A Comparison of Field-Based Similarity Searching Methods: CatShape, FBSS, and ROCS

Kirstin Moffat,[†] Valerie J. Gillet,*[,†] Martin Whittle,[†] Gianpaolo Bravi,[‡] and Andrew R. Leach[‡]

Department of Information Studies, University of Sheffield, Western Bank, Sheffield, S10 2TN, United Kingdom, and GlaxoSmithKline, Gunnels Wood Road, Stevenage, SG1 2NY, United Kingdom

Received April 11, 2007

Three field-based similarity methods are compared in retrospective virtual screening experiments. The methods are the CatShape module of CATALYST, ROCS, and an in-house program developed at the University of Sheffield called FBSS. The programs are used in both rigid and flexible searches carried out in the MDL Drug Data Report. UNITY 2D fingerprints are also used to provide a comparison with a more traditional approach to similarity searching, and similarity based on simple whole-molecule properties is used to provide a baseline for the more sophisticated searches. Overall, UNITY 2D fingerprints and ROCS with the chemical force field option gave comparable performance and were superior to the shape-only 3D methods. When the flexible methods were compared with the rigid methods, it was generally found that the flexible methods gave slightly better results than their respective rigid methods; however, the increased performance did not justify the additional computational cost required.

INTRODUCTION

Similarity methods for searching databases of chemical structures were first introduced over two decades ago. Such methods are normally applied early in a drug discovery program when little is known about the biological target of interest. Once one or more lead compounds have been identified, more sophisticated methods can be used to improve on the properties of the initial lead(s). The first approaches to similarity searching were based on two-dimensional (2D) properties of molecules.[1,2] While these early methods proved very effective at finding close analogues, the relationship between structure and biological activity is a complex one which has led to the development of a wide range of different similarity searching methods.[3]

The first 2D approaches were based on describing molecules at the atom and bond levels. Subsequent work has focused on describing the properties of atoms rather than element type, and a number of such descriptors have been developed, for example, the early binding property pairs[4] and the more recent Chemically Advanced Template Search (CATS),[5] Similog,[6] and the Functional Connectivity Fingerprints (FCFPs) descriptors from SciTegic.[7] Other approaches have been based on graph representations of molecules.[8–10]

There has also been considerable interest in representing molecules by their three-dimensional properties. The three-dimensional (3D) methods can be divided into alignment-independent methods where the molecules are typically represented by vectors which can be compared using techniques similar to those used for 2D methods, for example, using the Tanimoto coefficient.[11,12] The more computationally demanding methods require that the molecules are first aligned prior to calculating their similarity (a review of alignment methods is provided by Lemmen and Langauer[13]). While appealing from a molecular recognition perspective, a major challenge for 3D methods is the handling of conformational flexibility.

The typical approach to comparing similarity searching methods is to carry out retrospective searches based on compounds of known activity. The active compounds are seeded into a database of inactive compounds, or decoys; one of the actives is used as the query compound, and the methods are compared on their ability to rank highly compounds sharing the same activity. When studies of this type have compared 2D methods with 3D methods, the 2D methods have typically resulted in the highest recalls.[14–16] However, a more recent criterion on which methods are evaluated is their ability to identify compounds with similar bioactivity that have different molecular frameworks.[5,9,17–19] This has become known as "scaffold hopping"[20] and is of considerable interest since it may lead to different chemical series that could be exploited should one series reach a dead-end, for example, due to difficult chemistry or poor absorption, distribution, metabolism, and excretion properties. Recent studies have reported on the effectiveness of 3D methods for scaffold-hopping applications.[17,18,21]

Here, three field-based similarity methods are evaluated for similarity searching. They consist of two commercially available programs, namely, the CatShape[22] module of CATALYST and ROCS (Rapid Overlay of Chemical Structures),[18,23] and an in-house program developed at the University of Sheffield called FBSS (Field-Based Similarity Searcher).[24–26] UNITY 2D fingerprints are also used to provide a comparison with a more traditional approach to similarity searching, and simple whole-molecule descriptors are used to provide a baseline for the more sophisticated

* Corresponding author e-mail: v.gillet@sheffield.ac.uk.
[†] University of Sheffield.
[‡] GlaxoSmithKline.

10.1021/ci700130j CCC: $40.75 © 2008 American Chemical Society
Published on Web 03/20/2008

searches.[27] The methods are compared on their hit rates and also on their ability to recognize actives from different lead series.

3D Similarity Methods. The FBSS program[28] uses a genetic algorithm (GA) to search for the alignment of two molecules that maximizes the similarity between their steric, electrostatic, or hydrophobic fields, or any combination thereof. The fields are calculated from the molecular electrostatic potentials, electron densities, or molecular lipophilicity potentials, respectively, and are represented by atom-centered Gaussian approximations according to the method developed by Good et al.[29,30] The GA encodes translations and rotations of the database molecule relative to the query molecule, and the fitness function measures the similarity of the two fields on the basis of the alignment encoded in a chromosome. Similarity is calculated using the Carbó coefficient applied to Gaussian representations of the fields. The GA is configured to search for the alignment that maximizes the similarity. When more than one field is being used to calculate the similarity between two molecules, the mean of the similarities calculated for each individual field is used, and the final alignment is that which maximizes this mean similarity.[25] FBSS can operate in *rigid* mode or *flexible* mode. In rigid mode, the conformations of both the query molecule and the database molecule are fixed. In flexible mode, the conformation of the query is usually fixed and the conformation of the database molecule is varied by including its torsion angles within the chromosome. A simple van der Waals bump check is used to eliminate high-energy conformations.[28]

The CatShape[22] module in CATALYST is a shape-based similarity searching method. The van der Waals surface of a molecule is calculated and represented as a set of points of uniform average density on a grid. The surface points enclose a volume on the grid. The geometric center of the set of points is computed along with the three principal component vectors passing through the center. The maximum extents along each principal axis and the total volume are calculated. These provide shape indices that can be compared with the query and used in an initial screening step to eliminate poor matches from further consideration, on the basis of user-defined tolerances. Any database molecules which pass the screening step are then processed further. A molecule is aligned with the query by superimposing their geometric centers and aligning the major and minor axes. A steepest-descent optimization algorithm is then used to optimize the volume overlap. This iterative process applies a series of rotational and translational realignments until the volume overlap cannot be improved. The grid volumes are then compared using a Tanimoto score which is the intersection divided by the union of the query and target grid volumes. In contrast to the on-the-fly approach to handling conformational flexibility implemented in FBSS, in CatShape conformational flexibility is handled by precomputing an ensemble of conformers for each compound and comparing each conformer in turn.

ROCS[23] is a shape-based similarity method based on molecular volume. The volume overlap of two molecules is based on the overlap of Gaussians which have been parametrized to provide close approximations to hard-sphere atomic volumes. Two molecules are aligned by overlaying the geometric centers of the molecules and then aligning their maximum, intermediate, and minimum steric quadrupoles.[31] A rotation of 180° around each of the axes gives another orientation. All combinations of 180° rotations give rise to four unique orientations, each of which is used as a starting point for a gradient-based optimization. The gradient for the overlap is calculated using the atom-centered Gaussians with the maximum overlap obtained using the Tanimoto coefficient as for CatShape. It is also possible to include chemical force fields in the superposition and similarity score to allow the comparison of molecules on both shape and chemical complementarity. The chemical force field is also a Gaussian function centered on the atoms of the molecule, and in this case, the gradient for the overlap is calculated using the Gaussians representing volume and chemical properties. As with CatShape, conformational flexibility can be taken into account by precomputing an ensemble of conformers and comparing each in turn.

