J Mol Model (2011) 17:1343–1351
DOI 10.1007/s00894-010-0826-1

ORIGINAL PAPER

Singular value decomposition analysis of the torsional angles of dopamine reuptake inhibitor GBR 12909 analogs: effect of force field and charges

Deepangi Pandit · Anna Fiorentino · Supreet Bindra · Carol A. Venanzi

Received: 15 May 2010 / Accepted: 17 August 2010 / Published online: 14 September 2010

Abstract Three-dimensional quantitative structure-activity relationship (3D-QSAR) analysis of large, flexible molecules, such as the dopamine reuptake inhibitor GBR 12909 (**1**), is complicated by the fact that they can take on a wide range of closely-related conformations. The first step in the analysis is to classify the conformers into groups. Over 600 conformers each of a piperazine (**2**) and piperidine (**3**) analog of **1** were generated by random search conformational analysis using the Merck Molecular Force Field (MMFF94). Singular value decomposition (SVD) was used to group the conformers of **2** and **3** by the similarity of their non-ring torsional angles. SVD uncovered subtle differences in their conformer populations due to that fact that the conformers separate along different principal components, and ultimately to the fact that different torsional angles are the chief contributors to these components. The results were compared to our previous SVD analysis (Fiorentino, et al., Journal of Computational Chemistry, 2006, 27, 609-620) of conformer populations of **2** and **3** generated by the Tripos force field and Gasteiger-Hückel charges. Except for the dominant contribution of angle B3 to principal component 8 seen with both force fields, the angles which are chiefly responsible for the grouping of the conformers of **2** and **3** are different with both force fields. This illustrates that SVD is useful in identifying unique groupings of conformers in large data sets of flexible molecules—a first step in selecting representative conformers for 3D-QSAR modeling studies.

Keywords Dopamine reuptake inhibitors · GBR 12909 · Principal component analysis · Singular value decomposition · 3D-QSAR

D. Pandit · A. Fiorentino · S. Bindra · C. A. Venanzi (✉)
Department of Chemistry and Environmental Science,
New Jersey Institute of Technology,
Newark, NJ 07102, USA
e-mail: venanzi@adm.njit.edu

Present Address:
D. Pandit
BioMaPS Institute for Quantitative Biology,
Rutgers, The State University of New Jersey,
610 Taylor Road,
Piscataway, NJ 08854-8087, USA

Introduction

The "dopamine hypothesis" [1] implicates the dopamine transporter (DAT) in cocaine abuse and addiction. Structure-activity relationship (SAR) studies for several classes of dopamine reuptake inhibitor [2–14] provide a wealth of data for three-dimensional quantitative structure-activity relationship (3D-QSAR) modeling [15–26]. The GBR 12909 class of DAT ligands has been extensively studied as a potential treatment for cocaine dependence. To date, the DAT binding affinity of hundreds of analogs of GBR 12909, 1-{2-[bis(4-fluorophenyl)methoxy]ethyl}-4-(3-phenylpropyl)piperazine, (**1**, Fig. 1), has been determined [3, 6, 10, 27–40].

In a typical 3D-QSAR modeling protocol used by a technique such as Comparative Molecular Field Analysis (CoMFA), [41] one would select a rigid analog with good binding affinity to act as the molecular template for superposition of the conformers of other analogs. However, CoMFA studies on GBR 12909 analogs are complicated by the fact that these molecules have several rotatable bonds and, as a result, take on a wide range of closely-related conformations. Since CoMFA results are sensitive to the conformation used as the template structure, [41, 42] one approach, used by us in a study of GBR 12909 analogs, [43] is to carry out a series of CoMFA calculations, each using a different representative structure as the template.

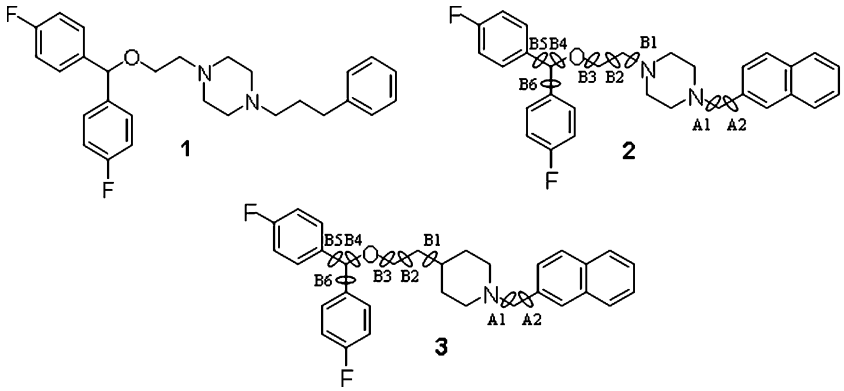

Fig. 1 Structure of GBR 12909, **1**, and analogs **2** and **3**. Torsional angles used in the study are labeled A1, A2, B1...B6

For flexible molecules, however, the question arises as to how to group the many conformations in order to select a representative conformer from each group. Hierarchical clustering (HC), [44] fuzzy clustering, [45, 46] or fuzzy relational clustering (FRC) [47] and singular value decomposition (SVD) [48] are possible approaches. We applied HC, [49] FRC, [50, 51] and SVD [52] to the classification of conformers of two GBR 12909 analogs (**2** and **3**, Fig. 1). They each have eight rotatable bonds (A1, A2, B1...B6) and are consequently very flexible. Both analogs were protonated on their common nitrogen [28]. The A-side (containing torsional angles A1 and A2 and the naphthyl group) is slightly less flexible than **1**, while the B-side (containing torsional angles B1...B6 and the bisphenyl group) is exactly the same as the B-side of **1**. The analogs differ only by a small change in the central heterocyclic ring, yet have different binding and selectivity characteristics at the DAT (Table 1). We showed that all three techniques were able to uncover natural groups in the data resulting from rotational minima around specific torsional angles. The SVD technique grouped conformations based on the similarity of their non-ring torsional angles [52]. The advantage of using torsional angles, which determine the overall shape of the molecule, over using features such as atomic coordinates is that the analysis is independent of molecular alignment. SVD was able to uncover subtle differences in the conformer populations of the two analogs due to nitrogen inversion at the unprotonated piperazinyl nitrogen in **2**.

