number of clusters, k, equals $N - L + 1$, as noted above. This is also called the actual number of clusters, to differentiate it from the effective number of clusters. The actual number of clusters is the sum of all clusters, large and small.
- Effective number of clusters, k^*, corresponds to the number of large clusters, calculated as $k^* = \exp(S_{\mathrm{cl}})$, where:

$$S_{\mathrm{cl}} = -\sum_{i=1}^{k} x_i \ln x_i \tag{1}$$

and x_i equals the fraction of the data set in the cluster and k equals the actual number of clusters at the clustering level.

- Minimum separation ratio: The minimum separation ratio, R_{L}, is the ratio between the critical threshold distance of the current level, and the critical threshold distance of the next highest clustering level (the level with one less cluster):

$$R_{\mathrm{L}} = d^*{}_L / d^*_{L+1} \tag{2}$$

The effective number of clusters is a measure of how many large clusters exist at a particular clustering level. The effective number of clusters reaches its maximum when all clusters are of equal size. The maximum effective number of clusters for a given level is the actual number of clusters, which can be achieved at individual levels, but not maintained across all clustering levels. Qualitatively, the effective number of clusters measures how many ''large clusters'' are present at a particular clustering level for a clustering study, but the more uneven the sizes of the clusters, the lower the effective number of clusters. The minimum separation ratio shows how comparatively dissimilar the most similar clusters are in a set of clusters; this ratio is important when it rises above 2, indicating a clear separation between the most similar clusters [54]. These basic statistics are provided as part of the standard XCluster output file, although not all were used in the present study to determine the best clustering level.

Although the minimum separation ratio is noted by Shenkin and McDonald to be a figure of merit [54], it may not be useful in selecting the best clustering level for all data sets [54, 60]. Large, flexible molecules like the GBR 12909 analogs can have a continuum of conformations, such that the minimum separation ratio does not vary much with clustering level. As will be seen below, this is the situation encountered with the present data. As an additional measure of cluster separation, the present work defines the novel derived statistic, percentage change in the effective number of clusters (%ΔEff), which may be a more useful statistic for such data sets. The %ΔEff at level L is calculated as:

$$\%\Delta\mathrm{Eff} = (\mathrm{Abs}[k^*{}_L - k^*{}_{L-1}]/k_L) \times 100 \tag{3}$$

where %ΔEff = percentage change in the effective number of clusters; k^*_L = effective number of clusters at clustering level L; k^*_{L-1} = effective number of clusters at clustering level $L - 1$; k_L = actual number of clusters at clustering level L; Abs = absolute value of the term in square brackets.

Consecutive application of the distance map and clustering statistics criteria outlined above led to determination of the best feature set/alignment option and the optimal number of clusters.

Energy profile of data sets and clusters

To further explore the attributes of the identified clusters, an energy profile analysis was performed for each analog. For each data set of conformers from the random search conformational analysis, the percentage of conformers in various energy ranges was displayed by a color-coded pie chart; the percentage of each major cluster in various energy ranges was illustrated by color-coded histograms. The histograms were given in terms of cluster percentages instead of absolute number of conformers to facilitate comparison of clusters of varying size. This analysis was applied to each of the major clusters individually and to the minor clusters, taken as a group.

Identification and comparison of cluster representatives

Once the optimal feature set/alignment option and clustering level were chosen, the representative conformer for each cluster was determined as the conformer with feature set atom locations closest to the average feature set atom locations within the cluster. The representative conformers of **2** and **3** in each cluster were compared by calculating the RMSD values for the fit of the feature set atoms, as well as all heavy atoms. The comparison was accomplished by converting each conformer of **3** into a *pseudo* **2** structure by replacing the carbon atom with a nitrogen and removing the ''extra'' hydrogen. This allowed Database RMSD Fit calculations using SYBYL to be performed, which can only be done on a database of identical molecules. A visual comparison of the pairs of most closely-related **2** and **3** representative conformers was also made, using superimposed structures.

An analysis of the conformational coverage of both the identified clusters and the representative conformers was performed. The radii of gyration of the conformer sets and the clusters were calculated and compared. The numbers of conformers in each cluster were compared as well.

Results

Review of distance maps and clustering statistics

Distance maps created from the three clustering feature set/alignment combinations (see options *a*, *b*, and *c* of Table 1) are shown in Figs. 2 and 3. Small RMSDs, representing high similarity, are shown by the color black (where the RMSD is equal to zero) or blue. The colors graduate from blue for the most similar conformer pairs, through green to yellow to orange to red, representing the least similar conformer pairs, which have the highest RMSDs. The scale varies based on the RMSD range of the study; there are 10 divisions with the maximum values noted in the figure legends. The distance maps were reviewed to select those with a small number of large, well-defined clusters for further analysis. The distance map in Fig. 2a shows several large, off-diagonal blue regions, which group all but one of the six clusters with another cluster, indicating poor clustering. Similarly, the distance map in Fig. 3a shows many small clusters on the diagonal, as well as a few off-diagonal blue areas representing similarity between clusters. Neither map was selected for further analysis. Comparison of the distance maps for options *b* and *c* for