Data Sets. The relative performance of the methods was compared using activity classes extracted from the Protein Data Bank (PDB),[32] referred to here as the PDB data sets, and from the MDL Drug Data Report (MDDR),[33] referred to here as the Briem and Lessel data sets. The PDB data sets consist of two activity classes, HIV reverse transcriptase (HIV-RT) and thrombin inhibitors, where the ligands have been extracted from protein–ligand complexes in the PDB in their bound conformations. The ligands were extracted using Relibase+[34] and clustered in SYBYL[35] according to their Comparative Molecular Field Analysis (CoMFA)[36] steric fields, and one compound was selected from each cluster to form a set of queries. These two data sets provide a somewhat unrealistic experiment since the ligands are present in their bound conformations; however, they provide a sanity check for the methods. These data sets were used in a previous study by Patel et al.[37]

The Briem and Lessel data sets include the five activity classes investigated in their study.[38] 3D structures were not available for these compounds, and so they were clustered using Selector in SYBYL, and a diverse subset was selected from each activity class for use as queries. Conformations of the molecules were generated using CONCORD. The active conformations of these compounds are not known, and hence these provide a more realistic test than the above.

The activity classes are shown in Table 1, which indicates the size of each data set and the number of queries selected in each case.

A 1000-molecule subset of MDDR was selected at random for use as inactive compounds. The subset was found to contain a small number of additional actives corresponding to the Briem and Lessel data sets (hence, the numbers shown in Table 2 differ slightly from those in their publication). Note that the run time of FBSS was the limiting factor on the size of the data sets considered. Conformations were generated using CONCORD, with 11 compounds failing, leaving a final set of 989 inactive compounds. The distributions of various physicochemical properties (molecular weight, hydrogen-bond donors, hydrogen-bond acceptors, ring counts, and rotatable bonds) in the 989-compound subset were compared with the whole of MDDR, and no significant differences were found (data not shown). Thus, the inactives were considered to be representative of MDDR as a whole. It has been suggested recently that the decoys used in each enrichment experiment should be selected to have similar

Table 1. Activity Classes and Number of Queries Used As Targets[a]

activity class	abbreviated name	no. of actives	no. of queries	UNITY pairwise similarities
5HT3 antagonists	5HT3	47	7	0.374
ACE inhibitors	ACE	40	7	0.546
HIV reverse transcriptase	HIV-RT	31	10	0.311
HMG-CoA reductase inhibitors	HMG	108	11	0.397
PAF antagonists	PAF	128	12*	0.326
thrombin inhibitors	thrombin	36	7	0.355
TXA2 antagonists	TXA2	48	7	0.352

[a] The UNITY pairwise similarities are also shown to indicate the heterogeneity of each activity class. The initial selection of PAF queries included some large hyperflexible fatty acid moieties which are not druglike and would be difficult for any 3D methods due to their flexibility, and so a subset of 7 PAF queries was created by removing compounds 147515, 162976, 164002, KO-286011, and 179150. This is referred to as the PAF subset. All query compounds are available in the Supporting Information.

Table 2. The Number of Unique Scaffolds Represented by Each Activity Class

activity class	no. of compounds	no. of scaffolds	av. no. mols per scaffold
5HT3	47	20	2.4
ACE	40	10	4.0
HIV-RT	31	11	2.8
HMG	108	23	4.7
PAF	128	46	2.8
thrombin	36	16	2.3
TXA2	48	22	2.2

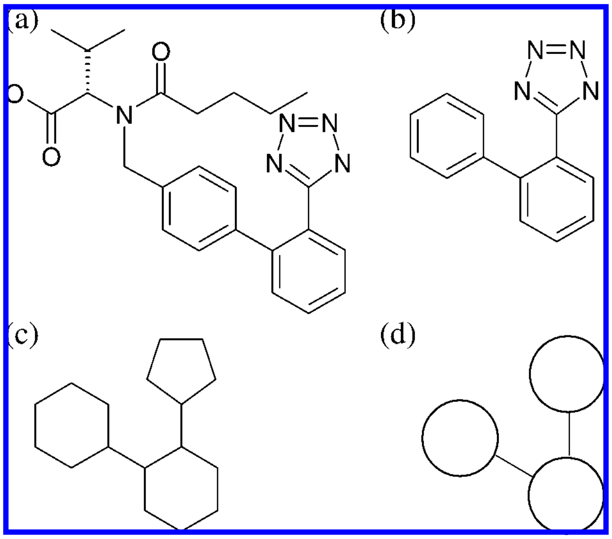


Figure 1. (a) Diovan, (b) the Murcko atomic framework, (c) the Murcko graph framework, and (d) the reduced graph scaffold.

physicochemical properties as the active compounds in order to provide a background that is equivalent for all active sets. An alternative strategy has been adopted here whereby similarity searches have also been carried out on the basis of simple whole-molecule properties to provide an above-random baseline for the more sophisticated search methods.

EXPERIMENTAL METHODS

Rigid Searches. Similarity searches were carried out for each activity class using each query compound in turn with all compounds represented by a single fixed conformation. For the HIV-RT and thrombin activity classes, the rigid searches used the bound conformations of the active compounds including the queries, whereas the inactive compounds were in their CONCORD conformations. As discussed above, this search is somewhat unrealistic of the usual situation in which similarity searching would be used. For the Briem and Lessel activity classes, the rigid searches were carried out with all compounds (actives and inactives) in their CONCORD conformations.

Flexible Searches. As mentioned earlier, there is a significant difference in the way in which FBSS considers conformational flexibility compared to CatShape and ROCS. In FBSS, conformational space is explored on-the-fly, by encoding the torsion angles of the database molecule within the chromosome. The starting conformations for the PDB activity classes were their bound conformations, whereas the starting conformations for all other compounds (the Briem and Lessel activity classes and the inactive compounds) were the CONCORD-generated structures. The queries were treated as rigid and consisted of bound conformations (PDB data sets) and CONCORD conformations (Briem and Lessel data sets).

CatShape and ROCS both handle flexibility through the use of precomputed conformations. In order to allow comparison of the shape-searching aspects of the two methods, the same set of conformers was used in each case. The FAST mode of Catalyst was used to generate conformers with up to 50 generated per molecule with a threshold of 20 kcal/mol. The bound conformations were also included for the PDB data sets, and the CONCORD-generated structures were included for the Briem and Lessel data sets and for the inactive compounds. As for FBSS, the queries were treated as rigid and consisted of bound conformations (PDB data sets) and CONCORD conformations (Briem and Lessel data sets). It should be noted that the choice of Catalyst for conformer generation may result in a bias toward CatShape, which was presumably optimized using Catalyst conformers.

Similarity searches were also carried out using UNITY fingerprints so that the 3D methods could be compared with a more traditional 2D search. A baseline for all of the searches was provided by carrying out similarity searches based on simple whole-molecule descriptors. These were calculated using MOE software[39] (v2008.06) and consisted of numbers of hydrogen-bond donors, hydrogen-bond acceptors, aromatic bonds and ring bonds, molecular weight, and the second κ shape index. The descriptors were normalized in the range 0−1, and similarity was calculated using the Tanimoto coefficient.