The HC, FRC, SVD and CoMFA analyses described above were applied to a data set of conformers of **2** and **3** generated by conformational searching (random search) [53] using the Tripos force field [54] and Gasteiger-Hückel atomic charges [55]. In order to investigate the sensitivity of the conformer populations to the force field used to generate them, [55] the random search procedure was also carried out using the Merck Molecular Force Field (MMFF94) and associated charges [56–59]. Plots of the conformational ensembles of **2** and **3** in torsional angle space showed that subtle differences in the treatment of the lone pairs on N and O by the two force fields led to different physical characteristics of the conformer populations [55]. In particular, the MMFF94 force field did not allow the nitrogen inversion seen with the Tripos force field. The two force fields also predicted slightly different minima for rotation around the $C(sp^3)$ – $O(sp^3)$ bond involved in angles B3 and B4 (see Fig. 1). Because the two force fields had different effects on the planarity of the unprotonated nitrogen in **2** (involving B1) and the rotational minima around B3 and B4 in 2 and 3, this resulted in different orientations of the B-side side chain. The two different force fields were shown to generate somewhat different conformer populations for **2** and **3**. [55]

In the present work, the two sets of conformer populations of **2** and **3** generated by the two different force fields were used to investigate the sensitivity of SVD to variations in the data set of molecular conformations. As will be seen below, plots of the conformational ensembles of **2** and **3** in principal component space showed that SVD could identify differences in the conformer populations of **2** and **3** generated by the two force fields.

In this study, we use a novel scaling technique for treating circular (torsional angle) data first described in our previous work [52]. To the best of our knowledge, this and our previous study [52] are the only applications of SVD to the analysis of conformations of very flexible molecules. SVD has been applied to phosphate backbone angles in DNA, but that is a system in which the differences in molecular shapes (A, B_I, B_{II}, and crankshaft) are well-defined [60]. The GBR 12909 analogs, however, are far more flexible than DNA, having a continuum of conformations rather than a discrete set of distinct forms.

Table 1 Binding affinity[a] of selected GBR 12909 analogs at the DAT and SERT labeled with [^{125}I] RTI-55

Analog	DAT	SERT	SERT/DAT
GBR 12909 (**1**) [31]	3.7 (±0.4)	126 (±27)	34
2 [29]	8.0 (±0.3)	312 (±15)	39
3 [31]	0.71 (±0.06)	229 (±21)	323

[a] Ki given in nanomolar(nM), standard deviation in parentheses

Table 2 Torsional angles[a] of global energy minimum conformers

Analog, force field[b]	A1	A2	B1	B2	B3	B4	B5	B6
2, MMFF94	202.2	80.5	54.3	62.4	197.3	339.7	356.7	252.0
3, MMFF94	202.2	80.2	50.2	59.5	192.6	346.3	177.8	252.4
2, Tripos	63.8	88.6	262.7	310.1	183.4	64.4	115.1	339.8
3, Tripos	298.7	291.8	51.5	58.5	189.6	306.7	67.9	6.5

[a] Angles in degrees

[b] MMFF94 force field and charges; Tripos force field and Gasteiger-Hückel charges

Methods

The SVD methodology is the same as in our previous work [52]. It is summarized below for the convenience of the reader.

Random search conformational analysis

The random search procedure locates minima on the conformational potential energy surface of the molecule. The random search procedure was carried out for the eight torsional angles (A1, A2, B1, . . . B6) of the protonated forms of **2** and **3** using the MMFF94 force field and charges with SYBYL Versions 6.9-7.1(available from Tripos, Inc., St. Louis, MO). Details of the procedure have been published elsewhere [55]. All calculations were performed on a SGI Origin 2000 with twenty 300 MHz processors, 20 GB of memory, using the IRIX "6.5.18 m UNIX"-based operating system.

Data pretreatment

The input data consisted of the eight torsional angles (A1, A2, B1 . . . B6) for each of the conformations of **2** and **3**. The conformer population of each of the analogs was analyzed separately. The torsional angles of the conformers in the original (raw) data ranged in value from -180° to 180° and were converted to values ranging from 0° to 360° following Reijmers et al. [61]. For each analog, the torsional angles of each conformer were defined relative to those of the global energy minimum (GEM) conformer identified by the random search. This GEM-scaling procedure [52] ensures that the difference between each of the (A1, A2, B1 . . . B6) torsional angles in a particular conformer and its counterpart in the GEM conformer falls between -180° and 180°.

Principal component analysis and SVD

Principal component analysis (PCA) is based on finding the directions in the data with the most variation, i.e., the eigenvectors corresponding to the largest eigenvalues of the covariance matrix, and projects the data onto these directions. SVD is a general form of PCA in which the data matrix is decomposed into a score (left singular) matrix and a loading (right singular) matrix. The SVD analysis was performed using MATLAB versions 7.0-7.1 for Windows (available from the Mathworks Inc., Natick, MA). The "svd" command was used to decompose the original data matrix **X** of dimensions $r \times c$ into three matrices, **U**, **S**, and **V**, where $\mathbf{X}=\mathbf{USV}^{\mathbf{T}}$ [62]. Each row (r) of the matrix represents a separate conformer of an analog and each of the eight columns (c) contains the GEM-scaled torsional angle (A1 . . . B6). The columns of **U** are left singular vectors of $\mathbf{XX}^{\mathbf{T}}$, whereas the rows of **V** are right singular vectors of $\mathbf{X}^{\mathbf{T}}\mathbf{X}$. **S** contains the ordered square roots of the eigenvalues from highest to lowest. The variance explained by each principal component is the sum of all the eigenvalues divided by the eigenvalue of the corresponding principal component. Consequently, the first principal component has

Table 3 Variance and angle/principal component correlation coefficients, analog **2**

