J Comput Aided Mol Des (2009) 23:853–868
DOI 10.1007/s10822-009-9302-3

Computation of 3D queries for ROCS based virtual screens

Gregory J. Tawa · J. Christian Baber · Christine Humblet

Received: 26 August 2009 / Accepted: 15 September 2009 / Published online: 26 September 2009

Abstract Rapid overlay of chemical structures (ROCS) is a method that aligns molecules based on shape and/or chemical similarity. It is often used in 3D ligand-based virtual screening. Given a query consisting of a single conformation of an active molecule ROCS can generate highly enriched hit lists. Typically the chosen query conformation is a minimum energy structure. Can better enrichment be obtained using conformations other than the minimum energy structure? To answer this question a methodology has been developed called CORAL (COnformational analysis, Rocs ALignment). For a given set of molecule conformations it computes optimized conformations for ROCS screening. It does so by clustering all conformations of a chosen molecule set using pairwise ROCS combo scores. The best representative conformation is that which has the highest average overlap with the rest of the conformations in the cluster. It is these best representative conformations that are then used for virtual screening. CORAL was tested by performing virtual screening experiments with the 40 DUD (Directory of Useful Decoys) data sets. Both CORAL and minimum energy queries were used. The recognition capability of each query was quantified as the area under the ROC curve (AUC). Results show that the CORAL AUC values are on average larger than the minimum energy AUC values. This demonstrates that one can indeed obtain better ROCS enrichments with conformations other than the minimum energy structure. As a result, CORAL analysis can be a valuable first step in virtual screening workflows using ROCS.

Keywords Ligand-based virtual screening · ROCS · Optimized query conformation · ROC curve analysis · Statistical significance · Virtual screening workflow

G. J. Tawa (✉) · C. Humblet
Chemical Sciences, Wyeth Research, CN 8000, Princeton, NJ, USA
e-mail: tawag@wyeth.com; gtawa@hotmail.com

J. C. Baber
Chemical Sciences, Wyeth Research, 200 Cambridge Park Drive, Cambridge, MA 02140, USA

Introduction

In many drug discovery projects no structure exists for the protein target. This is typically the case, for example, with GPCR targets [1–7]. However, even though there may be no known protein structure for the target of interest, there is often a known set of ligands active against the target of interest. As such, ligand-based efforts can be employed, i.e., finding new ligands by evaluating the *similarity* between a proposed ligand and the set of known active compounds [8]. In ligand based efforts, known actives are collected in order to develop queries for virtual screening or alignments for ligand-based design.

Many tools are available to perform these functions. One tool that is extensively used is ROCS (Rapid overlay of chemical structures) [9–11]. ROCS is a 3D method that aligns compounds based on shape and/or chemical similarity. ROCS can be used for virtual screening [8, 12] and ligand based optimization during later stages of a project.

The activities involved in running a ROCS virtual screen or performing a ROCS-based alignment for design always involve choosing a reference or bioactive conformation [13–19]. Finding good reference conformations for virtual screening is not always easy to do, especially for a diverse set of ligands. In this case, multiple queries or reference

conformations are necessary in order to cover the various structural classes of ligands. In addition, for ligand-based design, an alignment is necessary for each class.

To facilitate these ligand-based activities, we developed a suite of tools called CORAL (COnformational analysis, Rocs ALignment). CORAL will take an arbitrary set of ligands in 3D, perform a conformational expansion of each ligand, cluster the conformations into various groups, generate the centroid of each cluster, and then align all members of the cluster to the centroid. The centroids are then used as 3D conformational queries in ROCS-based virtual screening, and the cluster alignments can be used for 3D ligand-based design. The virtue of CORAL is its ability to navigate through the complexities of multi-molecule multi-conformation space and focus the user only on the most relevant regions defined by cluster centroids and their associated alignments.

Rapid overlay of chemical structures can generate highly enriched hit lists when the query is a single conformation of an active molecule [12, 16, 17, 20]. It is well known that ROCS performance is highly dependant on the selection of query molecules [17]. On the other hand, the current literature suggests that virtual screening performance of ROCS is essentially unaffected by choice of query conformation [16, 17]. As a result, the user typically chooses the minimum energy structure. CORAL, however, derives molecule query conformations that are usually different from the molecule's minimum energy structures. But are the so derived CORAL queries better than minimum energy structures for ROCS-based screening?

In order to answer this question, we evaluated CORAL and minimum energy queries in screening experiments using the DUD [21] data set. In the evaluation, CORAL was first used to derive one optimal query conformation for each of the 40 DUD ligand sets. In each case, 1/3 of the ligands were randomly chosen as the training set. These were used to derive the one CORAL optimized conformation. The associated minimum energy conformations were also determined. Both the CORAL and minimum energy conformations were then tested for their ability to retrieve known actives from screening databases composed of the 40 DUD decoy sets, each seeded with the remaining 2/3 of the active ligands. ROC curves [22, 23] were calculated, and the recognition capability of each query was quantified as the area under the ROC curve at 1–5, 10, 20, and 100% of the database retrieved.

Our results show that the CORAL AUC values are on average larger than the minimum energy AUC values. This demonstrates that one can indeed obtain better ROCS enrichments with conformations other than the minimum energy conformation. As a result, CORAL analysis can be a valuable first step in virtual screening workflows using ROCS.

Methods

Query preparation using CORAL

Once a given set of molecules is chosen, it is conformationally expanded using OMEGA [13, 24, 25] with the following parameters: nconfs = 50, rms = 0.6, Ewindow = 25. Then one selects a query conformation from this set for virtual screening. Query selection is idealized in Fig. 1. If several actives are known, each maps out a particular region of conformational space. This is shown by the vertical rectangles of Fig. 1. One can then ask the question, where are the conformations that have similar shape and chemistry orientation across the various conformation sets? These conformations lie in a narrow band shown in Fig. 1 as a horizontal yellow rectangle. Conformations in this band share similar shape and chemistry orientation. This narrow band is called the bioactive conformation set. The next step then is to find the best representative conformation in this set. In order to do this a matrix of overlap values is computed:

$$O_{ij}; \quad i \leq N,\ j \leq N \tag{1}$$

where O_{ij} = the pairwise combo score value between conformations i and j as aligned by ROCS [9–11] and N is the total number of conformations. After the matrix has been constructed the overlap average for each row of the matrix is computed by the formula

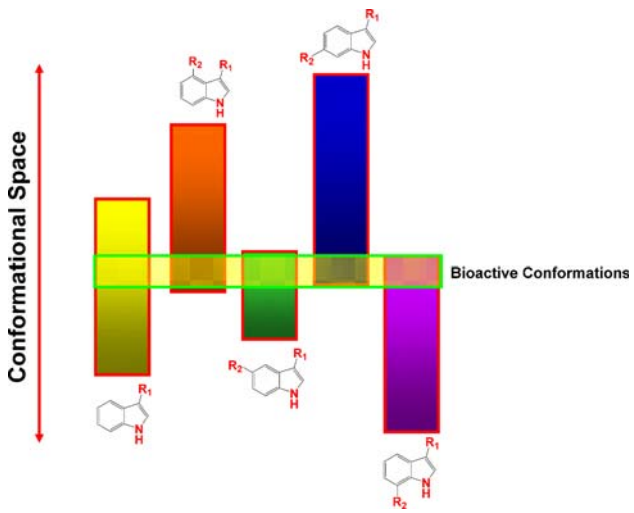


Fig. 1 Determination of bioactive conformations for a set of active ligands each spanning particular regions of conformational space (*vertical bars*). This can be represented by a narrow band of conformations (*horizontal yellow bar*) that share similar shape and chemistry orientation across all active ligands

$$\langle \Omega_i \rangle = \frac{1}{N} \sum_{j=1}^{N} O_{ij}; \quad i = 1 \rightarrow N. \tag{2}$$

The overlap average represents the degree of similarity that a given conformation i has with the rest of the conformations in the data set. The CORAL conformation is then computed as that conformation which has the largest overlap average

$$i_{\text{coral}} = \max[\langle \Omega_i \rangle]; \quad i = 1 - N. \tag{3}$$

The CORAL conformation is associated with one particular molecule of the data set. This molecule is determined, and its minimum energy structure, i_{min}, is also identified. The minimum energy structure is the lowest energy conformation in the set produced by OMEGA.

Virtual screening

The CORAL and minimum energy conformations are each tested for their ability to retrieve known actives from screening databases. For this study the DUD [21] data set is used to construct the screening databases. The DUD data set was originally designed for docking-based work. In its construction, Huang et al. [21, 26] remove a major pathology of commonly used screening data sets, that is, decoys being too different from the actives. In the DUD set decoys are chosen to match some simple molecular properties of the actives. These property matched decoys are a more difficult case than more general decoy sets.

Test set, training set, queries, and screening database construction

The work flow used to generate test sets, training sets, queries, and screening databases is shown in Fig. 2. In our experiment active ligand and decoy sets for the 40 DUD targets were used. For each DUD target (denoted by the index j in Fig. 2) the active ligands were partitioned into ten training sets and ten test sets (denoted by the index i, j in Fig. 2). Each training set was 1/3 the size of the active ligand set and the ligands were chosen at random from the active ligand set. Each test set was composed of the remaining 2/3 of the active ligands. The ten test sets were then seeded into the decoy set (indexed by j in Fig. 2) to create ten screening databases (indexed i, j). Each screening database was then conformationally expanded using OMEGA with parameters as described above. CORAL analysis was performed on the ten training sets to generate ten CORAL queries and ten associated minimum energy queries (both indexed i, j in Fig. 2) as previously described. The index i runs from 1 to 10, denoting the ten training, test, and screening databases generated per DUD target.