Enrichment factors at 2% and 10% of the ranked lists resulting from the similarity searches were calculated for the rigid searches, and enrichments at 10% were calculated for the flexible searches. The results were averaged over all queries in each activity class. The degree to which the 3D methods complement one another was also investigated by comparing hit lists. Finally, the extent to which the methods are able to recall active compounds that belong to different lead series (scaffold hop) was also investigated. All compounds (actives and inactives) were reduced to scaffolds by the procedure shown in Figure 1. First, the Murcko graph

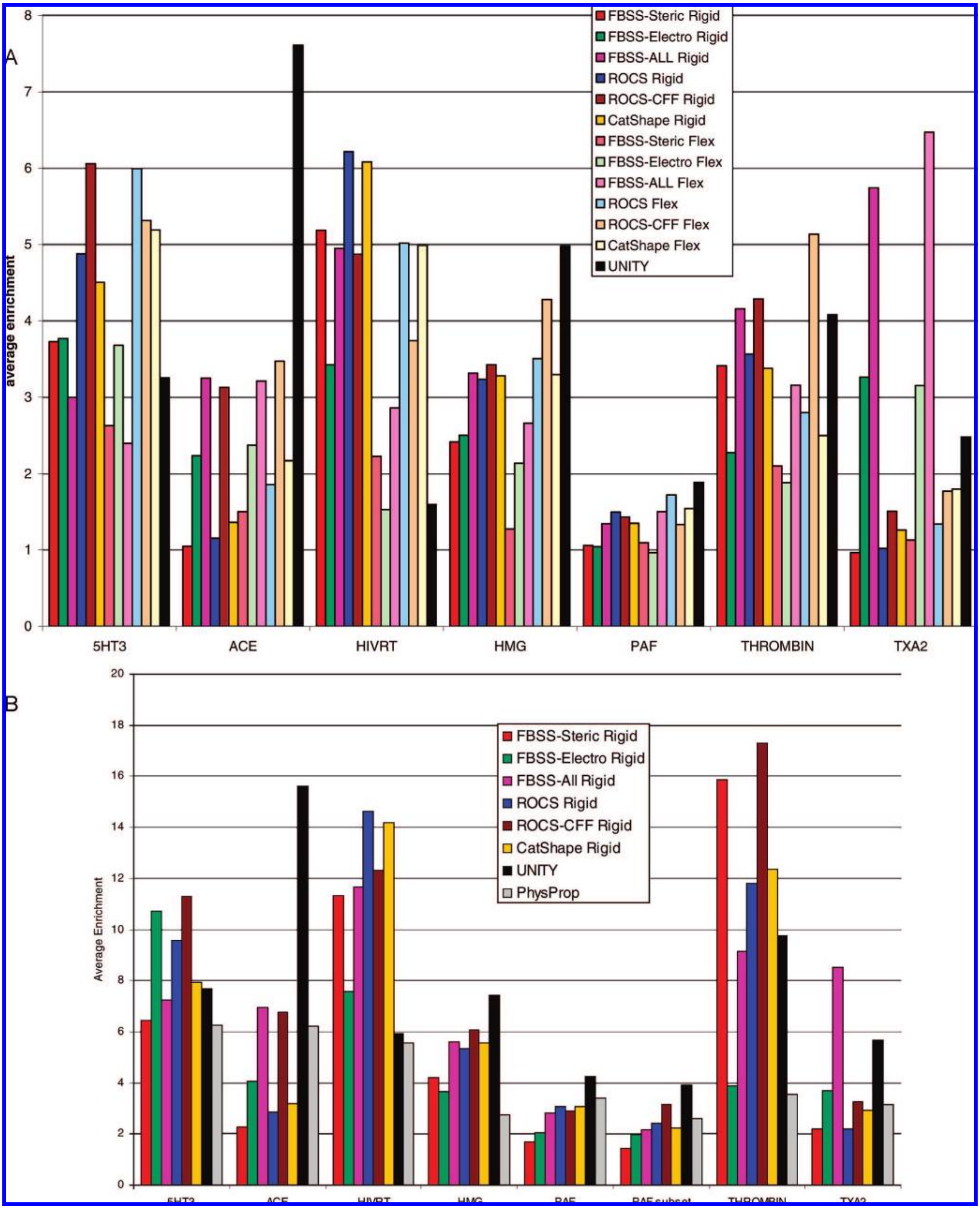


Figure 2. (a) Enrichments at 10% of the ranked list averaged over all queries in each activity class. Rigid and flexible searches are shown for each of the 3D methods. The FBSS searches include sterics only (FBSS-Steric); electrostatics only (FBSS-Electro); and steric, electrostatic, and hydrophobic fields combined (FBSS-ALL). The ROCS searches include shape only (ROCS) and shape and chemical complementarity combined (ROCS-CFF). The CatShape searches are shape only. (b) Enrichments at 2% of the ranked list averaged over all queries in each activity class for the rigid searches. Enrichments based on the simple whole molecule properties are also shown (PhysProp). Results are also presented for the PAF subset; see the text for details.

framework[40] of the compound was generated using Pipeline Pilot,[41] and this was then converted to a reduced graph using in-house software.[42] The resulting reduced graphs, referred to here as scaffolds, are similar to the reduced-skeletal cyclic

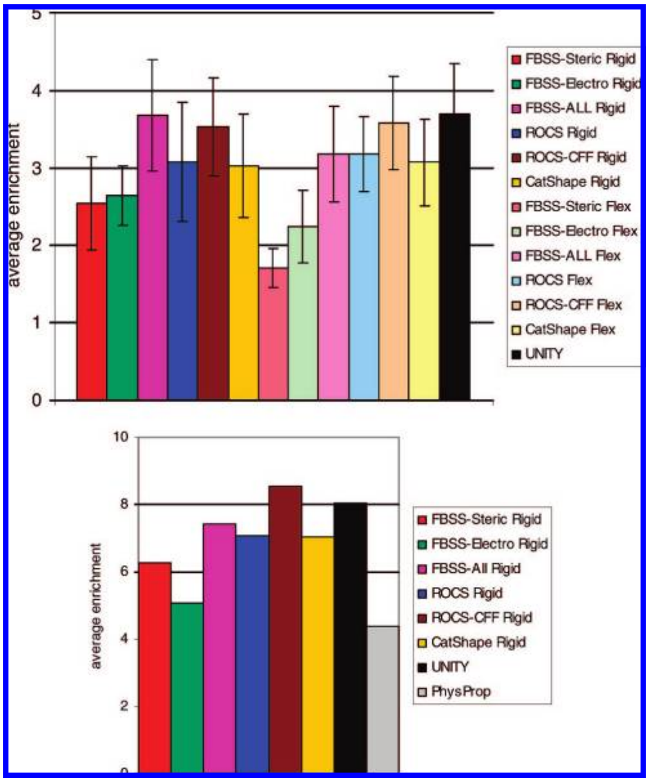


Figure 3. (a) Enrichments at 10% averaged over all activity classes. The searches carried out are as described in Figure 2. (b) Enrichments at 2% averaged over all activity classes. The searches carried out are as described in Figure 2b.

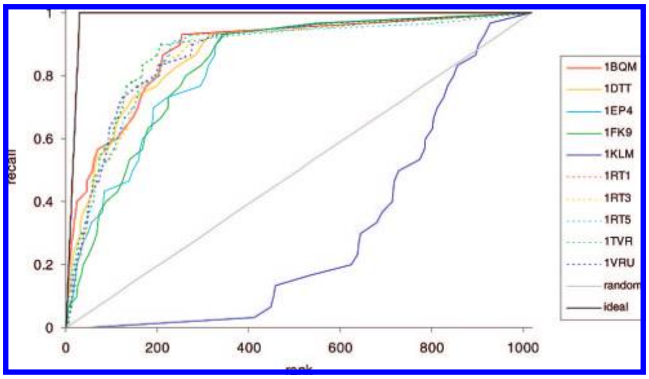


Figure 4. Enrichment plots for all of the queries in the HIV-RT data set using FBSS-Steric. Note that ROCS and CatShape also consider 1klm as an outlier.

systems described by Xu and Johnson,[43] and they provide abstract representations of molecules which describe their connectivity but disregard the type of atoms, the bond orders involved, and the number of atoms comprising a linker or ring system. The number of unique scaffolds in each activity class was determined and is shown in Table 2. For each query, all compounds (actives and inactives) having the same skeleton as the query were removed from the hit list, and the rate at which actives with unique scaffolds were accumulated was noted.

