Chemical Substitutions That Introduce Activity Cliffs Across Different Compound Classes and Biological Targets

Anne Mai Wassermann and Jürgen Bajorath*

Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Dahlmannstrasse 2, D-53113 Bonn, Germany

Received May 10, 2010

Applying the concept of matched molecular pairs, we have systematically analyzed the ability of defined chemical changes to introduce activity cliffs. Public domain compound data were systematically screened for matched molecular pairs that were then organized according to chemical transformations they represent and associated potency changes. From vast available chemical transformation space, including both R-group and core substructure changes, ~250 nonredundant substitutions were identified that displayed a general tendency to form activity cliffs. These substitutions introduced activity cliffs in the structural context of diverse scaffolds and in compounds active against many different targets. Activity cliff-forming transformations were often rather simple, including replacements of small functional groups. Moreover, in many instances, chemically very similar transformations were identified that had a much lower propensity to form activity cliffs or no detectable cliff potential. Thus, clear preferences emerged for specific transformations. A compendium of substitutions with general activity cliff-forming potential is provided to aid in compound optimization efforts.

INTRODUCTION

Traditionally, the design of analogs in hit-to-lead and lead optimization campaigns has largely depended on medicinal chemistry expert knowledge, experience, and intuition.[1] Without doubt, this subjective and expertise-driven approach has to this date been largely responsible for the success of medicinal chemistry programs, despite the increasing use of high-throughput technologies in drug discovery settings. However, attempts have also been made to more systematically address the process of analog design in order to aid medicinal chemists in the decision which compounds to synthesize next in the course of a compound optimization effort. By and large, these attempts have focused, and continue to focus, on analyzing the wealth of available analog and structure−activity relationship (SAR) data for diverse targets in order to identify, and ultimately predict, favorable chemical modifications. To these ends, rather different strategies have been pursued. In an early systematic study, Sheridan[2] screened the MDL Drug Data Report (MDDR)[3] for single-point modifications of compounds belonging to a specific activity class that retained biological activity. Going a step further, Haubertin and Bruneau compiled chemical replacements from a corporate compound deck and analyzed their effects on various compound properties.[4] The analysis covered substitutions involving a predefined set of 9000 functional groups. Furthermore, Raymond et al. applied a maximum common substructure-based algorithm to identify an exhaustive set of chemical changes within a collection of 2.7 million compounds and proposed a quantitative measure to assess the relevance of these modifications for optimization efforts.[5] This study was based on pairwise comparisons of compounds in order to identify substitutions. Moreover, Hajduk and Sauer investigated the effects of common chemical substitutions on ligand potency.[6] In this study, compound potency distributions resulting from 120 different chemical changes were compared, and substitutions identified that consistently shifted the distributions to higher or lower potency ranges. Despite their methodological differences, these studies had in common that relationships between small chemical modifications and compound properties were analyzed.

A useful concept to generalize chemical modifications and treat them in a consistent manner is provided by "matched molecular pairs" (MMPs).[7] An MMP is defined as a pair of compounds that differ only at a single site (e.g., a specific R-group or ring system) and are distinguished by a defined substructure. Hence, a characteristic of an MMP is that the compounds forming the pair are related to each other by a well-defined "transformation" (e.g., addition of an R-group or replacement of a ring). Recently, Hussain and Rea have introduced an efficient algorithm to systematically extract MMPs from compound data sets,[8] thus enabling a large-scale analysis of MMP distributions.

We have been interested in exploring the potential of defined chemical changes to introduce "activity cliffs"[9] in compound data sets. Activity cliffs are formed by structurally similar compounds having large differences in potency and are responsible for SAR discontinuity.[9,10] In particular, we wished to determine whether defined chemical transformations exist that display a general tendency to introduce activity cliffs across different compound classes and targets. Asking this question was at least in part inspired by our previous findings that a significant number of molecular scaffolds exist in public domain compound databases representing compounds that frequently form activity cliffs

* Corresponding author. Telephone: +49-228-2699-306. Fax: +49-228-2699-341. E-mail: bajorath@bit.uni-bonn.de.