Principal component	1	2	3	4	5	6	7	8
Variance	20.21	16.08	15.38	14.44	11.30	9.65	7.78	5.17
Cumulative variance	20	36	52	66	77	87	95	100
	PC1[a]	PC2	PC3	PC4	PC5	PC6	PC7	PC8
A1	−0.19	−0.20	0.09	0.32	0.15	**0.77**	**0.46**	0.12
A2	−0.01	**−0.48**	**−0.89**	0.05	−0.01	−0.03	−0.01	0.02
B1	**−0.36**	**0.80**	**−0.42**	−0.06	0.23	0.19	−0.10	−0.06
B2	0.04	0.12	−0.05	**0.90**	**−0.37**	−0.05	−0.11	−0.05
B3	0.06	−0.16	0.11	0.06	−0.03	**0.35**	−0.25	**−0.85**
B4	**0.53**	0.09	0.02	−0.04	−0.13	**0.45**	**−0.67**	0.30
B5	**0.68**	−0.11	0.01	0.25	**0.67**	−0.05	−0.05	−0.05
B6	**0.76**	0.16	−0.18	−0.33	**−0.38**	0.18	0.22	−0.10

[a] PC1...PC8: Principal components 1 through 8. Largest correlation coefficients are highlighted in boldface type

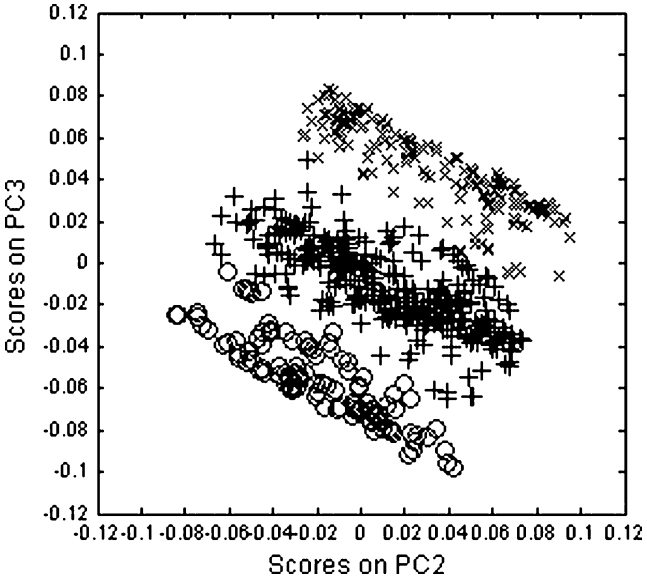

Fig. 2 Scores on PC2 vs. Scores on PC3 for **2**. The 'o', '+', and 'x' symbols were assigned by inspection to assist in visualization of the data separation

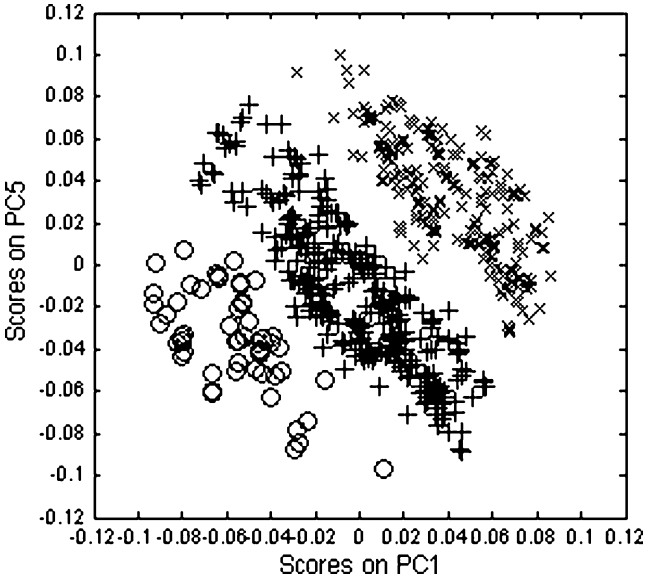

Fig. 4 Scores on PC1 vs. Scores on PC5 for **2**. The 'o', '+', and 'x' symbols were assigned by inspection to assist in visualization of the data separation

the highest eigenvalue, and the largest associated variance. These eigenvalues are the singular values.

The function corrcoef(x,y) was used to calculate correlation coefficients between the variables and the principal components where x is the column of the corresponding **U** matrix and y is the column of the variable from the data matrix **X**. The contribution of each variable to each principal component is given by correlation coefficients between the variables and the principal components. Large values of the correlation coefficient, with either negative or positive signs, indicate major contributors. The correlation coefficient matrix of the entire data set is used to obtain the correlation coefficients between the variables themselves.

Score plots are used to visualize data separation and loading plots to visualize relations between variables. Score plots were constructed for every possible combination of two principal components. Only those score plots which demonstrate data separation are discussed below. Since loading plots do not provide any additional information beyond that given in the tables of torsional angle/principal component correlation coefficients, no loading plots are given here.

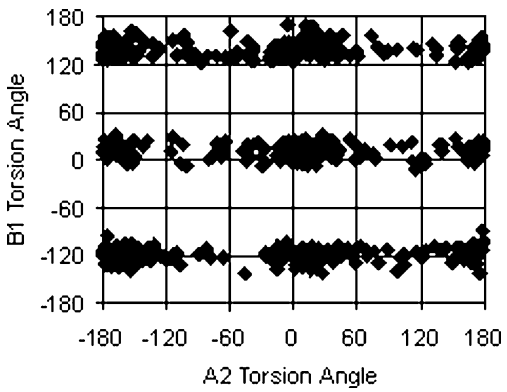

Fig. 3 Local minima of **2** in (A2, B1) space. Torsional angles are GEM-scaled and are given in degrees

Results

Analog 2

Conformational analysis

The conformer populations resulting from the random search conformational analysis of **2** and **3** using the MMFF94 force field are discussed in detail and compared extensively to the Tripos force field results in a separate publication [55]. Relevant points are summarized here. The random search produced 643 conformers for **2** with relative energy ranging from 0 to 19 kcal mol^{-1}. The torsional angles A1, A2, B1, . . . B6 of the GEM conformer of **2** are given in Table 2. These values were used to carry out the GEM scaling described in the Methods section.