Program Details. *FBSS.* FBSS was run using three different field types: steric fields; electrostatic fields; and steric, electrostatic, and hydrophobic fields combined (referred to as ALL fields). Hydrophobic fields were not used on their own since previous studies have shown that they are not as effective as steric or electrostatic fields.[16] The protonation states of the molecules were calculated to reflect likely protonation at physiological pH using a GlaxoSmithKline in-house program. Charges for the electrostatic fields were calculated using Mulliken charges and MOPAC AM1. FBSS was run with default settings (i.e., selection pressure of 1.1, GA population size of 125, number of iterations equal to 10 000).

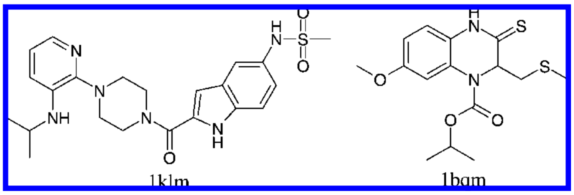


Figure 5. HIV-RT queries.

Table 3. Enrichments at 10% for the Thrombin Data Set[a]

FBSS-Flex	FBSS-Ensemble	ROCS	CatShape
2.10	2.68	2.80	2.50

[a] In FBSS-Flex, conformational space is explored on-the-fly. FBSS-Ensemble is based on rigid searching using the same set of conformers as used by ROCS and CatShape. The performance of FBSS is comparable to that of ROCS when using the same set of precomputed conformations. In contrast, performance is significantly reduced when conformational space is explored on-the-fly.

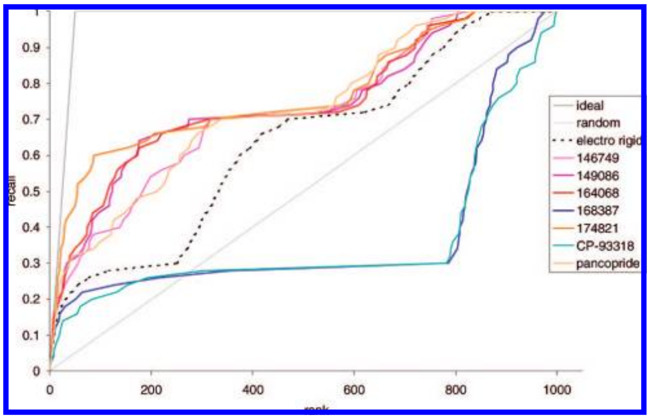


Figure 6. Enrichment plots for individual query compounds extracted from the 5HT3 queries for searches using the electrostatic fields in FBSS. The dashed line shows the average enrichment over all queries.

CatShape. The grid spacing was set to 1 Å, and the bit volume resolution (defining the size of the grid placed around the molecule) was 2 Å. The similarity tolerances were set to 0.0−1.0 so that all of the databases' compounds were ranked; that is, no screening was carried out. All of the parameters were set at the default values. All experiments were carried out with CatShape version 4.7.

ROCS. Searches were carried out using shape searching only and also with shape and chemical complementarity combined. The Tanimoto cutoff value was set to zero so that, as for CatShape, the entire set of molecules was ranked. Default values were used for all other parameters. The chemical complementarity searches used the ImplicitMills-Dean chemical force field, which defines six chemical types: hydrogen-bond donors, hydrogen-bond acceptors, hydrophobes, anions, cations, and rings. In ROCS, the molecules are first processed using a simple pK_a calculation so that the chemical types are assigned independent of the protonation

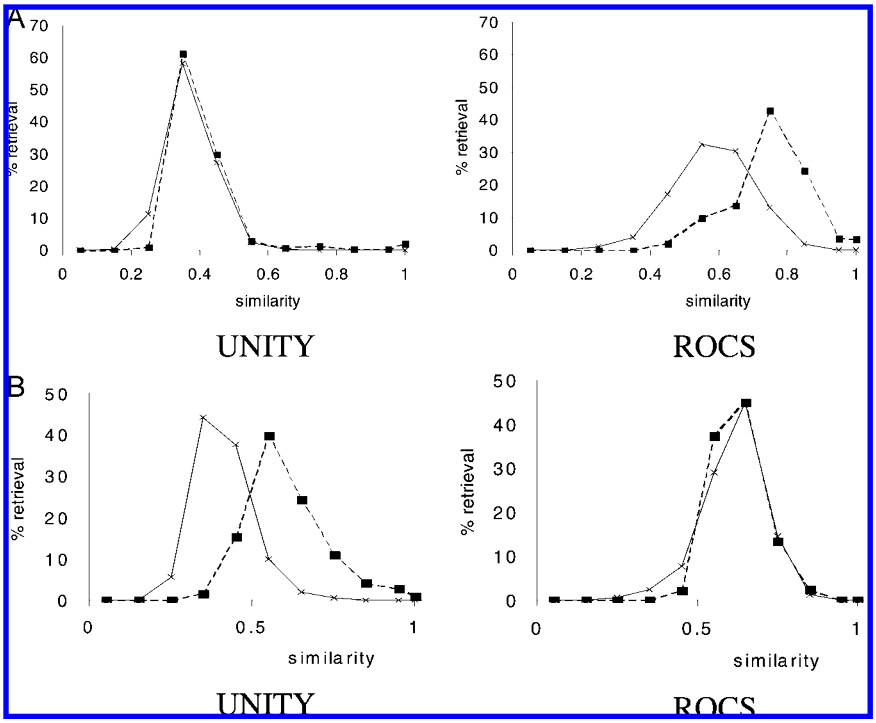


Figure 7. (a) Pairwise similarities for the HIV-RT data set. Distibutions of pairwise similarities between query compounds and actives (dashed lines) and the query compounds and the inactives (solid line) based on UNITY (left) and ROCS (right) similarities. (b) Pairwise similarities for the ACE data set. Distributions of pairwise similarities between query compounds and actives (dashed lines) and the query compounds and the inactives (solid line) based on UNITY (left) and ROCS (right) similarities.

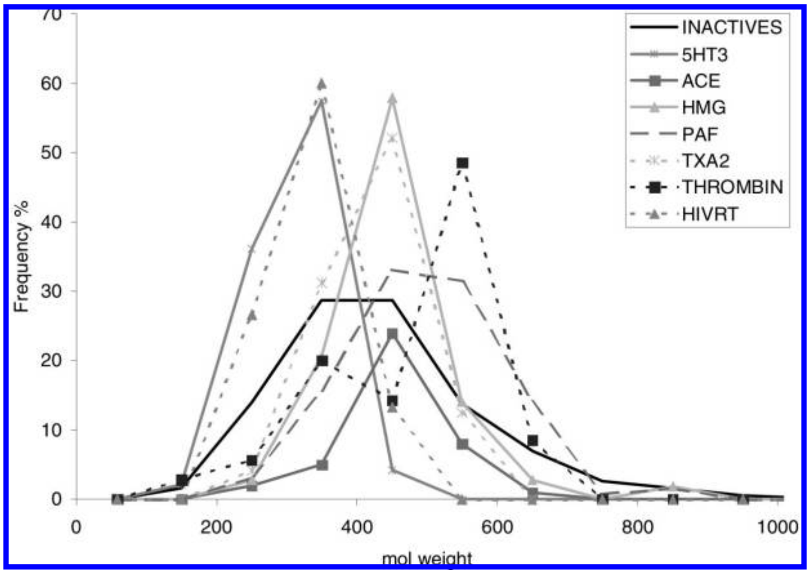


Figure 8. Distribution of molecular weights for the activity classes. The black line shows the distribution of molecular weights in the inactive compounds.

states specified in the input file. All experiments were carried out with ROCS2.2.