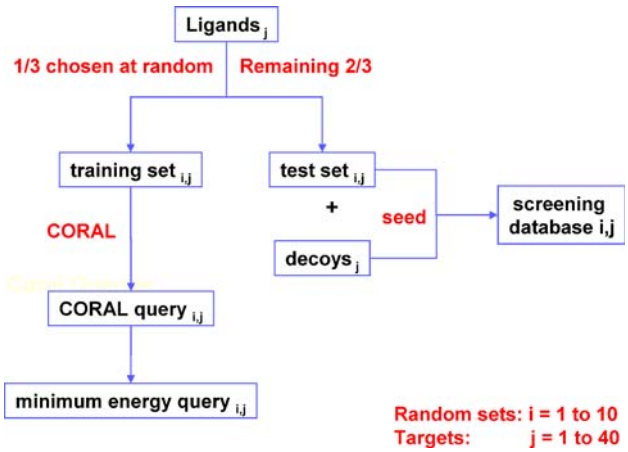


Fig. 2 Computation of test sets, training sets, queries, and screening databases using the actives and decoys for each of the 40 DUD targets. The index i ranges from 1 to 10 and it enumerates the various queries generated (either CORAL or minimum energy) for each DUD target. The index j ranges from 1 to 40 and it enumerates the particular DUD targets

The index j runs from 1 to 40, denoting the various DUD targets. So in total, 400 CORAL queries, 400 minimum energy queries and 400 screening databases were constructed. Each CORAL and minimum energy query is associated with one and only one of the 400 screening databases. This bootstrapping approach for generating training, test, and screening data is designed to generate multiple alternative versions of data. Using the alternative data sets, 400 screens for each type of query can be performed. The large number of screens enables better estimates of AUC distributions and therefore enhances our ability to resolve performance differences (if any) between each type of query.

Virtual screens, ROC curves, and AUC values

The virtual screening workflow is shown in Fig. 3. For each DUD target, denoted by j in Fig. 3, the ten CORAL queries and their associated minimum energy queries were used to screen the ten databases. Thus, for each target, 20 ROC curves were computed, ten for the CORAL queries and ten for the minimum energy queries.

Note in Fig. 3 that the index X designates the percentage of the database screened. Values of $X = 1$, 2, 3, 4, 5, 10, 20, and 100 were chosen to assess both early enrichment and overall performance. For each target, ten CORAL AUC values and ten minimum energy AUC values were computed at each X. Also ten AUC differences ($\Delta X_{i,j}$) were computed as CORAL $\text{AUC}_{i,j}$ − minimum energy $\text{AUC}_{i,j}$.

For each value of X then we have a distribution of 400 CORAL AUC values and a distribution of 400 minimum energy AUC values. In order to determine if these

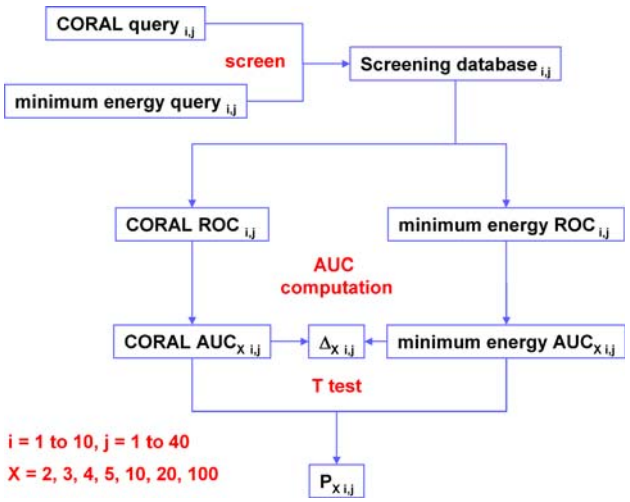


Fig. 3 Virtual screening, ROC AUC, and *T* test work flow. The index *i* ranges from 1 to 10 and it enumerates the various queries generated (either CORAL or minimum energy) for each DUD target. The index *j* ranges from 1 to 40 and it enumerates the particular DUD targets. The index *X* designates the percentage of the database screened

distributions are significantly different, a paired *T* test [27] was performed and significance value *p* was computed at the 95% confidence level ($\alpha = 0.05$).

Results

Performance of ROCS using CORAL derived queries and minimum energy structure queries

Results at the most general level are given in Tables 1 and 2. It shows average AUC values and *T* test statistics computed at $X = 1$, 2, 3, 4, 5, 10, 20, and 100% database retrieved for CORAL and minimum energy queries, respectively. The averages are compiled across 400 ROC curves for each query type. Results are also shown for choosing compounds at random. Note that at $X = 100\%$ the mean CORAL and minimum energy AUC values are 0.842 and 0.835, respectively. These AUC values are considerably better than the AUC value of 0.5 which is the case for a random selection of compounds. Overall ROCS obtains good virtual screening performance using either CORAL or minimum energy structures as queries.

Tables 1 and 2 also show that at all values of *X*, CORAL and minimum energy average AUC values are larger than the AUC values for a random selection of compounds. More importantly, however, is the fact that at all values of *X* the CORAL average AUC values are larger than the minimum energy average AUC values. This alone does not prove that CORAL queries outperform minimum energy queries. For that we appeal to the *T* test statistics listed in the table. For $X = 4$, 5, 10, 20, and 100%, $p(X) < 0.01$. This means that there is a significant difference between the CORAL results and minimum energy results. This combined with the fact that the CORAL AUC values are larger than the minimum energy AUC values means that CORAL queries outperform the minimum energy queries. For $X \leq 3\%$ the differences in CORAL and minimum energy AUC values are not significant because $p(X) > 0.05$. At $X = 3\%$, however, $p(X) = 0.06$ which is close to being statistically significant. To resolve the difference (if any) between CORAL and minimum energy AUC values at *X* values $\leq 3\%$ more screens are necessary.

A more detailed view of the results is given in Table 3. In this case, the results are compiled by individual DUD target. The internal diversity of each of the DUD active ligand sets, shown in column 2 of Table 3, was computed using the method of Turner and Willett [28]. In this computation only a single representative of each active compound was utilized. Different representatives of an active molecule appear contiguously in the DUD active ligand sets. The single retained representative was always chosen to be the first representative of a contiguous block. In the method of Turner and Willet the diversity of a ligand set is given by:

$$D(T) = 1 - \frac{\sum_{J=1}^{N(T)} \sum_{K=1}^{N(T)} \mathrm{SIM}(J,K)}{N(T)^2}. \qquad (4)$$

In our case *T* designates a particular DUD target (e.g. HIVPR, HMGA …), *D(T)* is the normalized diversity measure of the active ligand set associated with DUD target *T*, SIM*(J, K)* is the cosine similarity between two molecules *J* and *K* of the active ligand set, and *N(T)* is the number of molecules in the active ligand set. In our particular computation each molecule is represented by a vector of descriptors based on TRIPOS [29] fingerprints. A value of 1 for *D(A)* corresponds to a set of compounds that are as diverse as possible. A value of 0 for *D(A)* is just the opposite.

These diversity values are given in Table 3 along with average ROC AUC values and standard deviations for CORAL queries (columns 3 and 4, respectively) and minimum energy queries (columns 5 and 6, respectively). Column 7 shows average CORAL AUC − minimum energy AUC differences and column 8 shows the average percentage change in AUC when using CORAL queries as opposed to minimum energy queries. This value is given by

$$\%\,\mathrm{Change} = 100 \times \frac{\mathrm{Difference}}{\mathrm{Minimum\ energy\ AUC}} \qquad (5)$$

Finally, column 9 of Table 3 gives the *p* value at $X = 100\%$. The averages, standard deviations, and $p(100\%)$ values quoted in Table 3 are computed using the ten individual screening experiments performed for each DUD target. This is in contrast to Tables 1 and 2,

Table 1 Average AUC values and *T* test statistics for CORAL and minimum energy queries at 1–5% database retrieved

Query	AUC 1%		AUC 2%		AUC 3%		AUC 4%		AUC 5%	
	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
CORAL	0.000492	0.0004	0.003189	0.001	0.0078	0.003	0.0131	0.005	0.0183	0.008
Minimum energy	0.000489	0.0004	0.003186	0.001	0.0077	0.003	0.0129	0.005	0.0178	0.008
Random compound selection	0.00005		0.0002		0.00045		0.0008		0.00125	
T ($T_{critical} = 1.96$)	0.23		0.12		1.89		3.82		5.03	
p	0.82		0.90		0.06		<0.01		<0.01	

Averages are compiled across 400 ROC curves for each query type. Note that the results at 4 and 5% are statistically significant at the 95% confidence level whereas the results at 1, 2, and 3% are not statistically significant at the 95% confidence level. The number of decimal places shown for the AUC values is the minimal number necessary to show differences in the AUC values

Table 2 Average ROC AUC values and *T* test statistics for CORAL and minimum energy queries at 10, 20, and 100% database retrieved

Query	AUC 10%		AUC 20%		AUC 100%	
	Avg	SD	Avg	SD	Avg	SD
CORAL	0.051	0.02	0.122	0.04	0.842	0.1
Minimum energy	0.049	0.02	0.120	0.04	0.835	0.1
Random compound selection	0.005		0.02		0.5	
T ($T_{critical} = 1.96$)	6.47		5.63		3.71	
p	<0.01		<0.01		<0.01	

Averages are compiled across 400 ROC curves for each query type. Note that the results at all three levels are statistically significant at the 95% confidence level. The number of decimal places shown for the AUC values is the minimal number necessary to show differences in the AUC values

where each average was compiled across 400 total screening experiments, ten per DUD target × 40 DUD targets. Bold face black and red rows are for targets in which there is a statistically significant difference between the CORAL AUC distribution and the minimum energy AUC distribution at the 95 and 90% confidence levels, respectively. Note that in Table 3, the DUD targets are sorted decreasing by their diversity values. The diversity values range from 0.14 to 0.52 which means that, as far as diversity is concerned, the active ligands comprising the DUD data set are not very diverse.