10.1021/ci1001845
Published on Web 07/07/2010

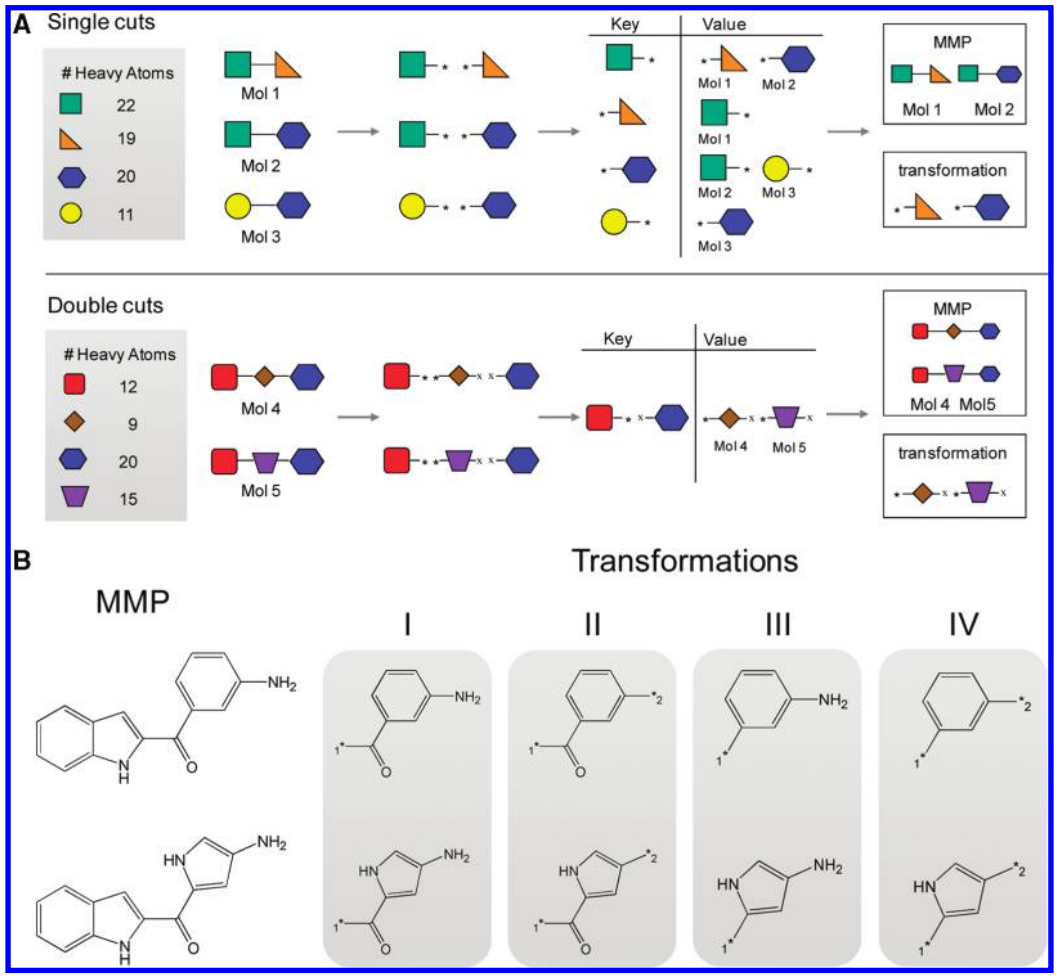

Figure 1. MMP generation scheme. (A) The steps leading to the identification of MMPs are illustrated for single and double cuts. Although molecules 2 and 3 in the single cut example share a common key fragment, the pair is not accepted as a MMP because the heavy atom counts of the two value fragments differ by more than eight atoms. (B) For the MMP shown on the left, multiple value fragment pairs of different sizes are obtained that represent alternative structural transformations, designated as transformations I−IV.

against multiple targets.[11] Hence, with this study, we intended to change the focus of activity cliff analysis from core structures to chemical transformations. Therefore, we have adapted the Hussain and Rea algorithm and systematically extracted MMPs from public domain compound sources. By grouping compound pairs according to their structural transformations and associating potency differences, we have identified structural changes that frequently formed activity cliffs. Filters were implemented to ensure that the analysis was not biased by ligand sets of individual targets or the presence of only very few compound pairs. On the basis of this analysis, we have identified a number of chemical transformations that introduced activity cliffs in different compound classes active against diverse targets. However, in many instances, similar structural transformations have also been identified that rarely formed activity cliffs.