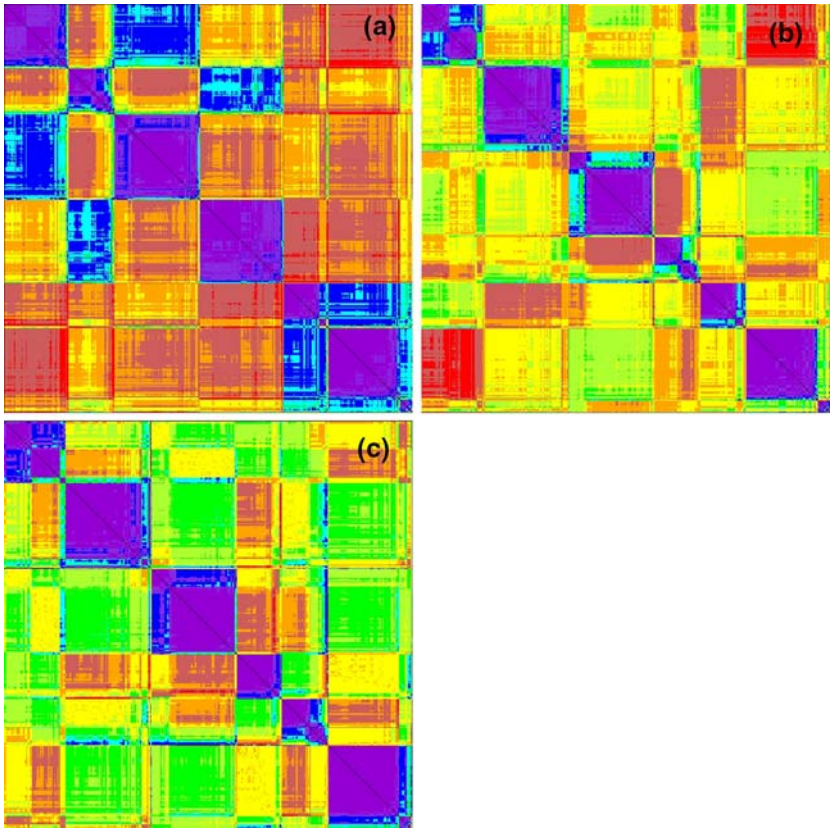


Fig. 2 Distance maps for clustering studies of **2**. (**a**) Feature set/alignment option *a*. Maximum distance equals 5.466 Å. (**b**) Feature set/alignment option *b*. Maximum distance equals 1.9127 Å. (**c**) Feature set/alignment option *c*. Maximum distance equals 181.3428°. Distance maps are color-coded by interconformational distances, in ascending order from black through blue to yellow to red as noted in the text. Blue boxes along the diagonal represent clusters of conformers; large squares contain many conformers, and small squares contain only a few conformers. Feature set/alignment options are identified in Table 1

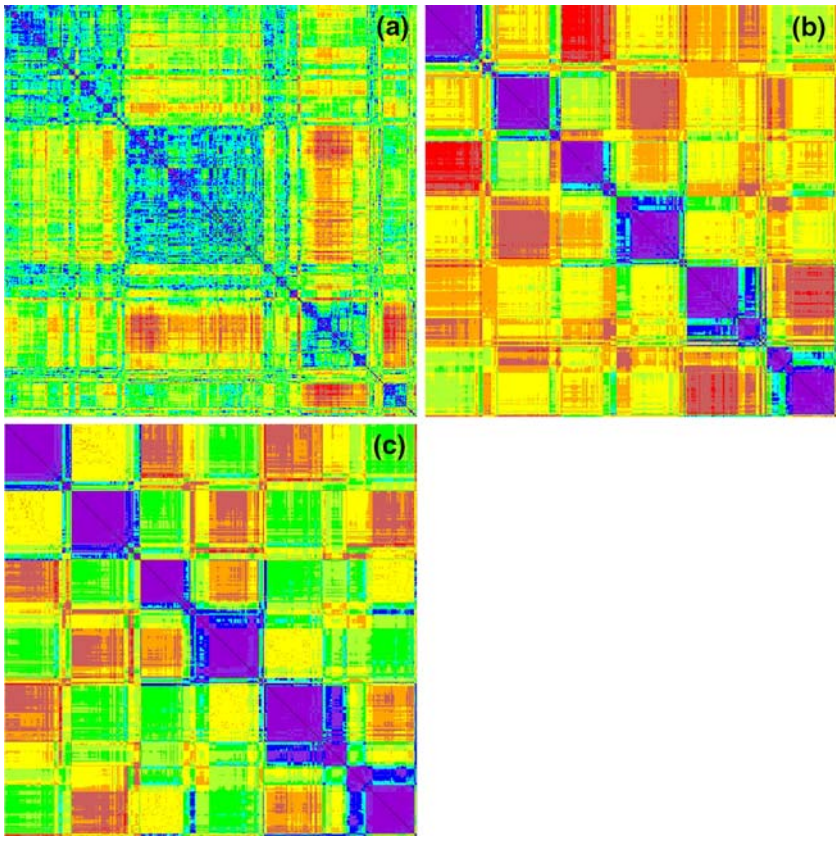


Fig. 3 Distance maps for clustering studies of **3**. (**a**) Feature set/alignment option *a*. Maximum distance equals 9.2589 Å. (**b**) Feature set/alignment option *b*. Maximum distance equals 1.9125 Å. (**c**) Feature set/alignment option *c*. Maximum distance equals 181.3208°. Distance maps are color-coded by interconformational distances, in ascending order from black through blue to yellow to red as noted in the text. Blue boxes along the diagonal represent clusters of conformers; large squares contain many conformers, and small squares contain only a few conformers. Feature set/alignment options are identified in Table 1

both analogs shows a strikingly similar clustering pattern of well-defined major clusters on the diagonal. These maps required further analysis to select the most promising feature set/alignment option.

Percentage of conformers in major clusters

The distance maps selected for analogs **2** (Fig. 2b, c) and **3** (Fig. 3b, c) all show six large clusters on the diagonal. From each map, the highest clustering level which resulted in six major clusters was selected to calculate the percentage of conformers in each major cluster. These results are reported in Table 2. The table shows that for **2**, option *c*, and **3**, option *b*, more than 99% of the conformers are included in the six major clusters. For **2,** option *b*, and **3**, option *c*, these values are 95% and 97%, respectively. These studies resulted in evenly populated clusters; the smallest of the major clusters for each feature set/alignment option contains 10% or more of the conformers in each data set. The largest minor clusters (selected from the clusters not included in the six major clusters) contain one-third or fewer conformers compared to the smallest major clusters.

Effective versus actual number of clusters

Plots of the effective number of clusters, k^*, versus the actual number of clusters, k, for the selected feature set/alignment options in Table 2 are given in Figs. 4 and 5 for the 20 highest clustering levels. The straight line where k equals k^* is also plotted for comparison. Again, the results are similar for both feature set/alignment options, although it is interesting to note where the large jumps in the effective number of clusters occur. The only clustering studies that separate into six effective large clusters ($k^* \cong 6$) within the 10 highest clustering levels, with the effective number of clusters closer to the actual number of clusters than the other studies listed in the table, are those for **2**, option *c*, and **3**, option *b*. This identifies options *c* and *b* as the most promising clustering studies for analogs **2** and **3**, respectively. However, because the present goal is to use the clustering studies to select representative conformers of **2** and **3** to be included in the *same* 3D-QSAR studies, selection of the same clustering feature set/alignment option is preferred. Option *c* was selected for both **2** and **3** because the conformer coverage of the six major clusters in option *c* for **2** is slightly better than that of option *b* for **3**, and the conformer coverage in option *c* for **3** is better than that of option *b*

Table 2 Percentage of conformers included at approximately six major clusters, selected feature set/alignment options

Analog (option)[a]	Clustering level	Actual number of clusters	Effective number of clusters	Smallest major cluster size	Largest minor cluster size	Total conformers not in major clusters
2(*b*)	717	12	6.81	10.4%	2.9%	4.7%
2(*c*)	719	10	6.02	11.0%	0.3%	0.7%
3(*b*)	731	9	6.23	13.4%	0.7%	0.9%
3(*c*)	729	11	6.62	13.3%	1.4%	3.0%

[a]Feature set/alignment options listed in Table 1

for **2** (3.0% not covered versus 4.7% not covered). Therefore, option *c* was used to confirm the optimum number of clusters, as detailed below.