RESULTS AND DISCUSSION

Enrichments at 10% of the data set for all of the methods and all of the activity classes are shown in Figure 2a with enrichments at 2% also shown for the rigid searches in Figure 2b. The latter plot also shows enrichments for the PAF subset and enrichments based on the simple whole-molecule descriptors. Enrichments averaged over all activity classes are shown at the 10% and 2% thresholds in Figure 3a and b, respectively. It can be seen that the enrichments vary greatly depending upon the data set and the method used for calculating the similarity. Almost all of the methods show enrichments that are better than random, although in some instances, the increase in performance is small.

When averaged over all of the data sets and at the 10% threshold, UNITY shows the overall best performance, closely followed by ROCS-CFF (rigid and flexible). Shape-only ROCS and CatShape show comparable performance when compared with each other with little difference between the rigid and flexible results. As for ROCS, the inclusion of functional fields improves the performance of FBSS over the shape-only results. However, in contrast to ROCS and

Table 4. The Highest Overall Enrichment and the Best Field-Based Enrichment Obtained for All of the Data Sets at the 10% Enrichment Threshold

	best result		best field-based result	
	method	enrichment	method	enrichment
TXA2	FBSS-ALL Flex	6.47	FBSS-ALL Flex	6.47
THROMBIN	ROC-CFF Flex	5.14	ROC-CFF Flex	5.14
PAF	UNITY	1.89	ROCS Flex	1.72
HMG	UNITY	4.98	ROCS-CFF Flex	4.28
HIV-RT	ROCS Rigid	6.22	ROCS Rigid	6.22
ACE	UNITY	7.61	ROCS-CFF Flex	3.47
5HT3	ROCS-CFF Rigid	6.06	ROCS-CFF Rigid	6.06

CatShape, the treatment of conformational flexibility in FBSS leads to degradation in performance (discussed further below). At the 2% enrichment threshold, the results averaged over all data sets are broadly similar to those at the 10% threshold, except for the FBSS searches where the relative performance of the three different methods changes. However, it should be noted that the small size of the data set (approximately 1000 compounds) means that the top 20 positions only of the ranked list are examined at the 2% threshold level and the enrichments are very sensitive to actives occurring near the threshold. The searches based on the whole-molecule properties also show significant enrichments compared to random; however, they are outperformed by all of the more sophisticated search methods.

Steric Fields—Rigid Searches. When the steric rigid searches (FBSS-Steric Rigid, ROCS Rigid, and CAT Rigid) are compared, ROCS and CatShape outperform FBSS for four of the data sets (HMG, HIV-RT, 5HT3, and PAF) at both enrichment thresholds. The performances of the 3D methods over the other three data sets (TXA2, thrombin, and ACE) are comparable at 10% enrichment. However, FBSS-Steric outperforms both ROCS and CatShape for thrombin at the 2% threshold, where it is very effective at ranking the

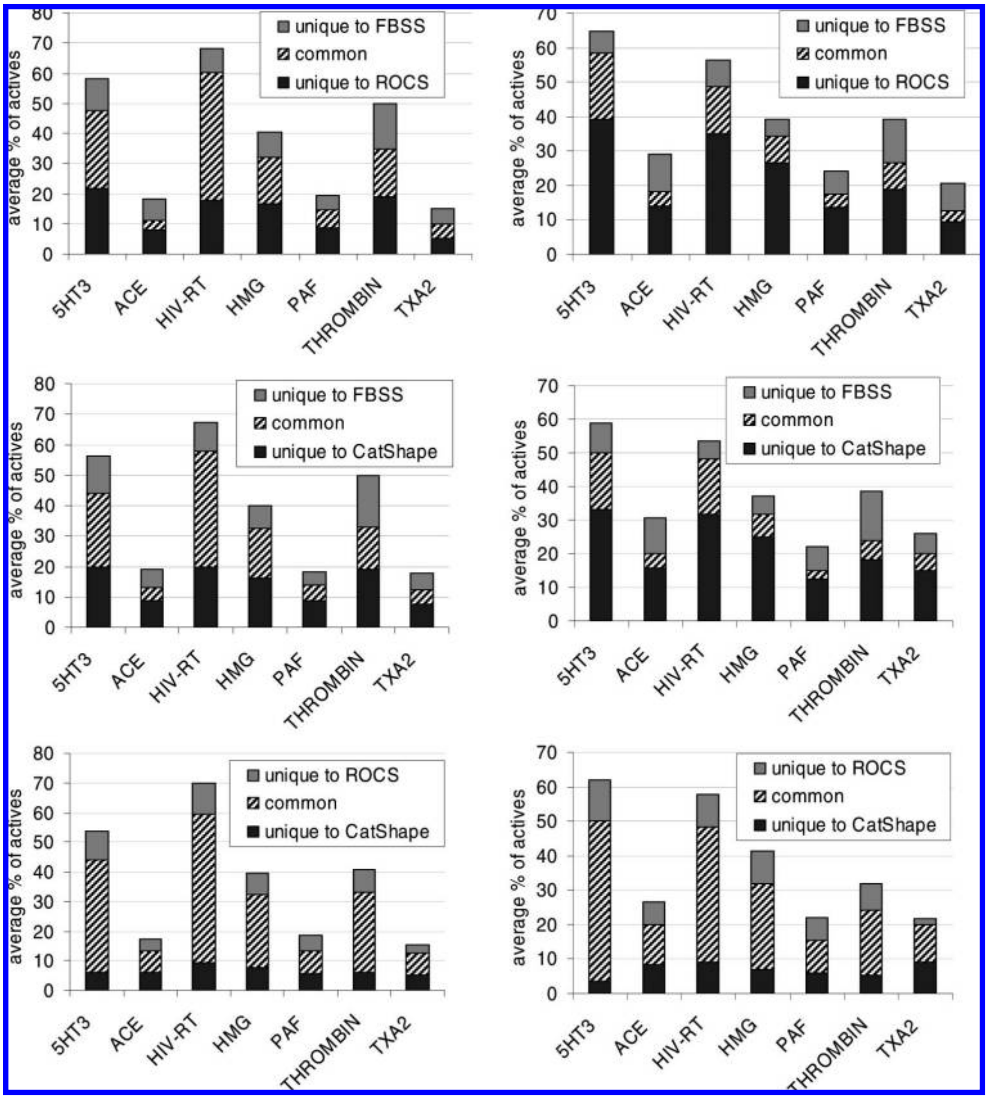
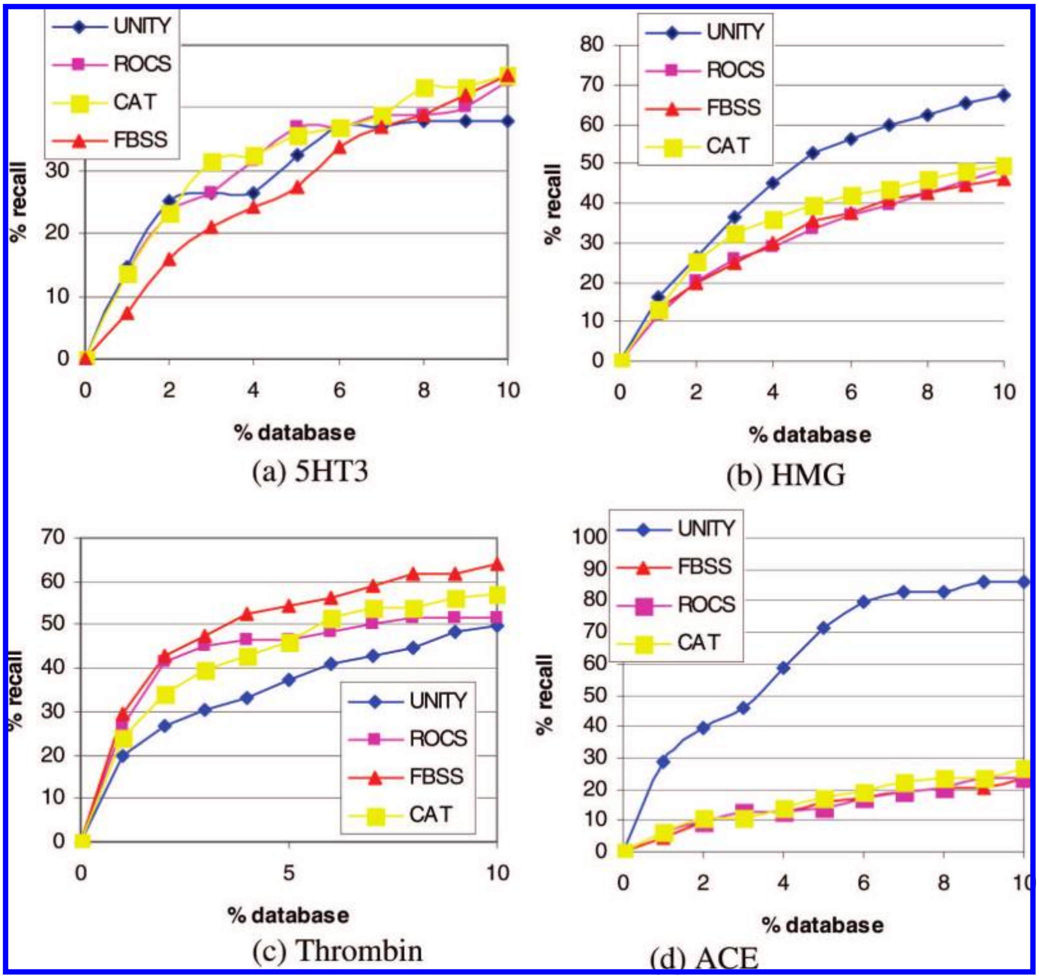