Diversity is considered in Table 3 in order to investigate potential relationships between it and the average AUC values. A plot of the average AUC values versus active ligand diversity index at $X = 100\%$ database retrieved is given in the upper frame of Fig. 4. The points are all located in the upper right triangle region of the plot. The lower frame of Fig. 4 shows the number of DUD targets that have active ligand sets contained within various diversity sub-ranges. This distribution is symmetric and centered on index range 0.3–0.4.

The upper plot of Fig. 4 can be understood by first considering several facts. For each DUD target the query structures are derived from 1/3 of the active ligand sets, the other 2/3 of the active ligand sets are seeded into the decoys and are to be found by virtual screening. As the 2D diversity of the active ligand set decreases, the structures to be found should start looking very similar to the query in 2D. But we know that high similarity in 2D also means high similarity in 3D. This was shown by Bostrum et al. [30]. In that work it was found that if two molecules are structurally similar ($Tanimoto_{MCSS}$ [31]) then their binding modes to the same protein are similar, i.e., they have shape Tanimoto values >0.8. In terms of Fig. 4, this means that only high AUC values can exist for ligand sets exhibiting low (<0.2) 2D diversity values.

When the 2D diversity of the ligand sets increases a variety of 3D structures is possible, some similar, some dissimilar to the query. Therefore, a range of AUC values is generated during screening. This range gets larger as the ligand set diversity increases. This is can be seen in Fig. 4 for diversity values >0.2. When the ligand set diversity is between 0.2 and 0.3 the AUC range is 0.91–0.98. For diversity values between 0.3 and 0.4 the AUC ranges from 0.73 to 0.99. For diversity values between 0.4 and 0.5 the AUC ranges from 0.5 to 0.95. Clearly these AUC ranges are getting larger. For diversity values >0.5, we expect the AUC range to be yet larger. However, the low sampling of active ligand sets for diversity >0.5 (see lower frame of Fig. 4) prevents this from happening. Note that the lower frame of Fig. 4 shows that there is also a low sampling of active ligand sets with diversity <0.2. However, the AUC range in this region is not expected to increase with more sampling of ligand sets because high similarity in 2D = high similarity in 3D [30].

Figure 5 shows a plot of the average AUC differences (column 7 of Table 3) versus DUD target. The targets are ordered from left to right by decreasing active ligand set diversity. Note that a clear majority of the differences are positive. This shows that most of the time (27 out of 40 cases) the CORAL AUC values are larger than the minimum energy AUC values. Column 8 of Table 3 reveals that

Table 3 Average ROC AUC values for CORAL and minimum energy queries at 100% database retrieved

Target	Diversity	CORAL		Minimum energy		<Difference>	% Change	*p*
		<AUC>	SD	<AUC>	SD			
HIVRT	0.52	0.6007	0.1	0.5683	0.2	0.0324	5.70	0.59
ALR2	**0.51**	**0.5371**	**0.2**	**0.5524**	**0.2**	**−0.0154**	**−2.78**	**0.02**
CDK2	0.49	0.7065	0.08	0.6934	0.1	0.0130	1.88	0.45
VEGFR2	**0.49**	**0.6138**	**0.09**	**0.5991**	**0.08**	**0.0147**	**2.46**	**0.04**
COX1	0.49	0.4978	0.07	0.4965	0.06	0.0013	0.27	0.91
AR	0.46	0.6650	0.06	0.6529	0.05	0.0121	1.86	0.34
COX2	*0.45*	*0.9506*	*0.006*	*0.9485*	*0.005*	*0.0021*	*0.22*	*0.08*
PDE5	0.43	0.7209	0.1	0.7051	0.1	0.0158	2.24	0.17
INHA	**0.43**	**0.7057**	**0.08**	**0.6331**	**0.1**	**0.0726**	**11.46**	**<0.01**
PDGFRB	0.43	0.7960	0.02	0.7922	0.03	0.0039	0.49	0.14
ER-AG	0.42	0.9097	0.03	0.9120	0.02	−0.0024	−0.26	0.24
P38	0.4	0.9012	0.01	0.9019	0.01	−0.0008	−0.08	0.52
HIVPR	0.39	0.9169	0.03	0.9258	0.01	−0.0089	−0.97	0.35
PR	0.39	0.9343	0.02	0.9327	0.02	0.0016	0.17	0.77
GPB	0.37	0.9344	0.04	0.9406	0.02	−0.0062	−0.66	0.39
ER-ANT	0.37	0.8284	0.04	0.8277	0.05	0.0007	0.085	0.92
ADA	0.37	0.7847	0.1	0.7891	0.1	−0.0044	−0.56	0.63
ACHE	**0.37**	**0.7823**	**0.03**	**0.7661**	**0.03**	**0.0162**	**2.12**	**0.05**
SRC	**0.36**	**0.8148**	**0.02**	**0.7867**	**0.03**	**0.0282**	**3.58**	**0.03**
GR	0.36	0.8503	0.04	0.8469	0.05	0.0035	0.41	0.14
HSP90	**0.34**	**0.8332**	**0.05**	**0.7712**	**0.04**	**0.0620**	**8.03**	**<0.01**
MR	*0.33*	*0.7778*	*0.2*	*0.7888*	*0.2*	*−0.0110*	*−1.39*	*0.10*
EGFR	*0.33*	*0.9579*	*0.01*	*0.9531*	*0.009*	*0.0048*	*0.50*	*0.06*
TK	**0.33**	**0.8325**	**0.02**	**0.8493**	**0.01**	**−0.0167**	**−1.97**	**0.03**
DHFR	*0.33*	*0.9929*	*0.003*	*0.9904*	*0.002*	*0.0025*	*0.25*	*0.06*
COMT	0.33	0.7355	0.2	0.7351	0.2	0.0004	0.05	0.89
ACE	0.32	0.8502	0.07	0.8471	0.06	0.0031	0.36	0.66
NA	0.31	0.9887	0.008	0.9861	0.005	0.0026	0.27	0.32
Thrombin	**0.31**	**0.9365**	**0.02**	**0.9016**	**0.03**	**0.0349**	**3.87**	**<0.01**
FXA	0.3	0.9123	0.03	0.9108	0.02	0.0015	0.16	0.74
AMPC	0.29	0.9481	0.04	0.9425	0.04	0.0056	0.59	0.44
FGFR1	0.28	0.9192	0.04	0.9285	0.02	−0.0093	−1.00	0.40
Trypsin	**0.26**	**0.9077**	**0.04**	**0.8792**	**0.03**	**0.0285**	**3.24**	**<0.01**
PPAR-GAMMA	0.26	0.9379	0.02	0.9414	0.008	−0.0035	−0.37	0.48
PNP	0.24	0.9562	0.04	0.9582	0.03	−0.0020	−0.21	0.66
HMGR	0.23	0.9239	0.03	0.9188	0.02	0.0051	0.55	0.52
SAHH	0.22	0.9546	0.01	0.9484	0.02	0.0062	0.65	0.18
PARP	0.22	0.9792	0.02	0.9842	0.02	−0.0051	−0.52	0.17
RXR-ALPHA	0.18	0.9642	0.02	0.9668	0.02	−0.0026	−0.27	0.68
GART	0.14	0.9152	0.03	0.9091	0.02	0.0061	0.67	0.36

Results are given for individual DUD targets. Averages for each query type are compiled across ten ROC curves generated for each DUD target. Bold and italic entries are statistically significant at the 95 and 90% confidence levels, respectively. The number of decimal places shown for the AUC values and differences equals that of the smallest difference in the table

in most cases the magnitude of the average percent changes are small. There are, however, some targets where CORAL outperforms minimum energy by a large amount, e.g., HIVRT, INHA, HSP90, and Thrombin, where the percent changes are 5.70, 11.46, 8.03, and 3.87%, respectively. For cases where the minimum energy query outperforms