SVD

Table 3 shows the percentage of the variance explained by each PC, along with the correlation coefficients between the PCs and each variable for **2**. As can be seen from the table, the first three PCs explain only 52% of

Table 4 Angle/angle correlation coefficients, analog **2**[a]

	A1	A2	B1	B2	B3	B4	B5	B6
A1	1	0.01	−0.02	0.09	**0.12**	−0.04	−0.01	**−0.14**
A2	0.01	1	0.01	0.02	−0.03	−0.05	0.02	0.04
B1	−0.02	0.01	1	−0.01	−0.09	−0.04	**−0.14**	−0.08
B2	0.09	0.02	−0.01	1	0.04	0.05	0.03	−0.08
B3	**0.12**	−0.03	−0.09	0.04	1	**0.11**	0.04	0.02
B4	−0.04	−0.05	−0.04	0.05	**0.11**	1	**0.20**	**0.33**
B5	−0.01	0.02	**−0.14**	0.03	0.04	**0.20**	1	**0.16**
B6	**−0.14**	0.04	−0.08	−0.08	0.02	**0.33**	**0.16**	1

[a] Largest correlation coefficients are highlighted in boldface type

variance and no single principal component is able to explain a large percentage of variance. The first principal component, PC1, explains 20% of variance and is correlated with the B1 (−0.36), B4 (0.53), B5 (0.68) and B6 (0.76) torsion angles. Principal components PC2 and PC3 are highly correlated to B1 (0.80) and A2 (−0.89), respectively, with somewhat lower correlation between A2 and PC2 (−0.48) and B1 and PC3 (−0.42). Table 3 shows that the chief contributor to PC4 is B2 (0.90), whereas B5 (0.67) and to a lesser extent B2 (−0.37) and B6 (−0.38) are the chief contributors to PC5. For PC6, A1 (0.77) is the chief contributor, followed by B4 (0.45) and B3 (0.35). For PC7, B4 (−0.67) and A1 (0.46) are the chief contributors, whereas B3 (−0.85) is the chief contributor to PC8.

Of all the score plots, only two show separation of the conformers into groups: PC2 vs. PC3 and PC1 vs. PC5. As can be seen from Fig. 2, the PC2 vs. PC3 score plot divides conformers roughly into three groups along the diagonal. This type of separation suggests that both PCs contribute to the separation. Since B1 is the chief contributor to PC2 and A2 to PC3, these two angles are responsible for the data separation. Also note that, while B1 is highly correlated to PC2 (0.80), A2 also has a large correlation coefficient (0.48), whereas for PC3 the trend is reversed: B1 (−0.42) and A2 (−0.89). Therefore, both angles contribute significantly to the data separation along both PCs. Since B1 influences the position of the B-side side chain and A2 influences that of the A-side side chain, it is reasonable that these two angles would contribute to separation of the conformers into groups.

Figure 3 plots the conformers in the space of the angles which are the chief contributors to PC2 (B1) and PC3 (A2). The data show very distinct separation into three groups along B1 due to rotational minima around the $C(sp^3) – N(sp^3)$ bond and little separation along A2 due to rotational minima around the $C(sp^3) – C (sp^2)$ bond. On the basis of viewing this type of conformer plot in torsional angle space, one might be tempted to group the conformers based on B1 alone. Yet principal component analysis uncovers more subtle trends. This illustrates the utility of SVD in uncovering subtle influences on conformer groups not seen by direct analysis of conformer plots in two-dimensional torsional angle space as in Fig. 3.

Figure 4 gives the PC5 vs. PC1 score plot and shows that the conformers divide roughly into three groups along the diagonal. As with the diagonal separation seen in Fig. 2, the same two angles (here, B5 and B6) are major contributors to both PCs: B5 (0.68) and B6 (0.76) for PC1, B5 (0.67) and B6 (−0.38) for PC5. B1 (−0.36) and B4 (0.53) are also contributors to PC1, and B2 (−0.37) to PC5. The data separation noted in the figure is due to these angles, with the largest influence stemming from those angles with the largest correlation coefficients to the PCs.

Table 4 gives the correlation coefficients between all eight angles for the analog **2** data. These coefficients represent full,

Table 5 Variance and angle/principal component correlation coefficients, analog **3**

Principal component	1	2	3	4	5	6	7	8
Variance	21.80	16.97	16.47	14.00	11.03	9.70	6.34	2.99
Cumulative variance	22	39	55	70	81	91	97	100
	PC1[a]	PC2	PC3	PC4	PC5	PC6	PC7	PC8
A1	0	−0.34	−0.09	−0.28	**0.79**	**−0.39**	−0.21	0.02
A2	0.28	0.09	**−0.95**	−0.02	−0.05	0.01	−0.04	0.01
B1	**−0.41**	**−0.81**	−0.18	0.31	0.01	0.29	0.04	−0.02
B2	**−0.55**	0.19	−0.05	**−0.69**	0.13	**0.50**	−0.02	0.02
B3	−0.16	−0.02	0.02	−0.12	0.10	−0.06	0.11	**−0.97**
B4	−0.23	0.06	−0.11	−0.02	0.33	−0.15	**0.91**	0.08
B5	**−0.76**	0.01	−0.22	−0.25	−0.33	**−0.42**	0.03	−0.06
B6	**−0.36**	**0.57**	−0.13	**0.57**	0.35	0.18	−0.04	−0.07

[a] PC1...PC8: Principal components 1 through 8. Largest correlation coefficients are highlighted in boldface type

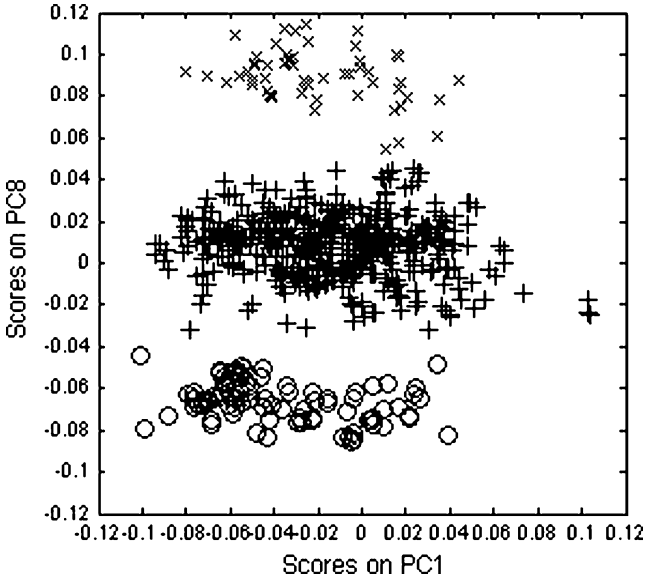

Fig. 5 Scores on PC1 vs. Scores on PC8 for **3**. The 'o', '+', and 'x' symbols were assigned by inspection to assist in visualization of the data separation

not partial, coefficients. The table shows that the highest correlation is between B4/B6 (0.33), followed by B4/B5 (0.20) and B5/B6 (0.16). Note that B4, B5 and B6 are the chief contributors of PC1, having correlation coefficients with PC1 of 0.53, 0.68, and 0.76, respectively. In other words, the angles which have the largest correlation coefficients with PC1 also have the largest angle-angle correlation coefficients. There is less correlation between A1/B3 (0.12), A1/B6 (−0.14), B1/B5 (−0.14), and B3/B4 (0.11).