Review of additional clustering statistics: minimum separation ratio and percentage change in effective number of clusters

Table 3 lists the 20 highest clustering levels for **2**, option *c*, in descending order, along with their associated data: minimum separation ratio, critical distance, actual and effective number of clusters, and %ΔEff. The highest minimum separation ratio is 1.23. Table 4 lists the same type of data for **3** and shows an even smaller range of minimum separation ratios for high clustering levels, the highest value being 1.17. Both values of minimum separation ratio fall significantly below the recommended value of 2 for identification of the optimal clustering level [54]. Therefore, it appears that the minimum separation ratio is not a useful criterion for these clustering studies.

The %ΔEff values were compared for the 20 highest clustering levels as well. This value is a numerical representation of the trends seen in Figs. 4 and 5, scaled by the number of actual clusters at the clustering level, as noted in Eq. 3. The %ΔEff is large for clustering levels where a large jump in the effective number of clusters occurs, as long as the actual number of clusters is low. Review of Table 3 identified clustering level 719 (approximately six effective clusters, 10 actual clusters, 13% change, minimum separation ratio 1.12) as the optimal clustering level for **2**, option *c*. Similarly, review of Table 4 identified clustering level 729 (approximately 6 effective clusters, 11 actual clusters, 11% change, minimum separation ratio 1.06) as optimal for **3**, option *c*.

Energy profile of the data sets and clusters

Figure 6a and b shows the distribution of relative energies of the conformers of **2** and **3**, respectively. Over half of the conformers of **2** have relative energies above 10 kcal/mol. Slightly over one-eighth of the conformers of **3** have relative energies in the same range. Forty percent of **3**'s conformers have a relative energy of 4 kcal/mol or less.

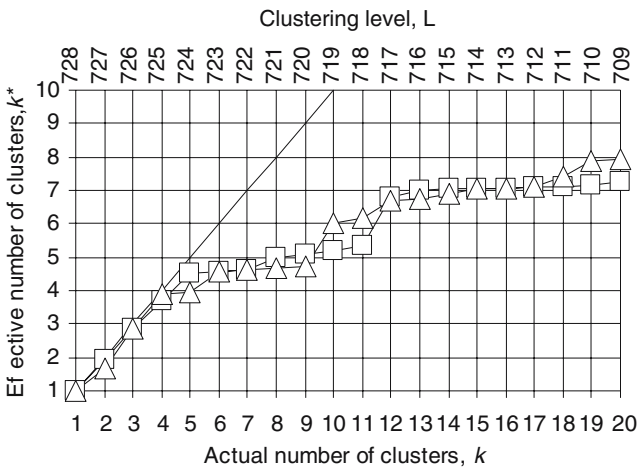


Fig. 4 Effective versus Actual Number of Clusters for **2**, feature set/alignment options *b* and *c*. Clustering level is noted on top *x*-axis for reference. Open square – Option *b*. Open triangle – Option *c*. Solid line – Theoretical maximum ($k = k^*$)

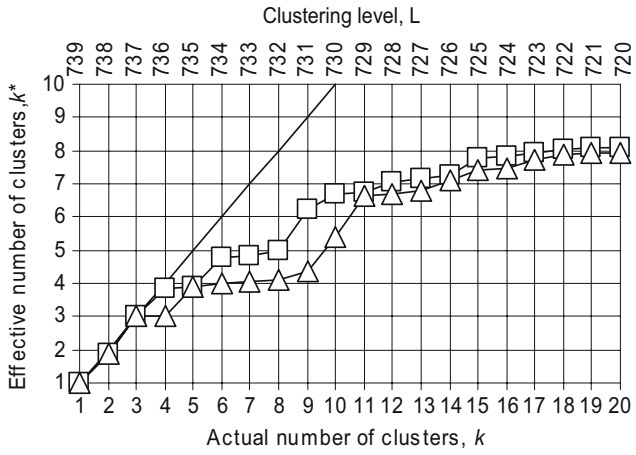


Fig. 5 Effective versus Actual Number of Clusters for **3**, feature set/alignment options *b* and *c*. Clustering level is noted on top *x*-axis for reference. Open square – Option *b*. Open triangle – Option *c*. Solid line – Theoretical maximum ($k = k^*$)

Table 3 Clustering statistics, **2**, feature set/alignment option *c*

Clustering level	Minimum separation ratio[a]	Critical distance[b]	Actual number of clusters	Effective number of clusters[c]	%ΔEff[d]
728	1.00	32.1	1	1.00	–
727	1.04	30.9	2	1.66	33%
726	1.18	26.3	3	2.87	40%
725	1.02	25.8	4	3.88	25%
724	1.05	24.6	5	3.93	1%
723	1.01	24.4	6	4.59	11%
722	1.23	19.8	7	4.63	1%
721	1.05	18.9	8	4.67	1%
720	1.02	18.6	9	4.71	0%
719	**1.12**	**16.7**	**10**	**6.02**	**13%**
718	1.05	15.9	11	6.18	1%
717	1.09	14.5	12	6.71	4%
716	1.01	14.4	13	6.72	0%
715	1.02	14.2	14	6.87	1%
714	1.03	13.7	15	7.05	1%
713	1.15	11.9	16	7.08	0%
712	1.10	10.9	17	7.09	0%
711	1.02	10.7	18	7.42	2%
710	1.01	10.6	19	7.86	2%
709	1.01	10.5	20	7.92	0%

[a] The ratio of the distances between the most closely-related clusters for a clustering level and those of the next highest level (the level with one less cluster)

[b] The cutoff distance associated with a clustering level; corresponds to the distance between the clusters just formed

[c] An estimate of how many large clusters are in the set of clusters

[d] Percentage change in effective number of clusters

Review of the energy profiles of the clusters reveals several interesting trends. Figure 7a shows the percentages of the six major clusters and the minor clusters within each energy range for **2**. The colored patterns for the first six bars are similar, indicating that the six clusters are basically homogeneous in terms of energy profiles. In contrast, the energy profile for the minor clusters is dominated by higher relative energies. A similar pattern is seen in Fig. 7b, where the conformers of **3** are distributed evenly among the six major clusters, with many higher relative energy conformers in the minor clusters.