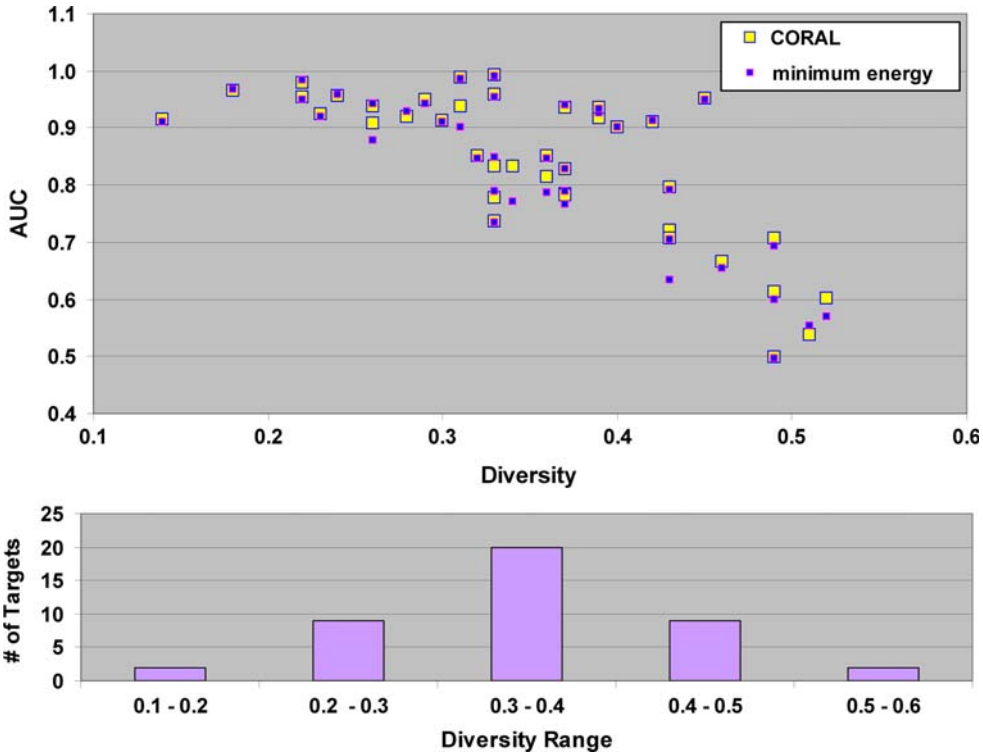


Fig. 4 *Top* average CORAL and minimum energy AUC values versus ligand set diversity index at 100% database retrieved. The AUC averages are compiled across the ten individual screening experiments performed for each DUD target. *Bottom* the number of DUD active ligand sets within five diversity sub-ranges. Sub-range 0.3–0.4 contains the most (18) DUD active ligand sets, sub-ranges 0.2–0.3 and 0.4–0.5 contain eight active ligand sets each, and the sub ranges 0.1–0.2 and 0.5–0.6 contain two active ligand sets each

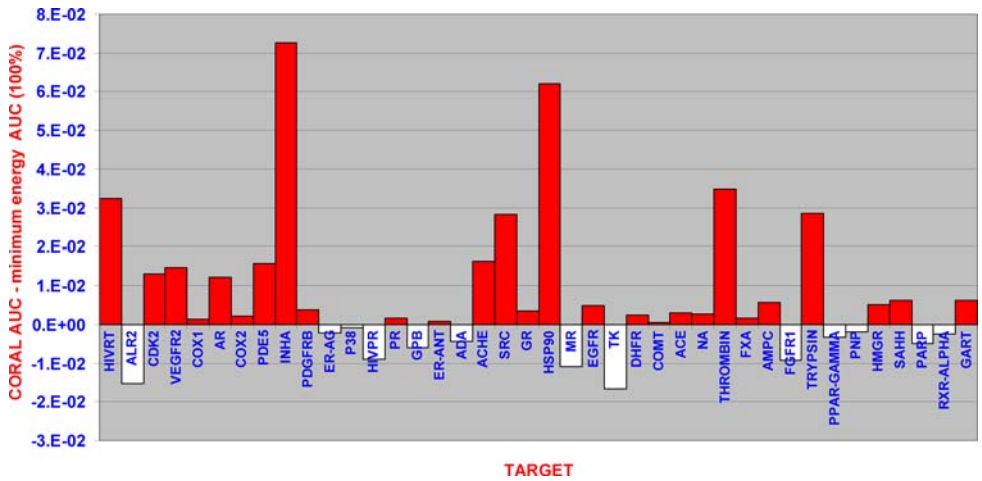


Fig. 5 The difference between the average CORAL AUC and the average minimum energy AUC for the 40 DUD Targets at 100% of the database retrieved. Averages are compiled across the ten individual screening experiments performed for each DUD target

CORAL, the magnitude of the percent change is typically much smaller. The target for which the minimum energy query most outperforms CORAL is ALR2 in which case the percent change is −2.78%.

The magnitudes of the percent changes should be reflected somewhat in the p values in column 9 of Table 3, i.e., for larger magnitude percent changes, statistical significance is more likely. In fact there is a statistically significant difference between the CORAL AUC and the minimum energy AUC for 9 targets, these are ALR2, VEGFR2, INHA, ACHE, SRC, HSP90, TK, Thrombin, and Trypsin. Of these 7 are cases in which CORAL outperforms minimum energy. Only two targets, ALR2 and TK, are for cases in which the minimum energy query outperforms CORAL.

There are 4 italicized rows in Table 3, these are for the targets COX2, MR, EGFR, and DHFR. These are targets that have p values >0.05 and ≤ 0.1. These cases are highlighted because there is a good chance that the addition of more screening experiments will render the p values ≤ 0.05 thereby making the differences between CORAL and minimum energy statistically significant. Note that in 3 out of 4 of these cases the CORAL conformation outperforms the minimum energy conformation.

Clearly from these examples it is seen that CORAL queries outperform minimum energy queries a majority of

the time. When the CORAL query outperforms the minimum energy query it can do so by a significant amount. On the other hand, when the minimum energy query outperforms CORAL, it does so by only a small amount. This is why more of the statistically significant differences are for cases where CORAL outperforms minimum energy queries. Even for the cases of near statistical significance, the majority of time it is when CORAL outperforms minimum energy. The general conclusion here is that CORAL queries on average perform equal to or better than the minimum energy queries.

Table 3 and Fig. 5 look similar at all values of X (percent database retrieved) for which the CORAL AUC values were statistically different from the minimum energy AUC values in Tables 1 and 2 (using all 400 screening experiments in the statistics). Taking a more detailed look at early enrichment, Table 4 and Fig. 6 show the results at $X = 4\%$ (the lowest percent of the database retrieved for which statistical significance was achieved). Again, the quantities quoted in Table 4 are computed using the ten individual screening experiments performed for each DUD target. These results are similar to the $X = 100\%$ results, in that CORAL queries outperform minimum energy queries. In fact CORAL queries outperform minimum energy queries for 29 out of the 40 targets considered. There are some differences, however. The first of these is that the targets for which CORAL performs best are different at $X = 4\%$ than at $X = 100\%$. For example at $X = 100\%$ and for the target Thrombin, CORAL queries outperform minimum energy queries by 3.87%. At $X = 4\%$, however, CORAL and minimum energy queries perform identically. Also, looking at column 8 of Table 4, it is seen that the percent changes are in general larger. For example, in the case of INHA, at 4% database retrieved the CORAL AUC is 21.80% larger than the minimum energy AUC. At 100% of the database retrieved, however, the CORAL AUC is 11.46% larger than the minimum energy AUC. In total, there are nine cases in Table 4 where the percent changes are greater than 5%. In 8 out of these 9 cases CORAL queries outperform minimum energy queries. In contrast, at $X = 100\%$ (Table 3) there are only three cases where the percent change is greater than 5%. In all three of these cases CORAL queries outperform minimum energy queries.

In Table 4 there is a statistically significant difference between the CORAL AUC and the minimum energy AUC for six targets; INHA, HIVPR, SRC, DHFR, COMT, and HMGR. Of these 5 are cases in which CORAL outperforms minimum energy. In only one case, HMGR, does the minimum energy query outperform CORAL.

There are five italicized rows in Table 4, these are for targets COX2, PDE5, ACHE, FGFR1, and Trypsin. These are targets that have p values > 0.05 and ≤ 0.1. These cases are highlighted because there is a good chance that the addition of more screening experiments will render the p values ≤ 0.05 thereby making the differences between CORAL and minimum energy statistically significant. Note that 4 out of 5 of these are cases for which CORAL outperforms minimum energy. It is clear that at $X = 4\%$ database retrieved the conclusion is the same as at $X = 100\%$, that is, CORAL queries on average perform equal to or better than the minimum energy queries.

This conclusion is seen another way by viewing the ROC curves for CORAL and minimum energy screens out to 4% inactives retrieved in Figs. 7, 8, 9, 10, 11, and 12. These ROC curves are compiled as averages over the ten screens done per DUD target. The particular DUD targets are chosen for which there was a statistically significant difference between the average CORAL and minimum energy AUC values. The targets chosen are INHA ($p < 0.01$), HIVPR ($p = 0.02$), SRC ($p = 0.05$), DHFR ($p < 0.01$), COMT ($p < 0.01$), and HMGR ($p = 0.02$).

Figures 7, 8, 9, 10, 11, and 12 each contain three ROC curves, one for the CORAL query, one for the minimum energy query, and one for the case of random compound selection. A look at the ROC curves in Figs. 7, 8, 9, 10, 11, and 12 shows that for the cases where CORAL outperforms minimum energy (INHA, HIVPR, SRC, DHFR, and COMT) that the CORAL curve is observed to be consistently above the minimum energy ROC curve. For the one case where the minimum energy query outperforms CORAL (HMGR), the ROC curves are nearly identical. Again these observations underscore the fact that on average CORAL queries are equal to or better than minimum energy queries.