Analog 3

Conformational analysis

The random search produced 632 conformers for **3** with relative energy ranging from 0 to 18 kcal mol^{-1}. The torsional angles A1, A2, B1, . . . B6 of the GEM conformer of **3** are given in Table 2. They are almost identical to those for **2**. These values were used to carry out the GEM scaling described in the Methods section.

SVD

As with **2**, Table 5 shows that for **3** no single principal component explains a large percentage of the variance; the first three PCs explain only 55% of the variance. PC1, which explains 22% of variance, is correlated to B2 (−0.55) and B5 (−0.76), with lesser correlations to B1 (−0.41) and B6 (−0.36). The chief contributors to PC2 are B1 (−0.81) and B6 (0.57), while A2 (−0.95) is the chief contributor to PC3. For the remaining principal components, the highest correlated angles are: B2 (−0.69) and B6 (0.57) for PC4, A1 (0.79) for PC5, B2 (0.50), A1 (−0.39) and B5 (−0.42) for PC6, B4 (0.91) for PC7, and B3 (−0.97) for PC8.

Score plots (not shown) constructed for all possible combinations of two PCs, from PC1 to PC7 do not show any data separation. This means that the variables responsible for those PCs do not separate the data into groups. However, the score plot of each PC vs. PC8 shows that the data separate into three groups along the PC8 axis. Fig. 5 shows the score plot of PC1 vs. PC8 and is typical of the other plots. Since the chief contributor to PC8 is B3 (−0.97), only that angle is responsible for the distinctive non-diagonal separation of conformers into three groups along PC8. The three groups are due to the staggered conformations that are the local minima for rotation around the $C(sp^3)$-$O(sp^3)$ bond of B3. The pattern of data separation in Fig. 5 for **3** is different than the diagonal separation seen in Figs. 2 and 4 for **2**. In those cases each PC has significant correlation coefficients with two or more angles and this group of angles is responsible for the data separation. In contrast, B3 makes by far the dominant contribution to PC8, leading to clean separation along the PC8 axis.

The correlation coefficients between all the eight angles for the analog **3** data are given in Table 6. The largest coefficient is 0.27 between B2 and B5, with lesser correlations between A1/B4 (0.12), A1/B6 (−0.12), B1/B5 (0.13), B2/B3 (0.11), B3/B4 (0.10), B4/B5 (0.13) and B4/B5 (0.14). As with **2** (Table 4), the angles which have the largest correlation coefficients with

Table 6 Angle/angle correlation coefficients, analog **3**

	A1	A2	B1	B2	B3	B4	B5	B6
A1	1	0.02	0.09	0.05	0.09	**0.12**	−0.02	**−0.12**
A2	0.02	1	−0.02	−0.05	−0.06	0.01	0.01	0.03
B1	0.09	−0.02	1	−0.02	0.02	0.03	**0.13**	−0.06
B2	0.05	−0.05	−0.02	1	**0.11**	0.08	**0.27**	0.04
B3	0.09	−0.06	0.02	**0.11**	1	**0.10**	**0.13**	0.01
B4	**0.12**	0.01	0.03	0.08	**0.10**	1	**0.13**	**0.14**
B5	−0.02	0.01	**0.13**	**0.27**	**0.13**	**0.13**	1	0.04
B6	**−0.12**	0.03	−0.06	0.04	0.01	**0.14**	0.04	1

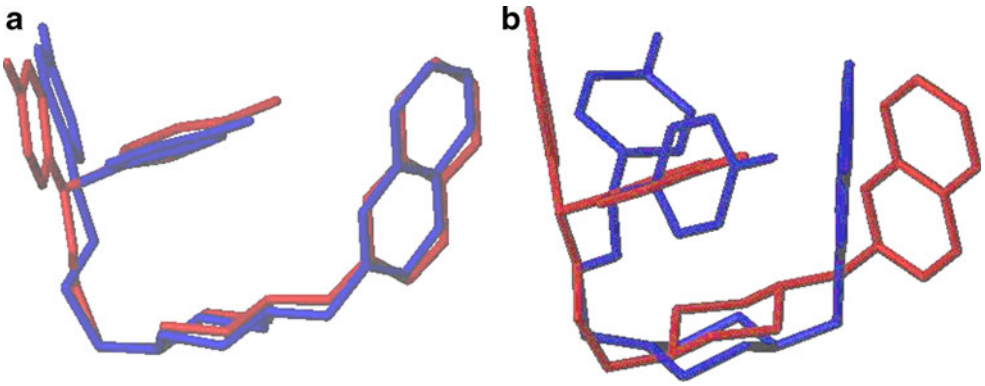

Fig. 6 GEM conformers of analog **2** (blue) and analog **3** (red); (**a**) MMFF94 force field and charges and (**b**) Tripos force field and Gasteiger-Hückel charges

PC1 (here, B5 (−0.76) and B2 (−0.55)) also have the largest angle-angle correlation coefficient.

Discussion

SVD of 2 and 3: MMFF94 force field

The random search conformational analysis produced conformer populations of about the same size for **2** and **3**. As shown in Fig. 6a, the GEM structures of each analog are almost identical. Comparison of Tables 3 and 4 shows that each principal component contributes almost the same percentage to the total variance for both **2** and **3**, and that no single principal component makes a dominant contribution. For both analogs, B5 is the chief contributor to PC1, B1 to PC2, A2 to PC3, B2 to PC4, B4 to PC7, and B3 to PC8. In both cases the angles which have large correlation coefficients with PC1 also have significant correlation coefficients with each other.