MATERIALS AND METHODS

Compound Sets. We analyzed the two major public domain repositories that collect compound optimization and SAR data, ChEMBL[12] (CDB) and BindingDB[13,14] (BDB). Pubchem bioassays[15] were not suitable as a compound source for our analysis because they contain screening but not compound optimization data. From both CDB and BDB, compounds with potency measurements against human targets were extracted. Whenever available, K_i values were selected (otherwise, IC_{50} values). For compounds with multiple potency values reported against the same target, the arithmetic mean was calculated to yield the final potency. From CDB, only measurements were selected having the highest target confidence level (CDB target confidence score 9) for direct interactions (target relationship type "D").[12] Very large molecules with more than 45 nonring single bonds were not considered for fragmentation. Ligand sets were assembled for targets for which at least five active compounds were available meeting our criteria, leading to the selection of a total of 33 497 CDB and 19 288 BDB compounds organized into 523 and 328 sets, respectively. Set sizes ranged from 5 to 1528 ligands for CDB and from 5 to 1503 ligands for BDB.

All ligand structures were standardized prior to fragmentation using the "washing" function of Molecular Operating Environment.[16] The standardization process also included the protonation of strong bases and deprotonation of strong acids.

Matched Molecular Pair Analysis. In order to identify MMPs, we have implemented and modified the Hussain and Rea algorithm.[8] The first step of the approach involves fragmentation of all test compounds. This is accomplished by marking all nonring single bonds between two nonhy-

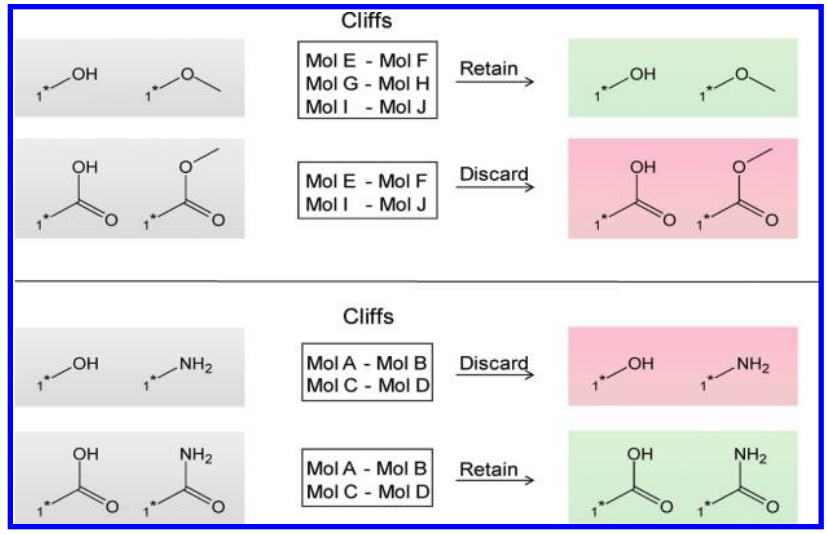

Figure 2. Determining cliff-forming transformations. In order to unambiguously define frequent cliff-forming transformations, rules are applied that relate the size of exchanged fragments and their number of occurrences in cliff-forming molecule pairs to each other.