Figure 9. Overlap of hit lists. The left-hand plots represent the rigid searches, and the right-hand plots represent the flexible searches.



(a) 5HT3

(b) HMG

(c) Thrombin

(d) ACE

Figure 10. Enrichment plots showing the rate at which novel scaffolds are accumulated averaged over all queries in each of the activity classes.

Table 5. Times Taken to Search One Query against the Data Set of 1000 Compounds[a]

	rigid	Flex
FBSS steric-field	7 h	42 h
CatShape	5 min	1.5 h
ROCS	1.5 min	30 min
UNITY	0.5 min	

[a] Note that the Flex time for FBSS includes the on-the-fly conformational search, whereas for ROCS and CatShape, the times do not include the time taken to generate the conformers. Also note that FBSS was run for a fixed number of generations (10 000) although convergence was reached after an average of 3000 generations so that the FBSS times could be reduced by one-third with a negligible impact on performance.

actives very high in the ranked list. Overall, ROCS (shape-only) and CatShape show similar results at both enrichment thresholds.

The relative enrichments are not only activity-class-dependent but are also dependent on the particular query used. For example, Figure 4 shows the cumulative recall plots for each of the HIV-RT queries using FBSS-Steric Rigid, where it can be seen that query 1klm is an outlier. The structure of 1klm is shown in Figure 5 together with 1bqm, which is more typical in size to the rest of the active ligands. It can be seen that 1klm is more bulky, and when bound to the receptor, it extends out of the binding pocket into the solvent, unlike most of the other ligands in the data set, despite the mode of binding to the receptor being the same as for the other ligands.[44] The 3D methods are based on whole molecules, and thus 1klm is perceived as dissimilar to the other molecules in the same activity class. This is true for all of the shape-based methods. UNITY performs poorly for all queries in this data set.

Steric Fields—Flexible Searches. The thrombin and HIV-RT data sets generally show greater enrichments for the rigid searches than for the flexible searches, across all of the 3D methods. However, the rigid searches are likely to be overestimates of the true performance since, in this case, the actives are present in the data sets in their bound conformations, whereas the inactives are in CONCORD-generated conformations.

Considering the Briem and Lessel data sets, the flexible searches of ROCS and CatShape generally give slightly better results than the rigid methods. For these data sets, the active conformations are not known, and in the rigid searches, all of the conformations are generated using CONCORD so that the active conformations are not necessarily present in the data sets. The presence of multiple conformers, therefore, increases the chances of finding conformations of the actives that are similar to the conformation of the query.

For the FBSS searches, there is no clear distinction between the rigid and flexible searches. Rigid FBSS outperforms flexible FBSS for the thrombin, HIV-RT, HMG, and 5HT3 data sets, whereas the opposite is true for the ACE, PAF, and TXA2 data sets. Apart from the TXA2 data set, the flexible FBSS searches give lower enrichments than either the flexible ROCS or CatShape methods. The generally poorer relative performance of the flexible searching in FBSS is likely to be due to limitations of the conformational search method implemented in FBSS. As mentioned previously, the only factor preventing the occurrence of highly strained conformers is a simple bump-checking routine.[16] Thus, it is possible that the inactives can be "twisted" into high-energy conformations which are more similar in shape to the target compound than both the original CONCORD conformations and the conformations explored in the active compounds. In contrast, the conformations considered by ROCS and CatShape are energetically reasonable due to the 20 kcal/mol threshold applied during conformer generation. A further explanation for the better performance of the ROCS and CatShape methods for flexible searches is that the conformations used in the rigid searches are also present in the multiconformer sets. Therefore, for any given pairwise comparison, the similarity returned should be at least as high as that achieved for the rigid search. The same is not true of the flexible FBSS searches, since there is no guarantee that the GA will explore the conformations used in the rigid searches.

The effectiveness of flexible FBSS searching can be improved by adopting the same strategy as used by ROCS and CatShape, that is, using precomputed conformations followed by rigid searching. The results are shown in Table 3, where it can be seen that the performance of steric FBSS on the thrombin data set is comparable to that of ROCS at the 10% enrichment threshold when the same set of precomputed conformers is used and FBSS is run in rigid mode considering each conformer in turn. However, in its current implementation, the run times of FBSS would make such an approach extremely time-consuming. The experiment serves to illustrate that the handling of conformational flexibility in FBSS can clearly be improved, for example, through the use of torsion libraries to prevent energetically unreasonable conformers from being generated.