Data separation is seen along the diagonal for the score plots of PC2 vs. PC3 and PC1 vs. PC5 for **2** because for those PCs, no single angle is the overwhelmingly dominant contributor. In contrast, date separation is seen along PC8 in the score plot of any PC vs. PC8 for **3** because B3 is the overwhelmingly dominant contribution to PC8. The score plots of the PCs for **2** and **3** show that the data separate differently for the two analogs. Torsional angles A2, B1, B4, B5, B6 are responsible for the data separation (or grouping of conformers) noted in Figs. 2 and 4 for **2**, whereas B3 is responsible for that seen in Fig. 5 for **3**. Therefore, SVD has identified those angles which are most responsible for the differences in the conformer populations of **2** and **3**. This illustrates the utility of SVD in uncovering subtle influences on conformer groups not seen by direct analysis of conformer plots in two-dimensional torsional angle space [55] as in Fig. 3.

SVD of 2 and 3: Tripos and MMFF94 force field comparison

The random search conformational analysis using the Tripos force field [55] produced conformer populations of about the same size for **2** as for **3**. The angles of the GEM structures are given in Table 2 and Fig. 6b shows that the GEM structures of **2** and **3** are somewhat different from each other, in contrast to the MMFF94 results. As in the MMFF94 case, each principal component contributes almost the same percentage to the total variance for both **2** and **3**, and no single principal component makes a dominant contribution [52]. For both analogs, A2 is the chief contributor to PC1, B2 and either B5 or B6 to PC2, A2 and B5 or B6 to PC3, and A1 to PC6 [52]. Both force fields show B4 and B3 to be the dominant contributors to PC7 and PC8, respectively. For **2**, the angles which have large correlation coefficients with PC1 in the Tripos force field also have significant correlation coefficients with each other.

As with the MMFF94 results, the score plots of the PCs for the conformer populations of **2** and **3** generated by the Tripos force field show that the data separate differently for the two analogs. Data separation occurs along PC1 (due to A1 and A2) for **2**, and along PC7 (due B4) and PC8 (due to B3) for **3** [52]. Except for the dominant contribution of B3 to PC8 in both force fields, the angles which are chiefly responsible for the grouping of the conformers of **2** and **3** are different in both force fields. This illustrates that SVD has uncovered subtle differences in the conformer populations of the two analogs generated by the two different force fields.

Summary

SVD uncovered subtle differences in the conformer populations of **2** and **3** generated by the MMFF94 force field. These differences are due to the fact that the conformers separate along different principal components in the case of **2** vs. **3**, and ultimately to the fact that different torsional angles are the chief contributors to these components. Separation of conformers into groups is the first step in selecting representative conformers as templates for CoMFA-type 3D-QSAR modeling.

Acknowledgments This work was funded in part by a grant DA018153 to C.A.V. from the National Institutes of Health. The authors would like to thank Jeelum Naik, Eun Kim, and Anuj Kumar for assistance with the calculations, and Kathleen Gilbert for helpful discussions.