The most detailed view of the early enrichment results is given in Fig. 13. Figure 13 shows AUC differences for all 400 individual screens (ten screens per DUD target $\times$ 40 DUD targets). In this case, no averaging has been done. The targets are ordered form left to right by decreasing diversity. The first ten AUC differences in Fig. 13 are associated with HIVRT (active ligand diversity = 0.52), the second 10 with ALR2 (active ligand diversity = 0.51) and so on. While there are cases where minimum energy queries outperform CORAL queries, the CORAL queries clearly outperform the minimum energy queries more frequently. In 93 of the screens shown in Fig. 13 CORAL queries outperformed minimum energy queries by 5% or more. In 36 of the screens, however, minimum energy queries outperformed CORAL queries by >5%.

Homing in on the details of the ROC curves for the individual screens we consider the CORAL and minimum energy queries associated with the targets INHA, SRC, and DHFR. These are targets in Table 4 for which the differences between CORAL and minimum energy AUC values are statistically significant. In addition for these targets

Table 4 Average ROC AUC values For CORAL and minimum energy queries at 4% database retrieved

Target	Diversity	CORAL		Minimum energy		<Difference>	% Change	*p*
		<AUC>	SD	<AUC>	SD			
HIVRT	0.52	0.003124	0.003	0.003146	0.003	−0.000022	−0.71	0.72
ALR2	0.51	0.009737	0.005	0.009691	0.005	0.000046	0.48	0.84
CDK2	0.49	0.007093	0.002	0.006931	0.002	0.000163	2.35	0.45
VEGFR2	0.49	0.005734	0.002	0.005573	0.003	0.000161	2.89	0.36
COX1	0.49	0.005004	0.002	0.004888	0.0008	0.000116	2.38	0.88
AR	0.46	0.010554	0.002	0.010685	0.002	−0.000130	−1.22	0.74
COX2	*0.45*	*0.018623*	*0.002*	*0.018509*	*0.002*	*0.000113*	*0.61*	*0.07*
PDE5	*0.43*	*0.007529*	*0.004*	*0.006896*	*0.004*	*0.000633*	*9.18*	*0.08*
INHA	**0.43**	**0.011649**	**0.005**	**0.009564**	**0.006**	**0.002085**	**21.80**	**<0.01**
PDGFRB	0.43	0.010871	0.002	0.010789	0.001	0.000082	0.76	0.18
ER-AG	0.42	0.012980	0.004	0.012751	0.004	0.000229	1.80	0.18
P38	0.4	0.016040	0.0007	0.015939	0.0004	0.000101	0.63	0.45
HIVPR	**0.39**	**0.015648**	**0.0009**	**0.015125**	**0.001**	**0.000523**	**3.46**	**0.02**
PR	0.39	0.018290	0.003	0.018080	0.003	0.000210	1.16	0.14
GPB	0.37	0.012800	0.003	0.013057	0.003	−0.000257	−1.97	0.49
ER-ANT	0.37	0.010602	0.001	0.009854	0.001	0.000749	7.60	0.22
ADA	0.37	0.010878	0.001	0.010210	0.002	0.000668	6.54	0.25
ACHE	*0.37*	*0.012978*	*0.002*	*0.011079*	*0.003*	*0.001899*	*17.14*	*0.10*
SRC	**0.36**	**0.016810**	**0.001**	**0.015983**	**0.003**	**0.000828**	**5.178**	**0.05**
GR	0.36	0.017576	0.0007	0.017576	0.001	0.000000	0.000	0.34
HSP90	0.34	0.014095	0.002	0.014057	0.002	0.000037	0.27	0.19
MR	0.33	0.005101	0.005	0.005258	0.005	−0.000156	−2.97	0.17
EGFR	0.33	0.018424	0.006	0.017762	0.006	0.000663	3.73	0.39
TK	0.33	0.005046	0.005	0.005042	0.005	0.000004	0.08	0.87
DHFR	**0.33**	**0.018932**	**0.005**	**0.017349**	**0.004**	**0.001584**	**9.13**	**<0.01**
COMT	**0.33**	**0.010831**	**0.007**	**0.010134**	**0.007**	**0.000697**	**6.88**	**<0.01**
ACE	0.32	0.007056	0.003	0.007112	0.003	−0.000056	−0.78	0.83
NA	0.31	0.015975	0.003	0.015518	0.003	0.000456	2.94	0.28
Thrombin	0.31	0.016908	0.001	0.016907	0.001	0.000001	<0.01	0.78
FXA	0.3	0.018151	0.0004	0.017976	0.0006	0.000175	0.98	0.52
AMPC	0.29	0.019380	0.002	0.019551	0.002	−0.000171	−0.87	0.65
FGFR1	*0.28*	*0.017297*	*0.0009*	*0.017448*	*0.001*	*−0.000150*	*−0.86*	*0.06*
Trypsin	*0.26*	*0.017172*	*0.001*	*0.016869*	*0.001*	*0.000303*	*1.80*	*0.07*
PPAR-GAMMA	0.26	0.018369	0.002	0.018143	0.001	0.000226	1.248	0.38
PNP	0.24	0.016958	0.002	0.017098	0.002	−0.000140	−0.818	0.39
HMGR	**0.23**	**0.018288**	**0.003**	**0.018625**	**0.003**	**−0.000337**	**−1.811**	**0.02**
SAHH	0.22	0.011834	0.003	0.011366	0.004	0.000467	4.11	0.72
PARP	0.22	0.019329	0.004	0.019408	0.005	−0.000079	−0.41	0.94
RXR-ALPHA	0.18	0.015946	0.004	0.015830	0.005	0.000116	0.73	0.94
GART	0.14	0.005872	0.005	0.006744	0.003	−0.000871	−12.92	0.80

Results are given for individual DUD targets. Averages for each query type are compiled across ten ROC curves generated for each DUD target. Bold and italic entries (11 altogether) are statistically significant at the 95 and 90% confidence levels, respectively. The number of decimal places shown for the AUC values and differences equals that of the smallest difference in the table

CORAL found two or more conformations of the same molecule that were different than the minimum energy structures. A look now at individual ROC curves for these cases will show that the screening results are indeed sensitive to the geometries utilized. Table 5 shows the identities of the CORAL queries considered. Figures 14, 15,

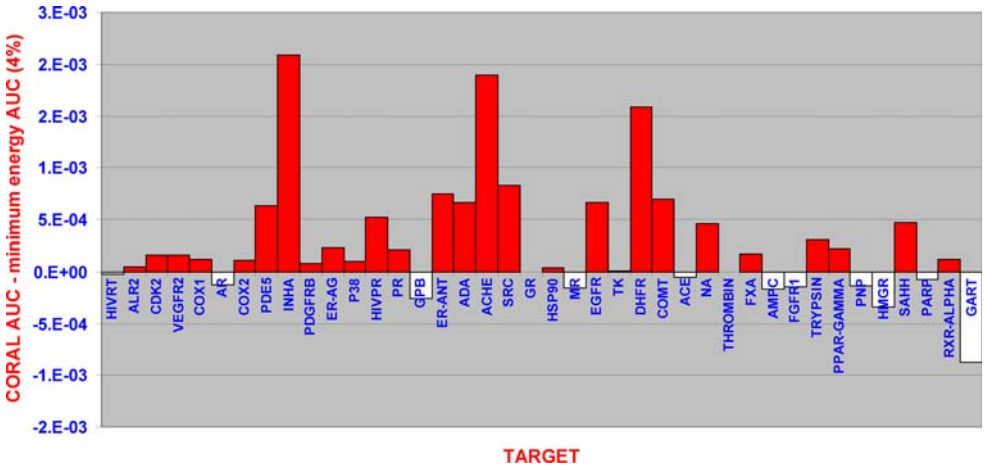


Fig. 6 The difference between the average CORAL AUC and the minimum energy AUC for the 40 DUD targets at 4% of the database retrieved. Averages are compiled across the ten individual screening experiments performed for each DUD target

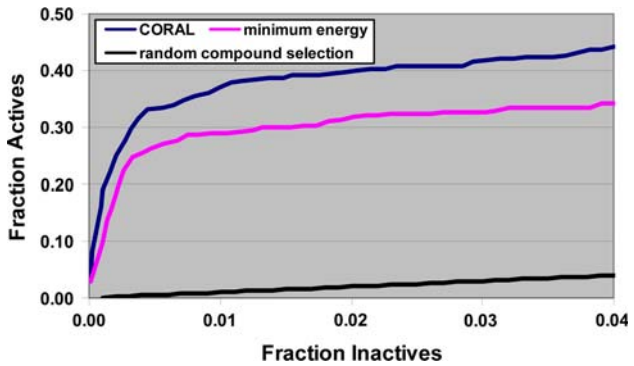


Fig. 7 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. Target = INHA. These ROC curves are compiled as averages over the ten screens done per DUD target

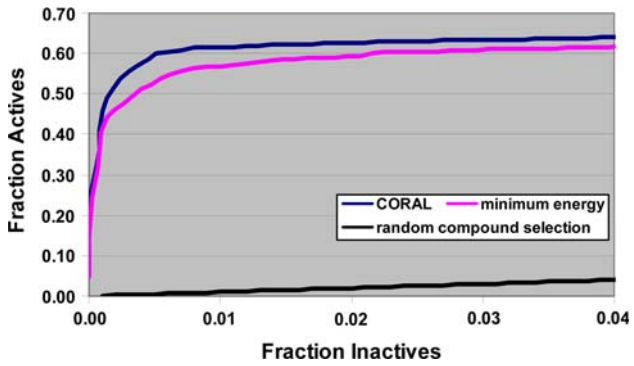


Fig. 9 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. Target = SRC. These ROC curves are compiled as averages over the ten screens done per DUD target