Representative conformers

Figure 8a and b displays the GEM conformer along with the six representative conformers for **2** and **3**, respectively. The A1 and A2 values of the representative conformers, ordered by increasing value on a scale of –180° to + 180°, are given in Table 5. The table and figure show that the A-sides of the representative conformers of the six clusters are very similar for the two analogs. This is important because A1 and A2 control the relationship between the nitrogen atom in the center ring

Table 4 Clustering statistics, **3**, feature set/alignment option *c*

Clustering level	Minimum separation ratio[a]	Critical distance[b]	Actual number of clusters	Effective number of clusters[c]	%ΔEff[d]
739	1.00	29.0	1	1.00	–
738	1.07	27.3	2	1.89	44%
737	1.07	25.6	3	2.99	37%
736	1.09	23.5	4	3.02	1%
735	1.17	20.1	5	3.88	17%
734	1.02	19.8	6	4.01	2%
733	1.11	17.8	7	4.04	1%
732	1.00	17.8	8	4.11	1%
731	1.08	16.5	9	4.34	3%
730	1.09	15.1	10	5.40	11%
729	**1.06**	**14.2**	**11**	**6.62**	**11%**
728	1.01	14.1	12	6.68	0%
727	1.00	14.1	13	6.77	1%
726	1.04	13.5	14	7.08	2%
725	1.02	13.2	15	7.40	2%
724	1.00	13.2	16	7.46	0%
723	1.07	12.4	17	7.75	2%
722	1.01	12.2	18	7.85	1%
721	1.01	12.1	19	7.92	0%
720	1.01	12.1	20	7.93	0%

[a] The ratio of the distances between the most closely-related clusters for a clustering level and those of the next highest level (the level with one less cluster)

[b] The cutoff distance associated with a clustering level; corresponds to the distance between the clusters just formed

[c] An estimate of how many large clusters are in the set of clusters

[d] Percentage change in effective number of clusters

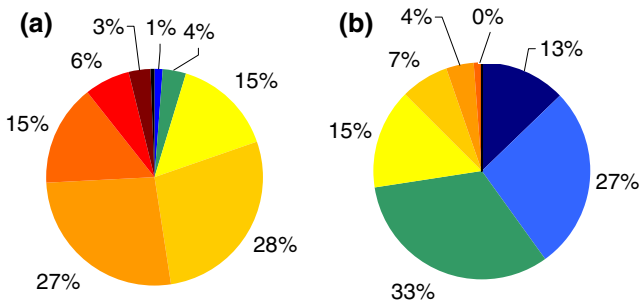


Fig. 6 Energy profiles of data sets, feature set/alignment option *c*. Percentages for each category are noted. The relative energy (in kcal/mol, relative to the energy of the global energy minimum conformer in each complete data set) is color-coded as follows: Dark blue: 0–2, Blue: 2–4, Green: 4–6, Yellow: 6–8, Gold: 8–10, Orange: 10–12, Red orange: 12–14, Red: 14–16, Dark red: 16–18, Black: 18–20. (**a**) Relative energies of the 728 conformers of **2**. (**b**) Relative energies of the 739 conformers of **3**

and the aromatic moiety (in these analogs, naphthalene) on the A-side, key pharmacophore characteristics of the majority of DA reuptake inhibitors. The representative conformers form three groups (based on their A1 values) of two conformers each (based on their A2 values). Each group has A1 values offset by approximately 120°. The two conformers in each group differ by approximately 180° in their A2 values. This can be seen more clearly in Fig. 8c and d for **2** and **3**, respectively. These figures plot the distribution of the entire set of conformers in (A1, A2) space. The representative conformer from each cluster is identified with a grey star. The figure shows that the conformers form clusters located about 120° apart on the A1 axis (at approximately –60°, +60°, and ±180°). The location of these clusters corresponds to rotational minima around the N(sp^3)–C(sp^3) bond in A1, and is typical of the rotational energy minima in aminomethane. The figure also shows that most of the conformers are found clustered along the A2 axis at A2 approximately equal to –90° and +90°, while the remaining conformers are spread along the axis at intermediate values of A2. The location of the clusters corresponds to rotational minima around the C(sp^3)–C(sp^2) bond in A2. This complex pattern of rotational minima, compared to that of methylbenzene, is due to the effect of the large substituent groups on the carbons in the C(sp^3)–C(sp^2) bond in A2. The combination of three clusters based on differences in A1, each separated into two clusters based on differences in A2, results in a total of six clusters. This shows that the hierarchical clustering approach has been able to uncover natural groups in the data.

Table 5 lists the data in terms of pairs of representative conformers for the six major clusters of **2** and **3**. The table shows that the A1 values are within 5° for each pair of representative conformers, with the A2 values showing a larger range, as is apparent in Fig. 9 which shows the representative conformer pairs aligned as in option *c*. The table also shows little difference in the range of the relative energies of the representative conformers for **2** (7.54–9.50 kcal/mol) and **3** (3.53–5.45 kcal/mol). Neither analog has its GEM as a representative conformer.

Table 6 lists the RMSD between the conformers of **2** and **3** in each of the six major clusters. The first RMSD column was calculated based on the same atoms used for feature set/alignment option *c* and shows that the five heavy atoms defining A1 and A2 have very similar locations in space for the representative conformer pairs, as seen in Fig. 9. The second RMSD column was calculated based on superposition of all heavy atoms in the molecules. The increase in the RMSD for all heavy atoms is due to differences in the B-side atoms.

Comparison of conformational space coverage

The radius of gyration of the conformational point cloud was used to measure the occupation of conformational space for each cluster. Large values indicate more coverage or "sweep" through conformational space, and

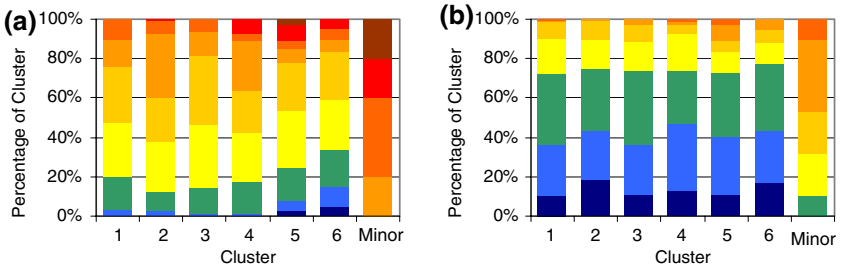


Fig. 7 Energy profile of clusters, feature set/alignment option *c*. Individual energy profiles are given for the six major clusters. All other clusters are grouped together in the "Minor" category. The percentage of each cluster with conformers within a particular range of relative energy (in kcal/mol) is color-coded as follows: Dark blue: 0–2. Blue: 2–4. Green: 4–6. Yellow: 6–8. Gold: 8–10. Orange: 10–12. Red orange: 12–14. Red: 14–16. Dark red: 16–18. Black: 18–20. (**a**) **2**. (**b**) **3**