References

1. Kuhar MJ, Ritz MC, Boja JW (1991) The dopamine hypothesis of the reinforcing properties of cocaine. Trends Neurosci 14:299–302
2. Singh S (2000) Chemistry, design, and structure-activity relationship of cocaine antagonists. Chem Rev 100:925–1024
3. Dutta AK, Davis MC, Reith MEA (2001) Rational design and synthesis of novel 2, 5-disubstituted *cis*- and *trans*-piperidine derivatives exhibiting differential activity for the dopamine transporter. Bioorg Med Chem Lett 11:2337–2340
4. Carroll FI (2003) 2002 Medicinal chemistry division award address: Monoamine transporters and opiod receptors. Targets for addiction therapy. J Med Chem 46:1775–1794
5. Dutta AK, Zhang S, Kolhatkar RB, Reith MEA (2003) Dopamine transporter as target for drug development of cocaine dependence medications. Eur J Pharmacol 479:93–106
6. Prisinzano T, Rice KC, Baumann MH, Rothman RB (2004) Development of neurochemical normalization ("agonist substitution") therapeutics for stimulant abuse: focus on the dopamine uptake inhibitor, GBR 12909. Curr Med Chem - CNS 4:47–59
7. Quirante J, Vila X, Bonjoch J, Kozikowski AP, Johnson KM (2004) 2, 3-Disubstituted 6-azabicyclo[3.2.1]octanes as novel dopamine transporter inhibitors. Bioorg Med Chem 12:1383–1391
8. Cini N, Danieli E, Menchi G, Trabocchi A, Bottoncetti A, Raspanti S, Pupi A, Guarna A (2006) 3-Aza-6, 8-dioxabicyclo[3.2.1]octanes as new enantiopure heteroatom-rich tropane-like ligands of human dopamine transporter. Bioorg Med Chem 14:5110–5120
9. Kim D-I, Deutsch HM, Ye X, Schweri MM (2007) Synthesis and pharmacology of site-specific cocaine abuse treatment agents: Restricted rotation analogues of methylphenidate. J Med Chem 50:2718–2731
10. Mishra M, Kolhatkar R, Zhen J, Parrington I, Reith MEA, Dutta AK (2008) Further structural optimization of cis-(6-benzhydryl-piperidin-3-yl)-benzylamine and 1, 4-diazabicyclo[3.3.1]nonane derivatives by introducing an exocyclic hydroxyl group: Interaction with dopamine, serotonin, and norepinephrine transporters. Bioorg Med Chem 16:2769–2778
11. Jin C, Navarro HA, Carroll FI (2008) Synthesis and receptor binding properties of 2β-alkynyl and 2β-(1, 2, 3-triazol)substituted 3β-(substituted phenyl)tropane derivatives. Bioorg Med Chem 16:5529–5535
12. Nielsen S, Pedersen CM, Hansen SG, Petersen MD, Sinning S, Wiborg O, Jensen HH, Bols M (2009) An extended study of dimeric phenyl tropanes. Bioorg Med Chem 17:4900–4909
13. Manning JR, Sexton T, Childers SR, Davies HML (2009) 1-Naphthyl and 4-indolyl arylalkylamines as selective monoamine reuptake inhibitors. Bioorg Med Chem Lett 19:58–61
14. Carroll FI, Blough BE, Abraham P, Mills AC, Holleman JA, Wolckenhauer SA, Decker AM, Landavazo A, McElroy KT, Navarro HA, Gatch MB, Forster MJ (2009) Synthesis and biological evaluation of bupropion analogues as potential pharmacotherapies for cocaine addiction. J Med Chem 52:6768–6781
15. Carroll FI, Gao Y, Rahman MA, Abrams P, Parham K, Lewin AH, Boja JW, Kuhar MJ (1991) Synthesis, ligand binding, QSAR, and CoMFA study of 3β-(p-substituted phenyl)tropane-2β-carboxylic acid methyl esters. J Med Chem 34:2719–2725
16. Carroll FI, Mascarella SW, Kuzemko MA, Gao Y, Abraham P, Lewin AH, Boja JW, Kuhar MJ (1994) Synthesis, ligand binding, and QSAR (CoMFA and classical) study of 3β-(3′-substituted phenyl)-, 3β-(4′-substituted phenyl)-, and 3β-(3′, 4′-disubstituted phenyl)tropane-2β-carboxylic acid methyl esters. J Med Chem 37:2865–2873
17. Lieske SF, Yang B, Eldefrawi ME, MacKerell JAD, Wright J (1998) (-)-3ß-Substituted ecgonine methyl esters as inhibitors for cocaine binding and dopamine uptake. J Med Chem 41:864–876
18. Zhu N, Harrison A, Trudell ML, Klein-Stevens CL (1999) QSAR and CoMFA study of cocaine analogs: Crystal and molecular structure of (-)-cocaine hydrochloride and *N*-methyl-3β-(p-fluorophenyl)tropane-2β-carboxylic acid methyl ester. Struct Chem 10:91–103
19. Newman AH, Izenwasser S, Robarge MJ, Kline RH (1999) CoMFA study of novel phenyl ring-substituted 3α-(diphenylmethoxy)tropane analogues at the dopamine transporter. J Med Chem 42:3502–3509
20. Robarge MJ, Agoston GE, Izenwasser S, Kopajtic T, George C, Katz JL, Newman AH (2000) Highly selective chiral *N*-substituted 3α-[*bis*(4′-fluorophenyl)methoxy]tropane analogues for the dopamine transporter: Synthesis and Comparative Molecular Field Analysis. J Med Chem 43:1085–1093
21. Davies HML, Gilliatt V, Kuhn LA, Saikali E, Ren P, Hammond PS, Sexton GJ, Childers SR (2001) Synthesis of 2β-Acyl-3β-(substituted naphthyl)-8-azabicyclo[3.2.1] octanes and their binding affinities at dopamine and serotonin transport sites. J Med Chem 44:1509–1515
22. Kulkarni SS, Newman AH, Houlihan WJ (2002) Three-dimensional quantitative structure-activity relationships of mazindol analogues at the dopamine transporter. J Med Chem 45:4119–4127
23. Kulkarni SS, Grundt P, Kopajtic T, Katz JL, Newman AH (2004) 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 47:3388–3398
24. Petukhov PA, Zhang J, Wang CZ, Ye YP, Johnson KM, Kozikowski AP (2004) Synthesis, molecular modeling, and biological studies of novel piperidine-based analogues of cocaine: Evidence of unfavorable interactions proximal to the 3α-position of the piperidine ring. J Med Chem 47:3009–3018
25. Yuan H, Kozikowski AP, Petukhov RA (2004) 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 47:6137–6143
26. Yuan H, Petukhov PA (2006) Improved 3D-QSAR CoMFA of the dopamine transporter blockers with multiple conformations using the genetic algorithm. Bioorg Med Chem Lett 16:6267–6272
27. Dutta AK, Xu C, Reith ME (1996) Structure-Activity relationship studies of novel 4-[2-[*bis*(4-fluorophenyl)methoxy]ethyl]-1-(3-phenylpropyl)piperidine analogs: Synthesis and biological evaluation at the dopamine and serotonin transporter sites. J Med Chem 39:749–756
28. Dutta AK, Meltzer PC, Madras BK (1993) Positional importance of the nitrogen atom in novel piperidine analogs of GBR 12909: Affinity and selectivity for the dopamine transporter. Med Chem Res 3:209–222
29. Matecka D, Lewis D, Rothman RB, Dersch CM, Wojnicki FHE, Glowa JR, De Vries AC, Pert A, Rice KC (1997) Heteroaromatic analogs of 1-[2-(diphenylmethoxy)ethyl]- and 1-[2-[*bis*(4-fluorophenyl)methoxy]ethyl]-4-3-phenylpropyl)piperazines (GBR 12935 and GBR 12909) as high-affinity dopamine reuptake inhibitors. J Med Chem 40:705–716
30. Bradley A, Izenwasser S, Wade D, Klein-Stevens C, Zhu N, Trudell ML (2002) Synthesis and dopamine transporter binding affinities of 3α -Benzyl-8-(diarylmethoxyethyl)-8-azabicyclo [3.2.1]octanes. Bioorg Med Chem Lett 12:2387–2390
31. Prisinzano T, Greiner E, Johnson IEM, Dersch CM, Marcus J, Partilla JS, Rothman RB, Jacobson AE, Rice KC (2002) Piperidine analogues of GBR 12909: High affinity ligands for the dopamine transporter. J Med Chem 45:4371–4374
32. Bradley A, Izenwasser S, Wade D, Cararas S, Trudell ML (2003) Synthesis and dopamine transporter selective 3-{2-(Diarylmethox-