Functional Fields. When considering electrostatics fields alone, the performance of FBSS improves for the 5HT3, TXA2, and ACE data sets relative to the steric searches using FBSS, ROCS, and CatShape. The performance based on electrostatic fields is further improved when the steric, electrostatic, and hydrophobic fields are combined (FBSS-ALL fields) for all activity classes, except for the 5HT3s at the 10% threshold and thrombin at the 2% threshold. Furthermore, the combined fields improve on the steric fields for five out of the seven data sets. In fact, on average, the ALL fields version of FBSS outperforms all of the 3D methods at the 10% threshold.

However, on inspection, it was found that the presence of a charge on a query has a strong influence on the electrostatic (and also the ALL fields) results. For example, the effect of charge on search performance for the 5HT3 data set is illustrated in Figure 6, which shows enrichment plots for the individual queries. A total of 25 out of the 40 actives contain a protonated amino group, with the remaining 15 being unprotonated. When the query compound is protonated, the 25 protonated actives are found near the top of the ranked list. However, when the query is unprotonated, the protonated actives are found to have low similarity to the target with all 25 appearing near the end of the ranked list. Thus, in effect, FBSS simply identifies the presence of a charge on a molecule rather than finding similarities over the whole of the molecular field. These effects were noted, to different extents, for all of the data sets. The two query compounds that result in a cumulative recall that falls below the random plot are both unprotonated (CP-93318 and 168387).

Inclusion of the chemical force field within ROCS (ROCS-CFF) also generally leads to improved performance compared to the shape-only searches. ROCS-CFF and FBSS-ALL fields show very similar enrichments over the majority of the data sets, with the exception of the TXA2 data set (all of which have a positive charge), where FBSS outperforms ROCS by a considerable margin, and the 5HT3 data set where the reverse is seen; that is, ROCS outperforms FBSS.

3D versus 2D. A comparison of the rigid 3D searches with the more traditional 2D method, UNITY, reveals considerable variation over the individual activity classes. UNITY outperforms all of the 3D methods by a wide margin for the ACE and HMG activity classes; however, the reverse effect is seen for the HIV-RT activity class, with the 3D methods outperforming UNITY.

Figure 7a shows the distributions of the UNITY pairwise similarities (left) and ROCS pairwise similarities (right) between the queries and the active compounds, and between the queries and inactive compounds for the HIV-RT data set. While the left-hand plot (UNITY) shows very little difference in the distributions, the right-hand plot (ROCS) shows a clear separation in two distributions. Both FBSS and CatShape show similar plots to that of ROCS. Figure 7b shows similar plots for the ACE active class. Here, the plots are reversed with little difference apparent between the two distributions for the ROCS plot (right), whereas a clear separation is evident in the UNITY plot.

The variation in UNITY enrichments over the activity classes corresponds closely to the average pairwise similarities: the ACE activity class shows high enrichment and low diversity, whereas the HIV-RTs are the most diverse and show the lowest enrichment. The enrichments seen for the shape methods follow the trend HIV-RT > 5HT3 > thrombin. Figure 8 shows the distributions of molecular weights in the activity classes, where it can be seen that the HIV-RT and 5HT3 actives are smaller on average than the inactives, while the thrombin actives tend to be larger than the inactives. Thus, as might be expected, the shape-based methods are better at discriminating on size than are the 2D methods. Given the differences in molecule size of these data sets, it might be expected that the whole-molecule properties would be effective at achieving good enrichments. However, although impressive enrichments are seen with the simple descriptors for 5HT3 and HIV-RT, they are outperformed by one or more of the 3D methods, and they perform poorly relative to all other methods on the thrombin data set. The 3D methods are no better than the simple descriptors for the ACE data set (where UNITY stands out), and all methods perform poorly on the PAF data sets.

Table 4 shows the best results for each of the data sets at the 10% enrichment threshold. It is seen that, overall, UNITY gives the best enrichments (three out of seven). When considering the 3D methods only, flexible ROCS-CFF gives the greatest enrichments.

Comparison of Actives Retrieved. Figure 9 shows the degree to which the actives retrieved by the different methods overlap for the steric searches at the 10% enrichment threshold. In general, ROCS and CatShape have the greatest overlap in the hits retrieved. Despite the overall poorer performance of FBSS in terms of enrichments, the method complements the other two approaches as shown by the significant numbers of hits retrieved by FBSS that are not found by either ROCS or CatShape.

Scaffold-Hopping Ability. Figure 10 shows the rate at which novel scaffolds are retrieved in four of the activity classes averaged over all queries. In the ACE and HMG activity classes, not only does UNITY perform better in terms of overall enrichments but it is also more effective at accumulating the novel scaffolds than the 3D methods. However, in the ACE class, this simply mirrors the overall poor performance of the 3D shape-based methods, which is barely above random, whereas UNITY is very effective. Furthermore, the ACE and HMG data sets are the least diverse, measured using both UNITY fingerprints and the average number of molecules per scaffold. In the 5HT3 class, the initial rates are similar for the 2D and 3D methods (ROCS and CatShape) with the 3D methods winning out as the lists are descended. In the thrombin activity class, the 3D methods perform better than UNITY at all ranks. Thus, there is some evidence that the 3D methods are more effective at scaffold-hopping for the more diverse data sets. It should be recognized here that the definition of scaffolds used has its limitations, as has been noted previously.[45] For example, the abstraction of atom and bond types can lead to artificial equivalences between compounds in different series, whereas the addition of a ring as a side chain can lead to compounds in the same series being represented by different scaffolds. The latter effect is apparent in the thrombin data set, which is represented by a relatively large number of scaffolds.

Timings. The timings for each program for one search of approximately 1000 compounds on a Silicon Graphics R10K Origin200 IRIX server running at 225 MHz are given in Table 5. ROCS is the fastest of the 3D methods, with CatShape approximately 3 times slower. FBSS is by far the slowest of the methods, especially when conformational flexibility is taken into account; however, it should be noted that the ROCS and CatShape times do not include the time taken to generate the conformers. Furthermore, in the experiments reported here, FBSS was run for a fixed number of generations (10 000), even though convergence was reached after around 3000 generations; thus, the FBSS times could be reduced by one-third with negligible impact on performance. However, even when both of these factors are taken into account, FBSS is considerably slower.

CONCLUSIONS

When averaged over all of the data sets, ROCS-CFF was found to give similar performance to the UNITY 2D fingerprints in terms of enrichments (although, as has been noted previously, the relative performance varied from one activity class to another and also across different queries selected from within an activity class). The shape-only 3D searches showed inferior performance relative to ROCS-CFF. These findings are consistent with a recent study by the Sheridan group,[46] which found that ROCS-CFF (referred to as ROCS-color in their work) gave similar performance to that of TOPOSOM (in-house 2D atom-pair descriptors). Furthermore, they also found that the inclusion of chemical atom-typing or functionality in ROCS improved its performance over the shape-only method. In our work, including multiple conformers in the 3D searches led to a slight improvement over the rigid searches; however, this was not sufficient to justify the extra computational cost involved. There was also some evidence to suggest that the 3D methods are more effective at scaffold hopping, for the more diverse data sets.

ACKNOWLEDGMENT

We gratefully acknowledge software support from Accelrys, Chemical Computing Group, Openeye, and SciTegic; MDL Information Systems Inc. for provision of the MDDR database; and funding support from GlaxoSmithKline and EPSRC. We thank Kristian Birchall for generating the reduced graph skeletons and Yogi Patel for generating the whole molecule descriptors.