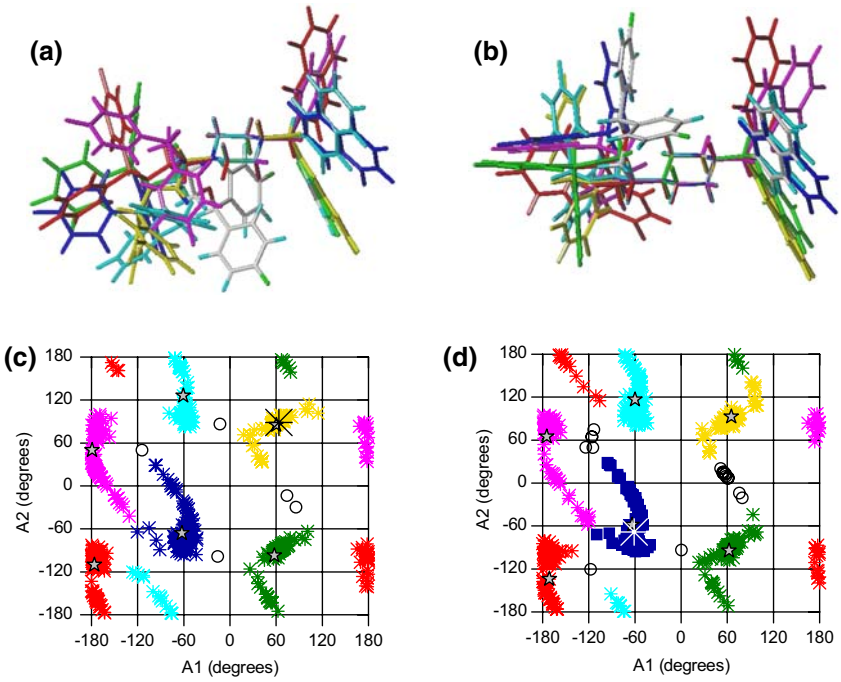


Fig. 8 Global energy minimum (GEM) conformer, representative conformers (feature set/alignment option c) and plots of conformer data sets in (A1, A2) space for **2** and **3**. GEM and representative conformers aligned by the center ring atoms are shown in (**a**) and (**b**). The complete data set of conformers is plotted in (A1, A2) space in (**c**) and (**d**). All representative conformers are color-coded by their cluster number, as noted in Table 5; the GEM conformer is grey. In (**c**) and (**d**), representative conformers are indicated by filled grey stars in the appropriate color-coded cluster. The GEM conformer for **2** is indicated by a large black asterisk overlapping the filled grey star in the yellow cluster. The GEM conformer for **3** is indicated by a large white asterisk overlapping the filled grey star in the blue cluster. Conformers not in a major cluster are indicated by open black circles. (**a**) GEM and six representative conformers for 2. (**b**) GEM and six representative conformers for **3**. (**c**) Plot of the 728 conformers of **2**. (**d**) Plot of the 739 conformers of **3**.

small values indicate less coverage. Tables 7 and 8 show the radius of gyration of the conformational point cloud for each of the six major clusters for **2** and **3**, respectively. The minor clusters (in italics) are included for comparison. The value for the radius of gyration for all conformers for **2** and the value for **3** are nearly identical (less than 3% difference) for each analog (72.1 vs. 74.0). The amount of coverage for each cluster is not directly related to the number of conformers in each cluster. For example, Cluster 3 of **2** has 82 conformers, and its radius of gyration is 33.7, whereas Cluster 4 of **2** has 151 conformers, and its radius of gyration is only 20.2. This equates to almost twice as many conformers in Cluster 4 covering less conformational space than Cluster 3.

Table 5 Representative conformers for major clusters of **2** and **3**, feature set/alignment option *c*

Cluster	Analog_conformer number[a]	A1 (degrees)	A2 (degrees)	Relative energy[b] (kcal/mol)	Cluster color in Figures[c]
1	**2**_256	–179.2	–133.0	9.50	Red
1	**3**_578	–174.6	–114.7	3.76	Red
2	**2**_453	–176.2	72.4	8.16	Magenta
2	**3**_653	–171.3	50.2	5.11	Magenta
3	**2**_593	–60.7	–55.6	8.19	Blue
3	**3**_246	–59.7	–64.8	3.53	Blue
4	**2**_546	–62.5	114.8	9.50	Cyan
4	**3**_258	–63.5	122.8	5.10	Cyan
5	**2**_101	60.1	–96.0	7.54	Green
5	**3**_489	65.2	–90.0	5.45	Green
6	**2**_573	57.9	84.6	8.34	Yellow
6	**3**_254	62.5	85.5	4.04	Yellow

[a] Conformer number in random search output

[b] Relative energy is given compared to the global energy minimum of the entire set of conformers for each analog