drogen atoms in a molecule, followed by systematic deletion of these bonds ("single cut") and their two- and three-bond combinations, called "double" and "triple cuts", respectively. A single cut results in two fragments F1 and F2 that are added to an index list. Two "key-value" pairs are built: fragment F1 is added as a "key" to the index with F2 as the corresponding "value" and vice versa. Double cuts result in a core and two terminal fragments. In this case, the core is considered the value, and the two terminal fragments together constitute the key. Of all possible triple cuts, only those are considered that result in a single core and three terminal fragments (i.e., triple cuts that result in two cores and two terminal fragments are not considered). In analogy to double cuts, the core is stored as the value with the three terminal fragments together as the key. For double and triple cuts, connectivity information of the fragments is retained. Furthermore, for all generated key-value pairs, source compound information is stored. Because all compounds sharing a particular key contain the corresponding fragment, the pairwise combination of these compounds yields MMPs, and the two value fragments define the structural transformation for each MMP. Stereochemical criteria are not considered.

In our implementation, a combination of two compounds was only considered an MMP if the heavy (nonhydrogen) atom counts of their distinguishing fragments differed by a maximum of eight atoms. This criterion was applied to ensure that compounds forming MMPs did not have large differences in size. By utilizing single and multiple cuts, transformations involving both R-groups and core structures can be assessed. A schematic outline of the algorithm is provided in Figure 1A. It should be noted that the algorithm often yields multiple, differently sized fragments that define the transformation of an MMP, as shown in Figure 1B.

Because transformations were associated with potency changes, MMPs were separately identified for compounds belonging to each ligand set. Our implementation also differed from the original method in that only those transformations were considered that occurred in at least two compound pairs in which the value fragment did not contain more heavy atoms than the combined key fragments (a criterion applied in analogy to rules for the derivation of maximum common substructures).[5]

For each accepted transformation, MMPs were grouped by ligand sets (targets), and logarithmic potency differences between compounds in each pair were recorded. If an MMP was present in more than one ligand set, i.e., if it was annotated with potency values for more than one target, potency differences for all targets were calculated. MMP analysis of active compounds does not account for mode-of-action differences.

All calculations were carried out with in-house written Perl or Scientific Vector Language[16] (SVL) scripts.

MMP sets identified for CDB and BDB and used for our analysis can be obtained via the following URL (downloadssection): http://www.lifescienceinformatics.uni-bonn.de.

Frequent Cliff-Forming Transformations. An MMP was considered to form an activity cliff if the potency of its compounds differed by at least two orders of magnitude. In order to identify transformations that most frequently introduced activity cliffs, the following filter criteria were applied:

(i) At least 1/15 (~7%) of the pK_i values recorded for a transformation represent activity cliffs (this value was empirically chosen to associate cliffs with multiple measurements).

(ii) The transformation introduces activity cliffs for at least four different targets.

(iii) The activity cliffs for the transformation are formed by at least four different MMPs.

Because several transformations might generate the same MMP, as illustrated in Figure 1B, we apply the following rules to clearly define frequent cliff-forming transformations and avoid redundancies:

(i) If several fragments/transformations yield the same MMPs, then the largest fragment pair is retained provided it corresponds to the same number of cliff-forming compound pairs than alternative fragments.

(ii) If a fragment pair accounts for more cliff-forming compound pairs than a larger fragment pair and if it includes all cliff-forming compound pairs of the larger fragment pair, then this pair is selected instead.

The transformation selection scheme is illustrated in Figure 2.

In order to identify activity cliff-forming transformations occurring in different compound classes, for MMPs repre-

senting the same transformation hierarchy scaffolds were calculated according to Bemis and Murcko[17] using Pipeline Pilot.[18]

Furthermore, in order to identify transformations that rarely introduce activity cliffs, we also ranked transformations using the following function:

$$\text{Score}(\text{transformation}_i) = \frac{\text{occ}_i \times \text{targets}_i}{(1 + \text{targetscliffs}_i) \times (1 + \text{cliffs}_i)} \quad (1)$$

where cliffs_i is the number of activity cliffs introduced by transformation_i, occ_i the sum of transformation_i in all ligand