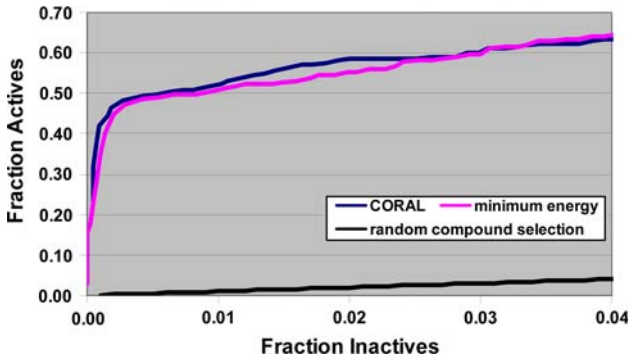


Fig. 8 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. Target = HIVPR. These ROC curves are compiled as averages over the ten screens done per DUD target

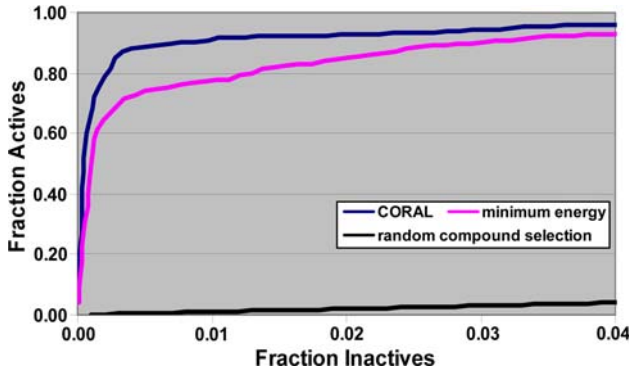


Fig. 10 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. TARGET = DHFR. These ROC curves are compiled as averages over the ten screens done per DUD target

and 16 show the ROC curves produced by the query conformations in Table 5 along with ROC curves of their associated minimum energy structures.

In Figs. 14, 15, and 16 the ROC curve is given for the CORAL queries and the associated minimum energy queries, followed by a 2D rendering of the query molecule, finally a table showing the conformation numbers of the query molecule is shown. The data in the tables shows the OMEGA energy relative to the minimum energy structure, the RMS relative to the minimum energy structure and the AUC difference at 4% inactives retrieved (CORAL AUC − minimum energy AUC).

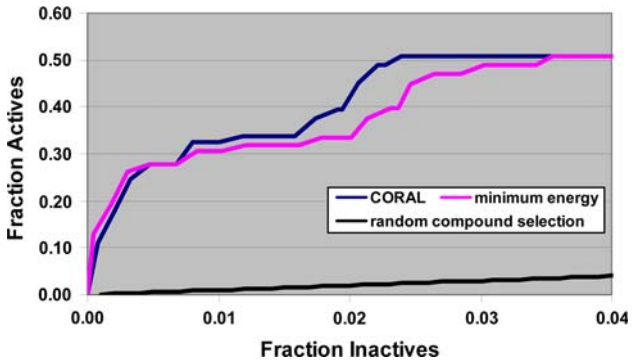


Fig. 11 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. Target = COMT. These ROC curves are compiled as averages over the ten screens done per DUD target

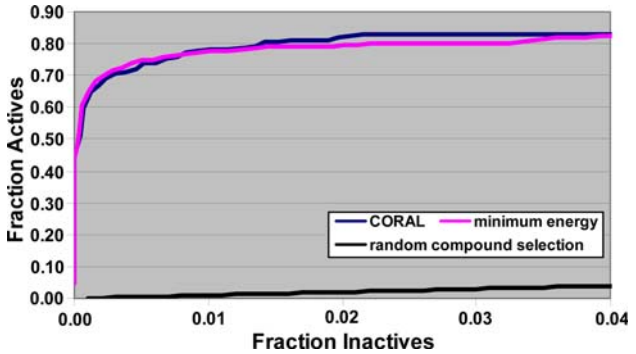


Fig. 12 ROC CURVES out to 4% of the inactives retrieved for CORAL and minimum energy queries. Target = HMGR. These ROC curves are compiled as averages over the ten screens done per DUD target

The first thing to notice concerning Figs. 14, 15, and 16 is that in all cases the CORAL ROC curves are consistently above the minimum energy ROC curves for values of fraction inactives out to 4%. Looking at the last column in the Tables in Figs. 14, 15, and 16 it is seen that all AUC differences are positive, i.e., in favor of the CORAL queries. Clearly, these observations support the conclusion that the CORAL geometries can give better virtual screening performance than the minimum energy conformations. Analysis of the RMS values in column 2 of the tables in Figs. 14, 15, and 16 show that the CORAL geometries are often very different than the associated minimum energy geometries. The general trend is that as the RMS deviation between the CORAL query and the minimum energy query increases the AUC difference gets more positive.

These observations reflect the fact that sometimes the choice of conformation does matter in ROCS virtual screening. Granted the ligand sets chosen for individual ROC analysis here, Figs. 14, 15, and 16, come from a few targets where there is a statistically significant difference between CORAL and minimum energy AUC values. However, the histogram plots of Figs. 5, 6, and 13 which show more positive than negative differences and the *T* tests (Tables 1, 2), which verify a significant difference between CORAL and minimum energy AUC distributions at a general level tell us that the conclusions drawn from Figs. 14, 15, and 16 are more generally valid and therefore apply to many of the other DUD ligand sets as well. All that needs to be done is to perform more screening computations in order to resolve the differences between the two query types for these other DUD ligand sets.

Discussion

The current literature suggests that the choice of query molecule is important in virtual screening performance of ROCS [17], but that the choice of conformation does not matter [16, 17]. After doing many hundreds of screens, however, we have shown that different geometries of the same molecule can yield very different results. The problem in our minds is that until now there has been no systematic way to select the best representative geometries and then to show that they do indeed give different results over a broad range of systems.

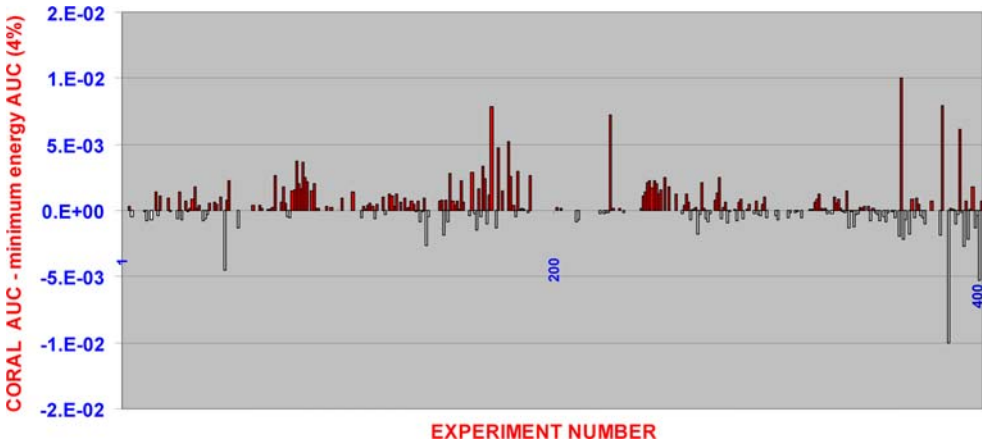


Fig. 13 The differences between the CORAL AUC and minimum energy AUC for all 400 individual screens (ten screens per DUD target × 40 DUD targets). In this case no averaging has been done. Results are given at 4% of the database retrieved

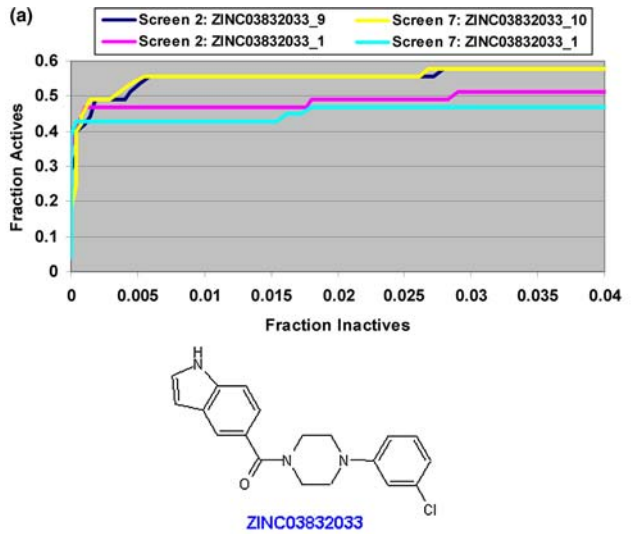
Table 5 CORAL query structures for selected DUD targets

INHA	SRC	DHFR
Screen 7: ZINC03832033_10	Screen 7: ZINC03832360_4	Screen 3: ZINC03814876_34
Screen 2: ZINC03832033_9	Screen 1: ZINC03832360_3	Screen 2: ZINC03814876_33
Screen 5: ZINC03833914_41		
Screen 3: ZINC03833914_12		

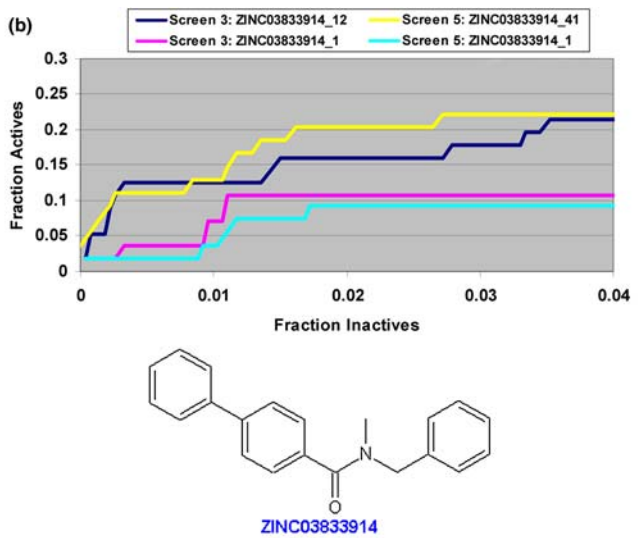
For us the solution to this problem is related to two recent lines of work. One regards proper experimental design for methods evaluation [22, 26, 32]. The other regards choice of query structure using available chemical information [17].