[c] Color of cluster in Fig. 8

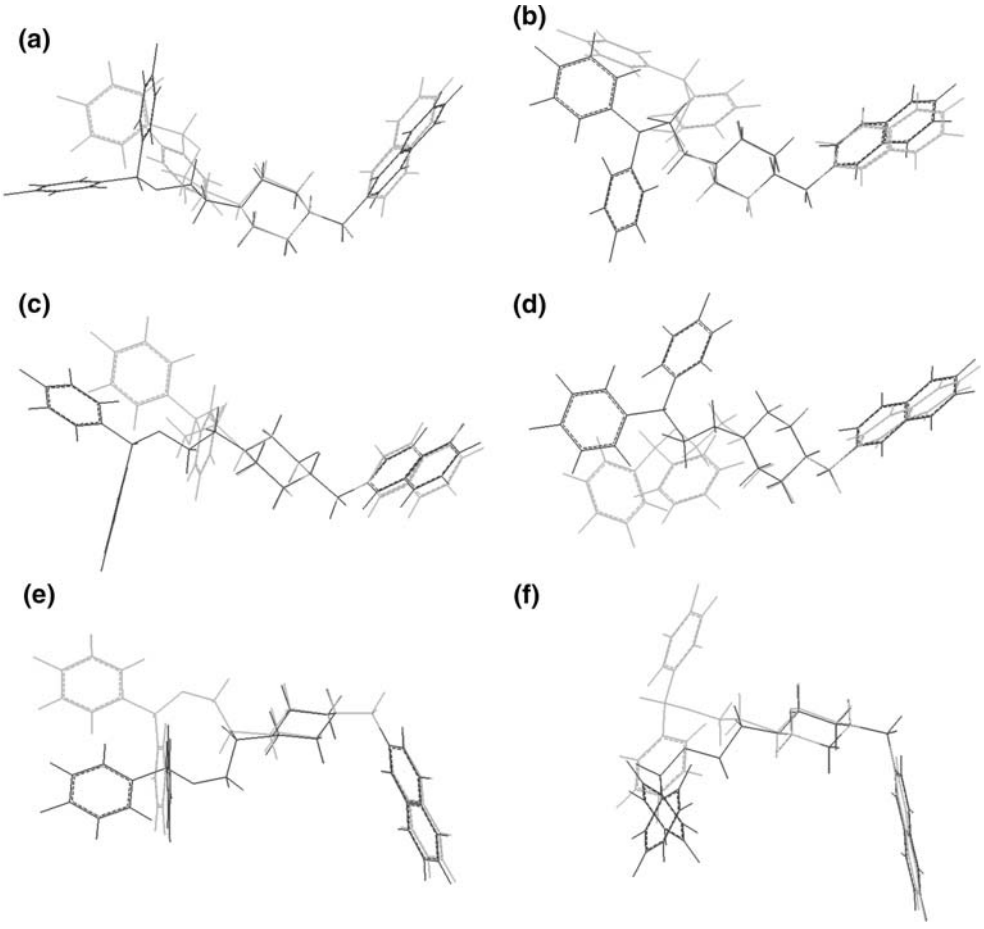


Fig. 9 Six pairs of representative conformers for Cluster 1 (**a**) through Cluster 6 (**f**), feature set/alignment option **c**. Each pair contains a conformer of **2** (shown in black) and a conformer of **3** (shown in grey), aligned as in feature set/alignment option *c* (Table 1)

Discussion

Clustering with feature sets determined by pharmacophore elements

Since the purpose of this work is to identify representative conformers to use as templates in 3D-QSAR studies of analogs of **2** and **3** with A-side substitutions, the clustering study used feature sets defined by the A-side pharmacophore elements common to most DAT inhibitors: a basic nitrogen and an aromatic ring. These were combined with different alignments to define options *a*, *b*, and *c*. Options *b* and *c* gave very similar distance maps with the same number of well-defined, large clusters along the diagonal for both **2** and **3**, even though the feature set for option *b* involves Cartesian coordinates, whereas that for *c* involves torsional angles. Shenkin and McDonald [54] noted similar behavior for their study of the roseotoxin B data set. In such cases where clustering is particularly strong and clear cut, they found that the same results appear to be obtained regardless of whether Cartesian or torsional angle space is used to define the distances for the proximity matrix. This seems to be the case for A-side clustering with options *b* and *c*.

Table 6 Comparison of representative conformers of **2** and **3**

Cluster	RMSD of atoms in feature set/ alignment option *c* (Å)	RMSD of all heavy atoms (Å)
1	0.22	2.37
2	0.31	2.24
3	0.12	1.77
4	0.10	1.95
5	0.13	2.00
6	0.08	1.79

The A1 and A2 torsional angles of the six representative conformers of **2** agree well with those of **3**. The distribution of the A1 torsional angle values for both the clusters and their representative conformers is equivalent for **2** and **3**; A2 torsional angle values agree to a lesser extent. The agreement in key A-side values for representative conformers is perhaps not surprising since **2** and **3** differ only by a substitution in the center ring on the B-side of the molecule.

Table 7 Conformational space coverage for **2**, feature set/alignment option *c*

Cluster	Number of conformers[a]	Conformation number of representative structure[b]	r_g^c
1	111	256	25.6
2	152	453	20.3
3	82	593	33.7
4	151	546	20.2
5	80	101	16.0
6	147	573	20.8
7	*1*	*375*	*0.0*
8	*1*	*527*	*0.0*
9	*1*	*687*	*0.0*
10	*2*	*94*	*7.3*
All	728	288	72.1

[a] Number of conformers in each cluster. Major clusters in regular typeface; minor clusters in italics

[b] Conformer number of representative conformer

[c] Radius of gyration (in Å) for all the conformers in the cluster; radius of gyration equals zero for single–conformer clusters

Table 8 Conformational space coverage for **3**, feature set/alignment option *c*

Cluster	Number of conformers[a]	Conformation number of representative structure[b]	r_g^c
1	98	578	31.2
2	125	653	27.6
3	143	246	22.2
4	122	258	24.4
5	111	489	14.7
6	121	254	21.7
7	*10*	*557*	*3.7*
8	*2*	*268*	*2.5*
9	*1*	*589*	*0.0*
10	*5*	*427*	*7.3*
11	*1*	*241*	*0.0*
All	739	522	74.0

[a] Number of conformers in each cluster. Major clusters in regular typeface; minor clusters in italics

[b] Conformer number of representative conformer

[c] Radius of gyration (in Å) for all the conformers in the cluster; radius of gyration equals zero for single-conformer clusters

Alternative feature set/alignment options

Hierarchical clustering with feature sets defined by the Cartesian coordinates of all heavy atoms (for full molecule clustering) or the Cartesian coordinates or torsional angles of the bisphenyl group (for B-side clustering) were investigated for the sake of completeness. These feature sets were combined with various alignments to give options *d* and *e* for full molecule clustering and options *f*–*k* for B-side clustering. The results are summarized in Table 9. The distance maps are not shown. Options *d* and *e* resulted in distance maps that are essentially featureless for both **2** and **3**. All four distance maps have numerous small squares on the diagonal, indicating that major clusters were not found for full molecule clustering. The large number of small squares is indicative of the diversity of the conformations when the analogs are viewed as a whole. Therefore, clustering with feature sets defined in terms of the A-side pharmacophore elements, as carried out in the present work, is useful because it allows one to classify the conformers of analogs of **1** in terms of a substructure region that is known to be common to other dopamine reuptake inhibitors. Since these other DAT inhibitors are much smaller than **1**, and no others except the benztropines contain the bisphenyl moiety, the basic nitrogen and proximal aromatic ring are of major importance to the binding of such compounds to the DAT.

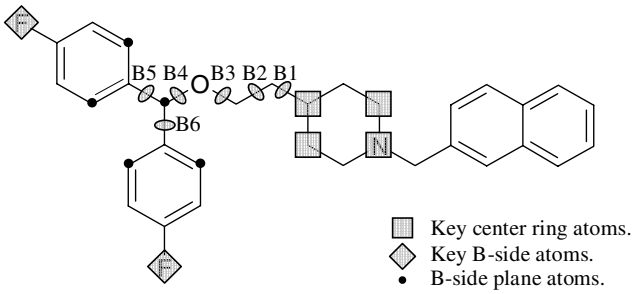
B-side clustering for both **2** and **3** with options *f*, *g*, and *j* resulted in featureless distance maps with many small clusters along the diagonal. Distance maps for options *h*, *i*, and *k* show from 3 to 15 major clusters of varying sizes on the diagonal. In contrast to the A-side distance maps (options *a*, *b*, and *c*) that are comparable for analogs **2** and **3**, the B-side distance maps for several feature set/alignment options do not match well for **2** and **3**. The distance maps for option *k* for both analogs show nine clearly-defined clusters, but the sizes and patterns of the clusters vary between the analogs. Of particular note is the comparison for option *i*, where a distinctly different pattern is seen in the presence of about 15 small clusters for **2** in contrast to a few large clusters for **3**. This may mean that analog **2** achieves more types of B-side conformations than analog **3**, due the presence of a nitrogen and lone pair in **2** versus a methyne group in **3** on the B-side of the central ring.