yethylidene)}-8-alkylaryl-8-azabicyclo[3.2.1]octanes. Bioorg Med Chem Lett 13:629–632
33. Kimura M, Masuda T, Yamada K, Mitani M, Kubota N, Kawakatsu N, Kishii K, Inazu M, Kiuchi Y, Oguchi K, Namiki T (2003) Syntheses of novel diphenyl piperazine derivatives and their activities as inhibitors of dopamine uptake in the central nervous system. Bioorg Med Chem Lett 11:1621–1630
34. Greiner E, Prizinzano T, Johnson IEM, Dersch CM, Marcus J, Partilla JS, Rothman RB, Jacobson AE, Rice KC (2003) 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 46:1465–1469
35. Ghorai SK, Cook CD, Davis MC, Venkataraman SK, Beardsley PM, Reith MEA, Dutta AK (2003) High affinity hydroxypiperidine analogues of 4-(2-benzhydroxyloxyethyl)-1-(4-fluorobenzyl)piperadine for the dopamine transporter: Stereospecific interactions in vitro and in vivo. J Med Chem 46:1220–12208
36. Kolhatkar RB, Ghorai SK, George C, Reith MEA, Dutta AK (2003) Interaction of *cis*-(6-benzhydrylpiperidin-3-yl)benzylamine analogues with monoamine transporters: Structure-activity relationship study of structurally-constrained 3, 6-disubstituted piperidine analogues of (2, 2-Diphenylethyl)-[1-(4-fluorobenzyl) piperidin-4-ylmethyl]amine. J Med Chem 46:2205–2215
37. Kolhatkar RB, Cook CD, Ghorai SK, Deschamps JR, Beardsley PM, Reith MEA, Dutta AK (2004) Further structurally constrained analogues of *cis*-(6-benzhydrylpiperidin-3-yl)benzylamine with elucidation of bioactive conformation: Discovery of 1, 4-diazabicyclo[3.3.1]nonane derivatives and evaluation of their biological properties for the monoamine transporters. J Med Chem 47:5101–5113
38. Zhang S, Zhen J, Reith MEA, Dutta AK (2004) 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 12:6301–6315
39. Greiner E, Boos TL, Prisinzano TE, De Martino MG, Zeglis B, Dersch CM, Marcus J, Partilla JS, Rothman RB, Jacobson AE, Rice KC (2006) Design and synthesis of promiscuous high-affinity monoamine transporter ligands: unraveling transporter selectivity. J Med Chem 49:1766–1772
40. Boos TL, Greiner E, Calhoun J, Prisinzano TE, Nightingale B, Dersch CM, Rothman RB, Jacobson AE, Rice KC (2006) Structure-activity relationships of substituted *N*-benzyl piperidines in the GBR series: Synthesis of 4-(2-(bis(4-fluorophenyl) methoxy)ethyl)-1-(2-trifluoromethylbenzyl) piperidine, an allosteric modulator of the serotonin transporter. Bioorg Med Chem 14:3967–3973
41. Cramer RD III, Patterson DE, Bunce JD (1988) Comparative Molecular Field Analysis (CoMFA).1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110:5959–5967
42. Kim KH, Greco G, Novellino E (1998) A critical review of recent CoMFA applications. In: Kubinyi H, Folkers G, Martin YC (eds) 3D QSAR in Drug Design: Recent Advances. Kluwer Academic, Dordrecht, pp 257–315
43. Gilbert KM, Boos TL, Dersch CM, Greiner E, Jacobson AE, Lewis D, Matecka D, Prisinzano TE, Zhang Y, Rothman RB, Rice KC, Venanzi CA (2007) DAT/SERT selectivity of flexible GBR 12909 analogs Modeled using 3D-QSAR methods. Bioorg Med Chem 15:1146–1159
44. Shenkin PS, McDonald DQ (1994) Cluster analysis of molecular conformations. J Comput Chem 15:899–916
45. Feher M, Schmidt JM (2001) Metric and multidimensional scaling: Efficient tools for clustering molecular conformations. J Chem Inf Comput Sci 41:346–353
46. Feher M, Schmidt JM (2003) Fuzzy clustering as a means of selecting representative conformers and molecular alignments. J Chem Inf Comput Sci 43:810–818
47. Dave RN, Sen S (2002) Robust fuzzy clustering of relational data. IEEE Trans Fuzzy Syst 10:713–727
48. Jackson JE, (1990) A user's guide to principal component analysis. John Wiley and Sons, New York.
49. Gilbert KM, Venanzi CA (2006) Hierarchical clustering analysis of flexible GBR 12909 dialkyl piperazine and piperidine analogs. J Comput-Aided Mol Des 20:209–225
50. Misra M, Banerjee A, Davé RN, Venanzi CA (2005) Novel feature extraction technique for fuzzy relational clustering of a flexible dopamine reuptake inhibitor. J Chem Inf Model 45:610–623
51. Banerjee A, Misra M, Pai D, Shih L-Y, Woodley R, Lu X-J, Srinivasan AR, Olson WK, Dave RN, Venanzi CA (2007) Feature extraction using molecular planes for fuzzy relational clustering of a flexible dopamine reuptake inhibitor. J Chem Inf Model 47:2216–2227
52. Fiorentino A, Pandit D, Gilbert KM, Misra M, Dios R, Venanzi CA (2006) Singular value decomposition of torsional angles of analogs of the dopamine reuptake inhibitor GBR 12909. J Comput Chem 27:609–620
53. Saunders M (1987) Stochastic exploration of molecular mechanics energy surfaces. J Am Chem Soc 109:3150–3152
54. Clark M, Cramer RD III, van Opdenbosch N (1989) Validation of the general purpose Tripos 5.2 force field. J Comput Chem 10:982–1012
55. Pandit D, Misra M, Gilbert KM, Skawinski WJ, Venanzi CA (2010) Conformational analysis of piperazine and piperidine analogs of GBR 12909: stochastic approach to evaluating the effects of force fields and solvent. J Mol Model. doi:10.1007/s00894-010-0712-x
56. Halgren TA (1996) Merck Molecular Force Field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17:490–519
57. Halgren TA (1996) Merck Molecular Force Field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. J Comput Chem 17:520–552
58. Halgren TA (1996) Merck Molecular Force Field. III. Molecular geometries and vibrational frequencies for MMFF94. J Comput Chem 17:553
59. Halgren TA (1996) Merck Molecular Force Field. IV. Conformational energies and geometries for MMFF94. J Comput Chem 17:587–615
60. Beckers MLM, Buydens LMC (1998) Multivariate analysis of a data matrix containing A-DNA and B-DNA dinucleoside monophosphate steps: Multidimensional Ramachandran plots for nucleic acids. J Comput Chem 19:695–715
61. Reijmers TH, Wehrens R, Buydens LMC (2001) Circular effects in representations of an RNA nucleotide Data set in relation with principal component analysis. Chemom Intell Lab Syst 56:61–71
62. Wall ME, Rechtesteiner A, Rocha LM (2003) Singular Value decomposition and principal component analysis. In: Berrar DP, Dubitsky W, Granzow M (eds) A practical approach to microarray data analysis, Kluwer,Dordrecht, pp 91-109