There are several important factors in the area of experimental design and methods evaluation. One is the choice of decoys and actives. We do not want the decoys to have significantly different physical properties than the actives because then it would be too easy to distinguish actives from decoys. This tends to equalize the performance various screening methods. As a result resolving the differences between methods becomes more difficult. The DUD dataset that is used in this work is good in this regard because the decoys are property matched to actives [21] thereby making it more difficult to distinguish actives from decoys. This will tend to augment the performance differences between various screening methods and make it easier to resolve the differences between methods.

Another important factor is quantification of virtual screening results. The goal is to identify which of two methods is better, either ROCS screening with a CORAL geometry or ROCS screening with a minimum energy geometry. Methods are most often quantified by computation of enrichment curves. We, on the other hand use ROC curves. ROC curves have multiple advantages over traditional enrichment curves, the most important of which is that the ROC AUC is independent of fraction actives present in a screening database. In the case of enrichment curves, on the other hand, AUC values are dependent on fraction of actives in a data set. As a result for enrichment curves, experiments performed on different data sets have different dynamical ranges (size of enrichment) and can not be compared. This is not the case for ROC curves. Using

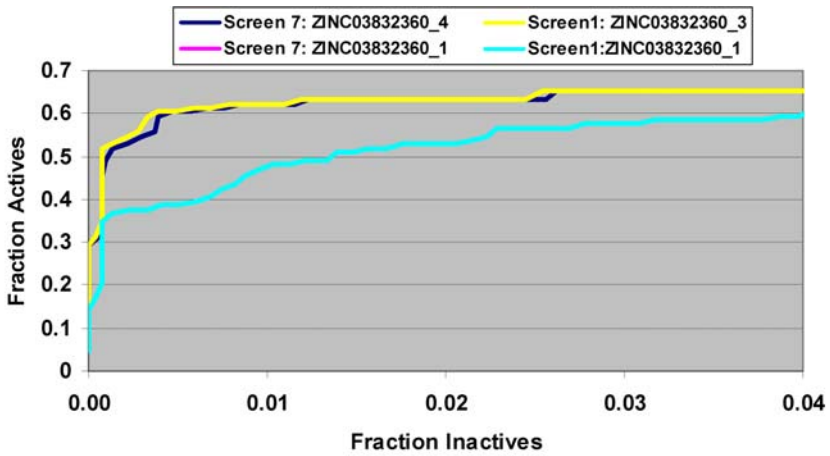


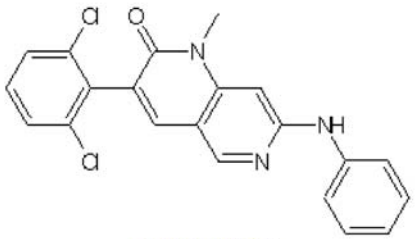
Conformation	OMEGA ENERGY	RMS	ΔAUC at 4%
ZINC03832003_1	0.00	0.00	0.0000
ZINC03832003_9	0.88	0.90	0.0027
ZINC03832003_10	3.48	2.05	0.0048



Conformation	OMEGA ENERGY	RMS	ΔAUC at 4%
ZINC03833914_1	0.00	0.00	0.0000
ZINC03833914_12	3.11	0.95	0.0027
ZINC03833914_41	9.84	1.64	0.0044

Fig. 14 **a** ROC curves for INHA up to 4% of the inactives retrieved. Two CORAL ROC curves are shown (conformations 9 and 10) along with ROC curves for the associated minimum energy structures (conformation 1). Active molecule = ZINC03832033. The active ligand diversity for INHA is 0.43. Energies and RMS values are given relative to the minimum energy structure. **b** ROC curves for INHA up to 4% of the inactives retrieved. Two CORAL ROC curves are shown (conformations 12 and 41) along with ROC curves for the associated minimum energy structures (conformation 1). Active molecule = ZINC03833914. The active ligand diversity for INHA is 0.43. Energies and RMS values are given relative to the minimum energy structure





ZINC03832360

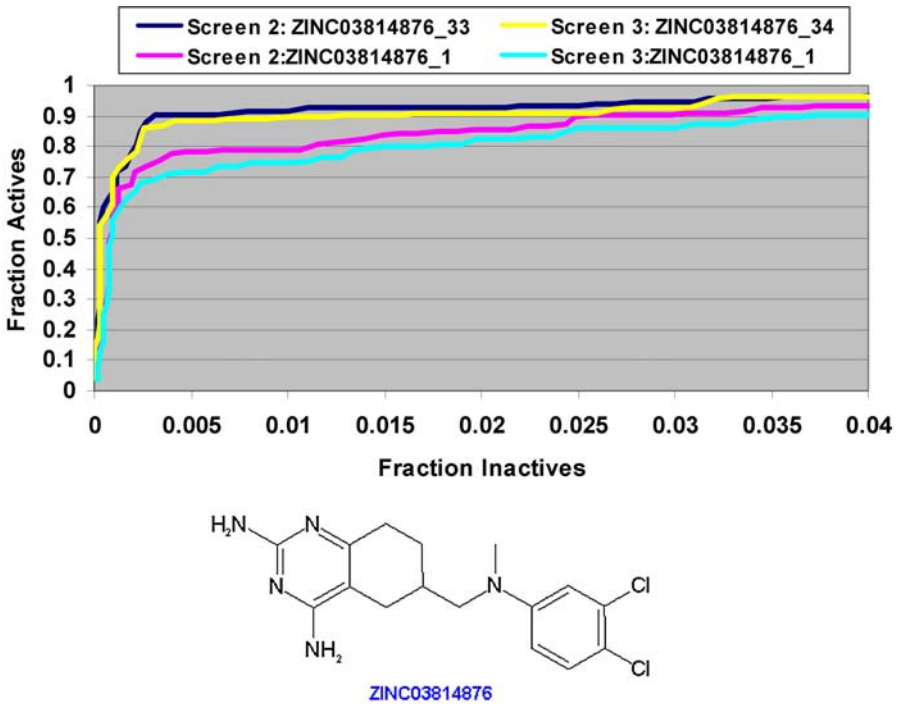
Fig. 15 ROC curves for SRC up to 4% of the inactives retrieved. Two CORAL ROC curves are shown (conformations 4 and 3) along with ROC curves for the associated minimum energy structures (conformation 1). Active molecule = ZINC03832360. The active ligand diversity for SRC is 0.36. Energies and RMS values are given relative to the minimum energy structure

Conformation	OMEGA ENERGY	RMS	ΔAUC at 4%
ZINC03832360_1	0.00	0.00	0.0000
ZINC03832360_3	3.48	1.58	0.0043
ZINC03832360_4	3.59	2.07	0.0045

ROC curves we can compare any two methods on any data set available.

This leads into the next point which is the ability to resolve the difference between two methods. This needs many computations. In our case using bootstrapping we perform ten screens per DUD target for each query type (CORAL and minimum energy). Since there are 40 DUD targets, this makes 400 total screening experiments for each query type. The large number of computations enhances resolution power. With the 400 computations at our disposal we can show significant differences between the CORAL AUC distribution and minimum energy AUC distribution down to 4% of the database screened. The results at 3% database screened are almost statistically significant. To obtain statistical significance at 2 or 1% database screened will need significantly more computations. One way to achieve this is to use a higher level of bootstrapping. To obtain statistical significance at the 3% level, not much has to be done. In this case the T value of 1.89 (Table 1) is just a little less than T_{critical} of 1.96. T_{critical} is the cutoff between retaining or rejecting the null hypothesis (that both query types yield identical AUC values). Since T increases roughly as the $\sqrt{N}$ (N being the sample size, currently 400) and looking at the size of T in relation to T_{critical} we determine that bootstrapping at 11 times per target (instead of the current 10) will lock in statistical significance at 3% database screened. Using similar arguments, locking in statistical significance at 1 or 2% database screened would require bootstrapping somewhere in the range of 73–260 times per target, respectively. These experiments though computationally intense can be done with the current methodology and would require anywhere from 2,920 to 10,400 screens, respectively.