Principal component analysis of the eight torsional angles of **2** and **3** using singular value decomposition [61] also uncovered differences in the two data sets that were traced to the fact that the nitrogen with its lone pair attempts to invert, so that the nitrogen in some conformers becomes almost planar with respect to its surrounding carbons. In contrast, the carbon of the methyne group in **3** remains tetrahedral; therefore, the carbon is nonplanar with respect to its neighboring hydrogen and carbons. This could account for the differences noted in the B-sides of **2** and **3**. It is also possible that some of the perceived differences in the data sets of the two analogs are due to incomplete conformational searching. This issue will be addressed in a future publication.

Comparison to fuzzy relational clustering results

The results of this study are supported by a fuzzy relational clustering (FRC) study performed by

Table 9 Alternate feature set/alignment options



Option	Feature set[a]	Alignment[b]	Distance map appearance	
			2	**3**
d	All heavy atoms	□ Center ring[c]	Many small squares	Many small squares
e	All heavy atoms	All heavy atoms	Many small squares	Many small squares
f	□ Center ring and ◇ B-side key atoms	□ Center ring[c]	Many small squares	Many small squares
g	□ Center ring and ◇ B-side key atoms	□ Center ring and ◇ B-side key atoms	Many small squares	Many small squares
h	◇ B-side key atoms	Oxygen and neighboring carbons[c]	Unevenly sized squares	Unevenly sized squares
i	◇ B-side key atoms and ● B-side plane atoms	Oxygen and neighboring carbons[c]	About 15 squares	About 3 large squares
j	B1 through B6	Nine heavy atoms defining B1 through B6[d]	Many small squares	Many small squares
k	B3 and B4	Five heavy atoms defining B3 and B4[d]	Nine squares	Nine squares

[a] Feature set: The atoms or torsional angles used to calculate the intermolecular distances. Key atoms (identified by symbols noted) and torsional angles are shown above. RMSD calculations for the distance matrix were carried out on the feature set atoms in the corresponding alignment. Options *a* and *b* are atom-based clusterings; option *c* is a torsional angle-based clustering

[b] Alignment performed by XCluster unless otherwise noted

[c] Alignment performed in SYBYL

[d] Alignment of conformers for the torsional angle studies is not necessary to calculate RMSD values, but alignment was carried out to properly visualize clusters

Venanzi and coworkers [59]. The FRC study was carried out only on the data set of conformers of **2.** In contrast to the hierarchical approach described here, the FRC study used different feature sets, each constructed from a combination of Cartesian coordinates of atoms and angles between molecular planes, as well as different superpositions. The FRC A-side results gave six representative conformers that fall into three sets of two conformers each. The conformers of each of the three sets differ from those of the other two by approximately 120° in A1, with each subset of two conformers differing from each other by about 180° in A2. This is similar to the hierarchical clustering results obtained with option *c*, supporting the identification of six clusters for A-side clustering. The FRC B-side clustering identified nine clusters (similar to the option *k* results in Table 9), while clustering with full molecule feature sets was unable to locate clusters (similar to the options *d* and *e* results in Table 9). The FRC and XCluster studies, therefore, produced comparable results, especially for A-side clustering, although different protocols and feature sets were used in each approach.

Alternative clustering statistics

Although a minimum separation ratio with a value greater than 2 is an excellent discriminator of the optimal clustering level, Shenkin and McDonald noted that this statistic is not definitive for all data sets [54]. A high value of minimum separation ratio indicates good clustering, but a low value does not rule out good clustering. In the present work, distance maps for option *c* showed well-defined, large clusters for **2** and **3**, yet neither case had a minimum separation ratio significantly above one. Therefore an alternative statistic, percent change in the effective number of clusters (%ΔEff), was defined. This novel,

alternative statistic was shown to be useful in determining the optimal clustering level.

Conformational flexibility

The flexibility of a molecule can be estimated by its number of rotatable bonds (RTB), calculated according to Oprea [62] by:

$$\mathrm{RTB} = N_{\mathrm{nt}} + \sum_i (n_i - 4 - \mathrm{RGB}_i - \mathrm{ShB}_i) \quad (4)$$

where RTB = number of rotatable bonds; N_{nt} = number of non-terminal single bonds (excluding single bonds in certain groups); n_i = number of single bonds in a non-aromatic ring i with six or more bonds; RGB_i = number of rigid bonds in ring i; ShB_i = number of bonds shared by ring i with any other ring.

Thus, for the three compounds of interest, ignoring aromatic rings:

1: RTB = 10 + (6 – 4 – 0 – 0) = 12
2: RTB = 8 + (6 – 4 – 0 – 0) = 10
3: RTB = 8 + (6 – 4 – 0 – 0) = 10

According to Oprea's review of drug-related chemical databases [62], over 70% of the drug-like molecules found in six well-known compound databases have a value of RTB between two and eight. The review also noted that the distribution of RTB among the drug-like databases shows that the Physician's Desk Reference compound database, containing over-the-counter and prescription drug information, has the middle 50% of its compounds with RTB from two to seven. Thus, **1**, **2**, and **3** appear to be more flexible than the average drug-like molecule.

Conclusion

The hierarchical clustering results show that when independent clustering studies are performed on these piperazine and piperidine analogs of **1**, the results are strikingly similar for the A-side analyses. The present conformational analysis forms the basis for 3D-QSAR studies of analogs of **2** and **3** with A-side substitutions. The 3D-QSAR work based on this study will help elucidate the effects of steric and electrostatic forces on the activity of the piperazine and piperidine analogs of **1**.

Acknowledgements This work was supported in part by NIH grant DA018153 (C.A.V.). K.M.G. acknowledges the support of Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship DA015555.