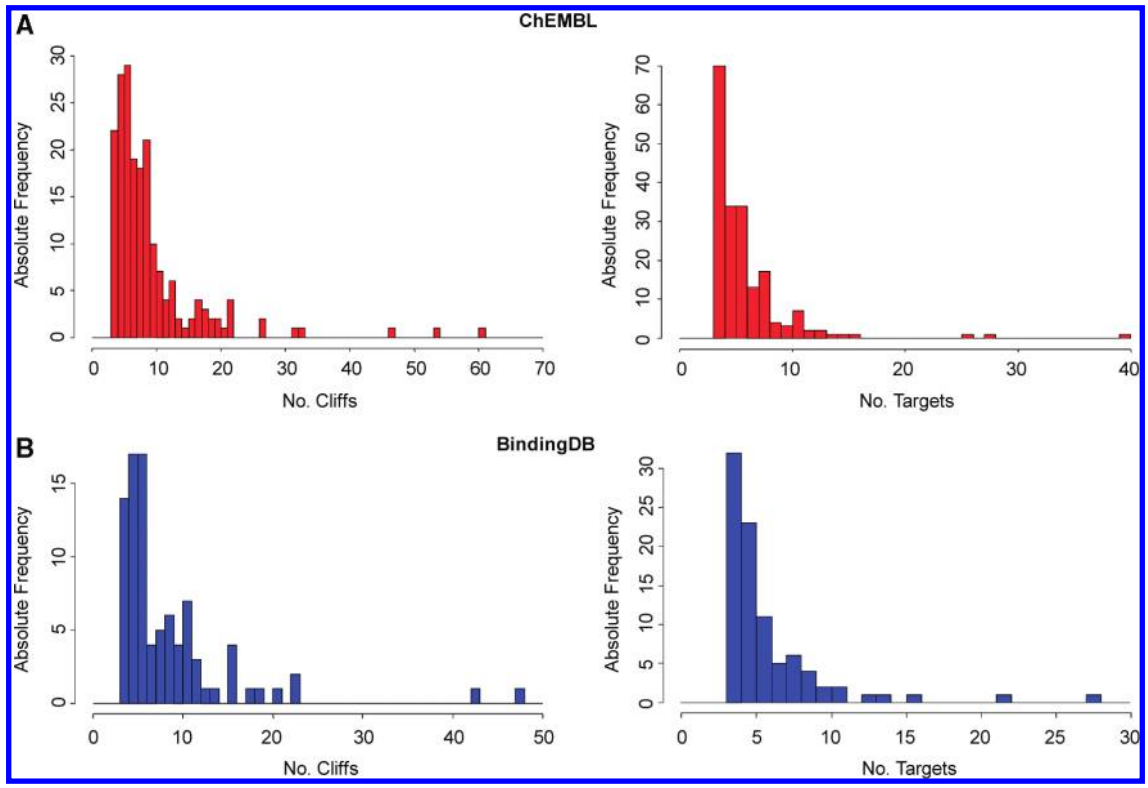

Figure 3. Distribution of cliff-forming transformations. For nonredundant frequent cliff formers, the total number of activity cliffs they form and the number of targets for which cliffs are introduced are reported: (A) 192 ChEMBL and (B) 90 BindingDB transformations.