Supporting Information Available: The query compounds on which the searches are based. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Willett, P.; Winterman, V.; Bawden, D. Implementation of nearest-neighbour searching in an online chemical structure search system. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 36–41.
(2) Carhart, R. E.; Smith, D. H.; Venkataraghavan, R. Atom pairs as molecular-features in structure activity studies - definition and applications. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 64–73.
(3) Sheridan, R. P.; Kearsley, S. K. Why do we need so many chemical similarity search methods. *Drug Discovery Today* **2002**, *7*, 903–911.
(4) Kearsley, S. K.; Sallamack, S.; Fluder, E. M.; Andose, J. D.; Mosley, R. T.; Sheridan, R. P. Chemical similarity using physiochemical property descriptors. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 118–127.
(5) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. "Scaffold-hopping" by topological pharmacophore search: A contribution to virtual screening. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2894–2896.
(6) Schuffenhauer, A.; Floersheim, P.; Acklin, P.; Jacoby, E. Similarity metrics for ligands reflecting the similarity of the target proteins. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 391–405.
(7) SciTegic, Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121.
(8) Raymond, J. W.; Gardiner, E. J.; Willett, P. Rascal: Calculation of graph similarity using maximum common edge subgraphs. *Comput. J.* **2002**, *45*, 631–644.
(9) Barker, E. J.; Cosgrove, D. A.; Gardiner, E. J.; Gillet, V. J.; Kitts, P.; Willett, P. Scaffold-hopping using clique detection applied to reduced graphs. *J. Chem. Inf. Model.* **2006**, *46*, 503–511.
(10) Rarey, M.; Dixon, J. S. Feature trees: A new molecular similarity measure based on tree matching. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 471–490.
(11) Abrahamian, E.; Fox, P. C.; Naerum, L.; Christensen, I. T.; Thogersen, H.; Clark, R. D. Efficient generation, storage, and manipulation of fully flexible pharmacophore multiplets and their use in 3-D similarity searching. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 458–468.
(12) Mason, J. S.; Morize, I.; Menard, P. R.; Cheney, D. L.; Hulme, C.; Labaudiniere, R. F. New 4-point pharmacophore method for molecular similarity and diversity applications: Overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures. *J. Med. Chem.* **1999**, *42*, 3251–3264.
(13) Lemmen, C.; Lengauer, T. Computational methods for the structural alignment of molecules. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 215–232.
(14) Brown, R. D.; Martin, Y. C. The information content of 2D and 3D structural descriptors relevant to ligand-receptor binding. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1–9.
(15) Brown, R. D.; Martin, Y. C. Use of structure-activity data to compare structure-based clustering methods and descriptors for use in compound selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572–584.

(16) Schuffenhauer, A.; Gillet, V. J.; Willett, P. Similarity searching in files of three-dimensional chemical structures: Analysis of the bioster database using two-dimensional fingerprints and molecular field descriptors. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 295–307.

(17) Bohl, M.; Dunbar, J.; Gifford, E. M.; Heritage, T.; Wild, D. J.; Willett, P.; Wilton, D. J. Scaffold searching: Automated identification of similar ring systems for the design of combinatorial libraries. *Quant. Struct.-Act. Relat.* **2002**, *21*, 590–597.

(18) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489–1495.

(19) Jenkins, J. L.; Glick, M.; Davies, J. W. A 3D similarity method for scaffold hopping from the known drugs or natural ligands to new chemotypes. *J. Med. Chem.* **2004**, *47*, 6144–6159.

(20) Böhm, H.-J.; Flohr, A.; Stahl, M. Scaffold hopping. *Drug Discovery Today: Technol.* **2004**, *1*, 217–224.

(21) Jenkins, J. L.; Glick, M.; Davies, J. W. A 3D similarity method for scaffold hopping from known drugs or natural ligands to new chemotypes. *J. Med. Chem.* **2004**, *47*, 6144–6159.

(22) Hahn, M. Three-dimensional shape-based searching of conformationally flexible compounds. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 80–86.

(23) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **1996**, *17*, 1653–1666.

(24) Wild, D. J.; Willett, P. Similarity searching in files of three-dimensional chemical structures. Alignment of molecular electrostatic potential fields with a genetic algorithm. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 159–167.

(25) Drayton, S. K.; Edwards, K.; Jewell, N.; Turner, D. B.; Wild, D. J.; Willett, P.; Wright, P. M.; Simmons, K. Similarity searching in files of three-dimensional chemical structures: Identification of bioactive molecules. *Internet J. Chem.* **1998**, *1*, CP3–U34.

(26) Thorner, D. A.; Willett, P.; Wright, P. M.; Taylor, R. Similarity searching in files of three-dimensional chemical structures: Representation and searching of molecular electrostatic potentials using field-graphs. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 163–174.

(27) Bender, A.; Glen, R. C. Discussion of measures of enrichment in virtual screening: Comparing the information content of descriptors with increasing levels of sophistication. *J. Chem. Inf. Model.* **2005**, *45*, 1369–1375.

(28) Thorner, D. A.; Wild, D. J.; Willett, P.; Wright, P. M. Similarity searching in files of three-dimensional chemical structures: Flexible field-based searching of molecular electrostatic potentials. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 900–908.

(29) Good, A. C.; Hodgkin, E. E.; Richards, W. G. Utilization of Gaussian functions for the rapid evaluation of molecular similarity. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 188–191.

(30) Good, A. C.; Richards, W. G. Rapid evaluation of shape similarity using Gaussian functions. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 112–116.

(31) Masek, B. B.; Merchant, A.; Matthew, J. B. Molecular skins - a new concept for quantitative shape-matching of a protein with its small-molecule mimics. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 193–202.

(32) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(33) MDL Information Systems Inc., Symyx Technologies, Inc., 3100 Central Expressway, Santa Clara, CA 95051.

(34) *Relibase+ v1.3.0*; Cambridge Crystallographic Data Centre: Cambridge, U.K.

(35) *SYBYL v7.0.* Tripos Inc., 1699 South Hanley Rd, St. Louis, MO, 63144.

(36) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular-field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.

(37) Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 653–681.

(38) Briem, H.; Lessel, U. F. In vitro and in silico affinity fingerprints: Finding similarities beyond structural classes. *Perspect. Drug Discovery Des.* **2000**, *20*, 231–244.

(39) *MOE*, version v2008.06; Chemical Computing Group: Montreal, Quebec, Canada.

(40) Bemis, G. W.; Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893.

(41) *Pipeline Pilot v6.1.1.* SciTegic, Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121.

(42) Gillet, V. J.; Willett, P.; Bradshaw, J. Similarity searching using reduced graphs. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 338–345.

(43) Xu, Y. J.; Johnson, M. Using molecular equivalence numbers to visually explore structural features that distinguish chemical libraries. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 912–926.

(44) Esnouf, R. M.; Stuart, D. I.; DeClercq, E.; Schwartz, E.; Balzarini, J. Models which explain the inhibition of reverse transcriptase by HIV-1-specific (thio)carboxanilide derivatives. *Biochem. Biophys. Res. Commun.* **1997**, *234*, 458–464.

(45) Schuffenhauer, A.; Ertl, P.; Roggo, S.; Wetzel, S.; Koch, M. A.; Waldmann, H. The scaffold tree - visualization of the scaffold universe by hierarchical scaffold classification. *J. Chem. Inf. Model.* **2007**, *47*, 47–58.

(46) McGaughey, G. B.; Sheridan, R. P.; Bayly, C. I.; Culberson, J. C.; Kreatsoulas, C.; Lindsley, S.; Maiorov, V.; Truchon, J. F.; Cornell, W. D. Comparison of topological, shape, and docking methods in virtual screening. *J. Chem. Inf. Model.* **2007**, *47*, 1504–1519.

CI700130J