Putting the experimental design aspects aside for a moment, the other line of work important for this discussion regards choice of query structures using available chemical information [17]. In their efforts to improve the performance of ROCS queries Kirchmair et al. [17] utilize information from multiple active molecules. To do this they compute the center molecule for each of the active DUD ligand sets. The center molecule is computed using standard clustering methods and ECFP4 [33, 34] fingerprints. Once the center molecule is determined, the rest of the actives are seeded into the decoy sets. The lowest energy conformation of each center molecule is generated using OMEGA. ROCS screening using the minimum energy structure of each center molecule was performed for all 40 DUD sets. Kirchmair states that the average AUC of 0.82 is obtained,



ZINC03814876

Fig. 16 ROC curves for DHFR up to 4% of the inactives retrieved. Two CORAL ROC curves are shown (conformations 33 and 34) along with ROC curves for the associated minimum energy structures (conformation 1). Active molecule = ZINC03814876. The active ligand diversity for DHFR is 0.33. Energies and RMS values are given relative to the minimum energy structure

Conformation	OMEGA ENERGY	RMS	ΔAUC at 4%
ZINC0381487_1	0.00	0.00	0.0000
ZINC0381487_33	17.54	2.72	0.0035
ZINC0381487_34	17.57	2.90	0.0039

which is 14% higher than using the ligand present in the respective x-ray structures.

Kirchmair emphasizes the fact that the center molecule of an active ligand set is better to use for virtual screening. For us the question remaining is the dependency of the results on query conformations of the center molecule. *Since the center molecule produces augmented AUC values it may very well be the case that there are augmented differences in AUC values for various conformations of the center molecule.* Our efforts are similar to Kirchmair's, but the center molecule and its optimal conformation (CORAL geometry) are computed simultaneously in 3D by finding the conformation that has the largest average combo score with the rest of the chosen active molecules (See Eqs. 1–3). After finding the optimal center molecule conformation, we also determine the minimum energy conformation of the center molecule. Then using the experimental design discussed above, an evaluation is made as to which is better, the center molecule optimal conformation (CORAL), or the center molecule minimum energy conformation. After 800 screening experiments a statistically significant difference was found between the utility of these two conformations of the center molecule, and we conclude that the center molecule optimized conformation (CORAL) is in general superior to the center molecule minimum energy conformation.

Extensions to the current work are can be defined by considering Fig. 4. In this figure it is seen that AUC values at 100% database retrieved (CORAL or minimum energy) can decreases as the diversity index of the DUD active ligand sets increases. It is obvious that a single, molecule geometry represented by either the CORAL or minimum energy query is less able to draw enriched hit sets if the active compounds to be found are diverse. This is the case because a diverse set of active compounds is represented in 3D by multiple active molecule conformational classes. At best (assuming that the query conformation shares similarity with one of the conformational classes) a single molecule geometry query will preferentially draw hits from one of these classes and not the others. This tends to flatten the ROC curve and reduce the AUC.

A good strategy to ameliorate the problems posed by Fig. 4 is to run a clustering computation on the conformationally expanded set of the known active molecules for each DUD target, then compute the 3D centroids of each cluster as described previously. The 3D centroids (one for each cluster) should then be used to run multiple ROCS searches for each DUD target. Hit compounds would be

ranked using their highest combo score with any of the centroid queries. In this way active conformations from all conformational classes get pushed forward in the hit list. This will lift the ROC curve and increase the AUC. This strategy is similar to one presented by Kirchmare [17] but different in that the center molecule and its optimum conformation are determined simultaneously by clustering at the conformational level.

Conclusions

Direct comparisons of CORAL AUC values and minimum energy AUC values from early (4%) to late (20%) screening stages show that on average CORAL queries outperform minimum energy queries in ROCS-based virtual screening. Therefore, different conformations of the same molecule can give different ROCS screening results. In some cases, choice of query conformation can significantly enhance screening results, e.g., INHA, SRC, DHFR.

Because the DUD datasets are smaller and more densely packed with actives than the typical virtual screening collection, it is difficult to directly assess the practical significance of the differences we found between the CORAL and minimum energy queries. In particular, it is difficult to choose a particular cutoff as the most relevant to compare to prospective virtual screening work. At the 4% cutoff, we find that in most cases the number of actives retrieved would be comparable between the CORAL and minimum energy queries. For 9 of the DUD targets, however, there is a greater than 5% discrepancy between the number of actives retrieved by the two queries. In 8 of these 9 cases, the CORAL query outperforms the minimum energy query. Thus, there is little chance that a priori CORAL analysis will significantly hurt the virtual screen performance whereas there is nearly a 1 in 4 chance that a CORAL analysis will significantly increase the number of actives retrieved. Given this gain and its simplicity, we suggest that CORAL be considered as a first step in virtual screening workflows using ROCS.

Acknowledgments The authors would like to thank Will Somers and Tarek Mansour of Wyeth Chemical Sciences for their support, Dave Diller for manuscript suggestions, Ramaswamy Nilikantan for help with the diversity analysis and Youping Huang for help in performing the statistical analysis.

References

1. Rai BK, Tawa GJ, Katz AH, Humblet C (2009) Modeling G protein-coupled receptors for structure-based drug discovery using low-frequency normal modes for refinement of homology models: application to H3 antagonist. Proteins (accepted for publication)
2. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: A G. protein-coupled receptor. Science 289:739–745
3. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human B2-adrenergic G protein-coupled receptor. Science 318:1258–1265
4. Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman AP, Stevens RC (2008) The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. Science 322:1211–1217
5. Kim D, Xu D, Guo JT, Ellrott K, Xu Y (2003) PROSPECT II: protein structure prediction program for genome-scale applications. Protein Eng 16:641–650
6. Petrey D, Xiang Z, Tang CL, Xie L, Gimpelev M et al (2003) Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. Proteins 53(6):430–435
7. Simons KT, Kooperberg C, Huang E, Baker D (1997) Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. J Mol Biol 268:209–225
8. Tresadern G, Bemporad D, Howe TA (2009) Comparison of ligand based virtual screening methods and application to corticotrophin releasing factor 1 receptor. J Mol Graph Model 27:860–870
9. ROCS 2.3.1, OpenEye Scientific Software, Santa Fe, NM, 2007. http://www.eyesopen.com
10. Grant JA, Gallard MA, Pickup BG (1996) A fast method of molecular shape comparison: a simple application of a Gaussian description of molecular shape. J Comput Chem 17:1653–1666
11. Nicholls A, Grant JA (2005) Molecular shape and electrostatics in the encoding of relevant chemical information. J Comput-Aided Mol Des 19:661–686
12. Freitas RF, Oprea TI, Montanari CA (2008) Two-dimensional QSAR and similarity studies on cruzain inhibitors aimed at improving selectivity over cathepsin L. Bioorg Med Chem 16: 838–853
13. Bostrom J, Greenwood JR, Gottfries J (2003) Assessing the performance of OMEGA with respect to retrieving bioactive conformations. J Mol Graph Model 21:449–462
14. Bostrom J (2001) Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. J Comput Aided Mol Des 15:1137–1152
15. Diller DD, Merz KM Jr (2002) Can we separate active from inactive conformations? J Comput Aided Mol Des 16:105–112
16. Hawkins PCD, Skillman GA, Nicholls A (2007) Comparison of shape-matching and docking as virtual screening tools. J Med Chem 50:74–82
17. Kirchmair J, Distinto S, Markt P, Schuster D, Spitzer GM, Liedl KR, Wolber G (2009) How to optimize shape-based virtual screening: choosing the right query and including chemical information. J Chem Inf Model 49:678–692
18. Perola E, Charifson PS (2004) Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. J Med Chem 45:2499–2510
19. Putta S, Landrum GA, Penzotti JE (2005) Conformation mining: an algorithm for finding biologically relevant conformations. J Med Chem 48:3313–3318
20. Rush TA (2005) Shaped-based 3-D scaffold hopping method and its application to a bacterial protein–protein interaction. J Med Chem 48:1489–1495

21. Huang N, Shoichet B, Irwin J (2006) Benchmarking sets for molecular docking. J Med Chem 49:6789–6801
22. Triballeau N, Acher F, Brabet I, Pin J-P, Bertrand H-O (2005) Virtual screening workflow development guided by the "Receiver Operating Characteristic" curve approach. Applications to high-throughput docking on metabotropic glutamate receptor subtype 4. J Med Chem 48:2534–2547
23. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36
24. OMEGA 2.2.1, OpenEye Scientific Software, Santa Fe, NM, 2007. http://www.eyesopen.com
25. Bostrom J (2002) Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. J Comput Aided Mol Des 15:1137
26. Hawkins PCD, Warren GL, Skillman AG, Nicholls A (2008) How to do an evaluation: pitfalls and traps. J Comput Aided Mol Des 22:179–190
27. Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W.H. Freeman, New York
28. Turner DB, Tyrell SM, Willett P (1997) Rapid quantification of molecular diversity for selective database acquisition. J Chem Inf Comput Sci 37:18–22
29. Patterson DE, Cramer RD, Ferguson AM, Clark RD, Weinberger LE (1996) Neighborhood behavior: a useful concept for validation of ''molecular diversity'' descriptors. J Med Chem 39:3049–3059
30. Bostrom J, Hogner A, Schmitt S (2006) Do structurally similar ligands bind in a similar fashion? J Med Chem 49:6716–6725
31. OEChem-C++ theory manual, OEMCSSEARCH. OpenEye Scientific Software: Santa Fe, NM, 2006. http://www.eyesopen.com
32. Nicholls A (2008) What do we know and when do we know it? J Comput Aided Mol Des 22:239–255
33. Hassan M, Brown RD, Varna-O'Brien S, Rogers D (2006) Cheminformatics analysis and learning in a data pipelining environment. Mol Divers 10:283–299
34. Scitegic Inc, Pipeline Pilot Version 7.5.2.300, 2009. http://www.scitegic.com