References

1. Kuhar MJ, Ritz MC, Boja JW (1991) TINS 14:299
2. Singh S (2000) Chem Rev 100:925
3. Ravna AW, Sylte I, Kristiansen K, Dahl SG (2006) Bioorg Med Chem 14:666
4. Sen N, Shi L, Beuming T, Weinstein H, Javitch JA (2005) Neuropharmacology 49:780
5. Glowa JR, Wojnicki FHE, Matecka D, Bacher J, Mansbach RS, Balster RL, Rice KC (1995) Exp Clin Psychopharmacol 3:219
6. Elmer GI, Brockington A, Gorelick DA, Carroll FI, Rice KC, Matecka D, Goldberg SR, Rothman RB (1996) Pharmacol Biochem Behav 53:911
7. Rothman RB, Mele A, Reid AA, Akunne HC, Greig N, Thurkauf A, de Costa BR, Rice KC, Pert A (1991) Pharmacol Biochem Behav 40:387
8. Mojsiak J (2003) Unpublished results. National Institutes of Health
9. Prisinzano T, Rice KC, Baumann MH, Rothman RB (2004) Curr Med Chem – Central Nerv Syst Agents 4:47
10. Kolhatkar R, Cook CD, Ghorai SK, Deschamps J, Beardsley PM, Reith MEA, Dutta AK (2004) J Med Chem 47:5101
11. Kolhatkar RB, Ghorai SK, George C, Reith MEA, Dutta AK (2003) J Med Chem 46:2205
12. Matecka D, Lewis D, Rothman RB, Dersch CM, Wojnicki FHE, Glowa JR, De Vries AC, Pert A, Rice KC (1997) J Med Chem 40:705
13. Dutta AK, Xu C, Reith MEA (1996) J Med Chem 39:749
14. Dutta AK, Reith MEA, Madras BK (2001) Synapse 39:175
15. Prisinzano T, Greiner E, Johnson EM II, Dersch CM, Marcus J, Partilla JS, Rothman RB, Jacobson AE, Rice KC (2002) J Med Chem 45:4371
16. Cramer RD III, Patterson DE, Bunce JD (1988) J Am Chem Soc 110:5959
17. Carroll FI, Gao Y, Rahman MA, Abrams P, Parham K, Lewin AH, Boja JW, Kuhar MJ (1991) J Med Chem 34:2719
18. Froimowitz M (1993) J Comput Chem 14:934
19. Carroll FI, Mascarella SW, Kuzemko MA, Gao Y, Abraham P, Lewin AH, Boja JW, Kuhar MJ (1994) J Med Chem 37:2865
20. Yang B, Wright J, Eldefrawi ME, Pou S, MacKerell AD Jr (1994) J Am Chem Soc 116:8722
21. Lieske SF, Yang B, Eldefrawi ME, MacKerell AD Jr, Wright J (1998) J Med Chem 41:864
22. Zhu N, Harrison A, Trudell ML, Klein-Stevens CL (1999) Struct Chem 10:91
23. Muszynski IC, Scapozza L, Kovar K-A, Folkers G (1999) Quant Struct Act Relationsh 18:342
24. Hoffman BT, Kopajtic T, Katz JL, Newman AH (2000) J Med Chem 43:4151
25. Davies HML, Gilliatt V, Kuhn LA, Saikali E, Ren P, Hammond PS, Sexton GJ, Childers SR (2001) J Med Chem 44:1509
26. Zhan CG, Zheng F, Landry DW (2003) J Am Chem Soc 125:2462
27. Paula S, Tabet MR, Keenan SM, Welsh WJ, Ball WJ Jr (2003) J Mol Biol 325:515
28. Paula S, Tabet MR, Farr CD, Norman AD, Ball WJ Jr (2004) J Med Chem 47:133
29. Yuan H, Kozikowski AP, Petukhov PA (2004) J Med Chem 47:6137
30. Kulkarni SS, Grundt P, Kopajtic T, Katz JL, Newman AH (2004) J Med Chem 47:3388

31. Robarge MJ, Agoston GE, Izenwasser S, Kopajtic T, George C, Katz JL, Newman AH (2000) J Med Chem 43:1085
32. Newman AH, Izenwasser S, Robarge MJ, Kline RH (1999) J Med Chem 42:3502
33. Fengyi L, Boli L, Zhaoxing M, Shijun Z (2004) J Mol Struct (Theochem) 712:207
34. Froimowitz M, Wu K-M, Rodrigo J, George C (2000) J Comput Aided Mol Des 14:135
35. Froimowitz M, George C (1998) J Chem Inf Comp Sci 38:506
36. Kulkarni SS, Newman AH, Houlihan WJ (2002) J Med Chem 45:4119
37. Froimowitz M, Patrick KS, Cody V (1995) Pharm Res 12:1430
38. Venanzi CA, Misra M, Gilbert KM, Buono RA, Schweri MM, Shi Q, Deutsch HM (2006) To be submitted for publication. New Jersey Institute of Technology
39. Gilbert KM, Skawinski WJ, Misra M, Paris KA, Naik NH, Deutsch HM, Venanzi CA (2004) J Comput Aided Mol Des 18:719
40. Wang S, Sakamuri S, Enyedy IJ, Kozikowski AP, Zaman WA, Johnson KM (2001) Bioorg Med Chem Lett 9:1753
41. Nicklaus MC, Wang S, Driscoll J, Milne GWA (1995) Bioorg Med Chem 3:411
42. Veith M, Hirst JD, Brooks CL III (1998) J Comput Aided Mol Des 12:563
43. Boström J, Norrby P-O, Liljefors T (1998) J Comput Aided Mol Des 12:383
44. Debnath AK (1999) J Med Chem 42:249
45. Perola E, Charifson PS (2004) J Med Chem 47:2499
46. Benedetti P, Mannhold R, Cruciani G, Pastor M (2002) J Med Chem 45:1577
47. Guarnieri F, Weinstein H (1996) J Am Chem Soc 118:5580
48. Hopfinger AJ, Tokarski JS (1997) In: Charifson PS (eds), Practical application of computer-aided drug design. Marcel Dekker, New York, 1997, pp 105–164
49. Barnett-Norris J, Guarnieri F, Hurst DP, Reggio PH (1998) J Med Chem 41:4861
50. Barnett-Norris J, Hurst DP, Lynch DL, Guarnieri F, Makriyannis A, Reggio PH (2002) J Med Chem 45:3649
51. Greenidge PA, Merette SAM, Beck R, Dodson G, Goodwin CA, Scully MF, Spencer J, Weiser J, Deadman JJ (2003) J Med Chem 46:1293
52. Bernard D, Coop A, MacKerell AD Jr (2003) J Am Chem Soc 125:3101
53. Bernard D, Coop A, MacKerell AD Jr (2005) J Med Chem 48:7773
54. Shenkin PS, McDonald DQ, (1994) J Comput Chem 15:899
55. Clark M, Cramer RD III, Van Opdenbosch N (1989) J Comput Chem 10:982
56. Berfield JL, Wang LC, Reith MEA (1999) J Biol Chem 274:4876
57. Skawinski WJ (2004) Unpublished results. New Jersey Institute of Technology
58. Dutta AK, Meltzer PC, Madras BK, (1993) Med Chem Res 3:209
59. Misra M, Banerjee A, Davé RN, Venanzi CA (2005) J Chem Inf Model 45:610
60. Shenkin PS, Erman B, Mastrandrea LD (1991) Struct Funct Gen 11:297
61. Fiorentino A, Pandit D, Gilbert KM, Misra M, Dios R, Venanzi CA (2006) J Comput Chem 27:609
62. Oprea TI (2000) J Comput Aided Mol Des 14:251