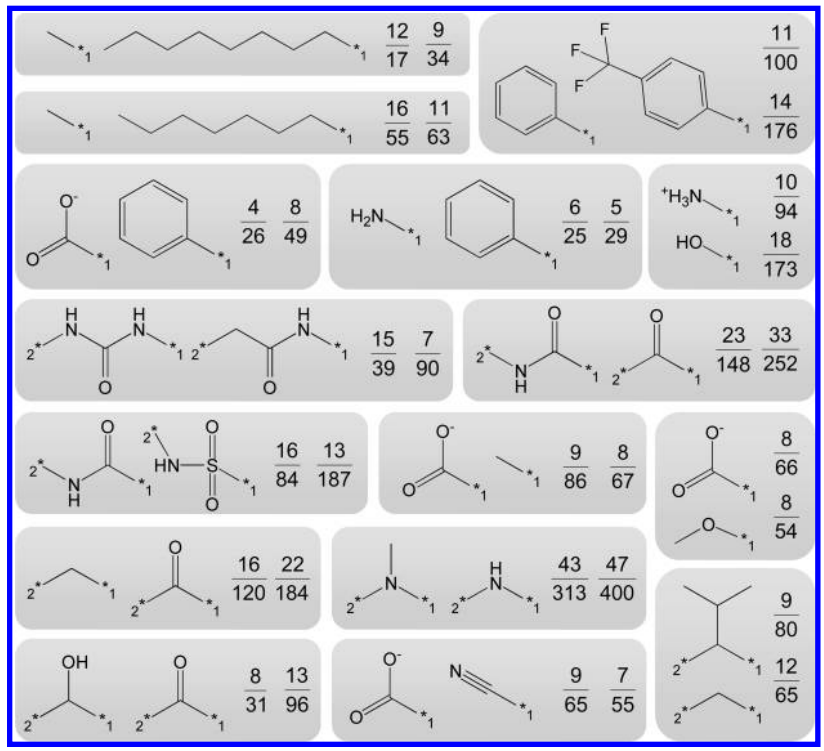

Figure 4. Common transformations. Shown are 16 structural transformations that frequently introduce activity cliffs and are found in both CDB and BDB MMPs. For all transformations, cliff-to-occurrence ratios are reported for the BDB ligand sets (left or top) and the CDB ligand sets (right or bottom).

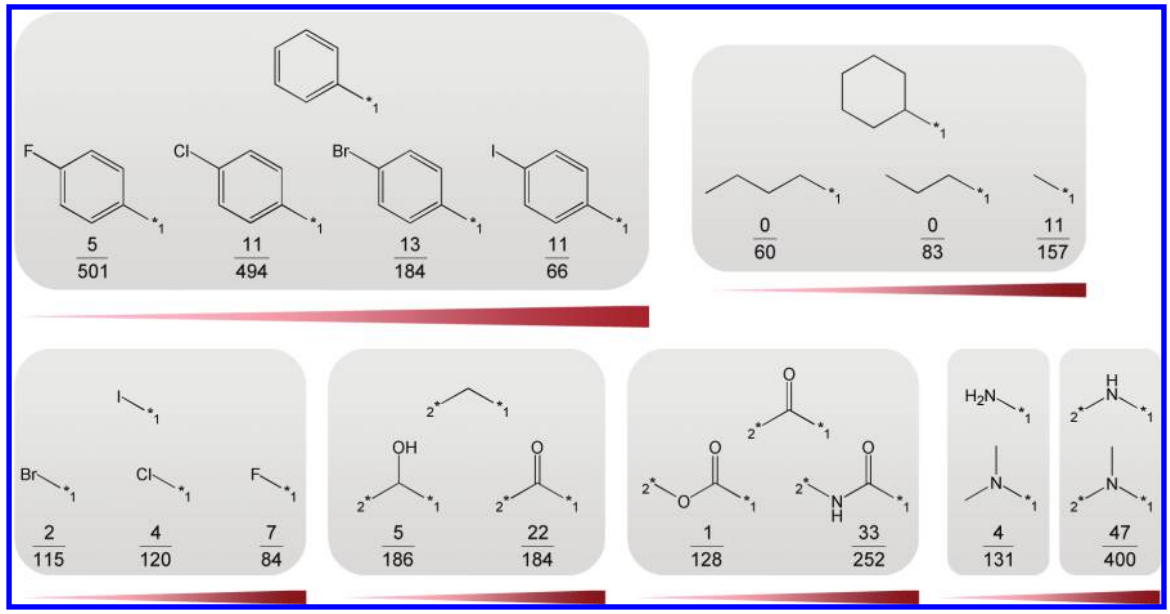

Figure 5. Transformations with different activity cliff potential. Structurally similar transformations are shown together with their relative frequency of cliff formation observed for CDB ligand sets. If more than two fragments are shown in a gray-shaded box, then the fragment at the top is exchanged with each of the fragments at the bottom. From left to right, these fragments are arranged in the order of increasing cliff-forming potential.

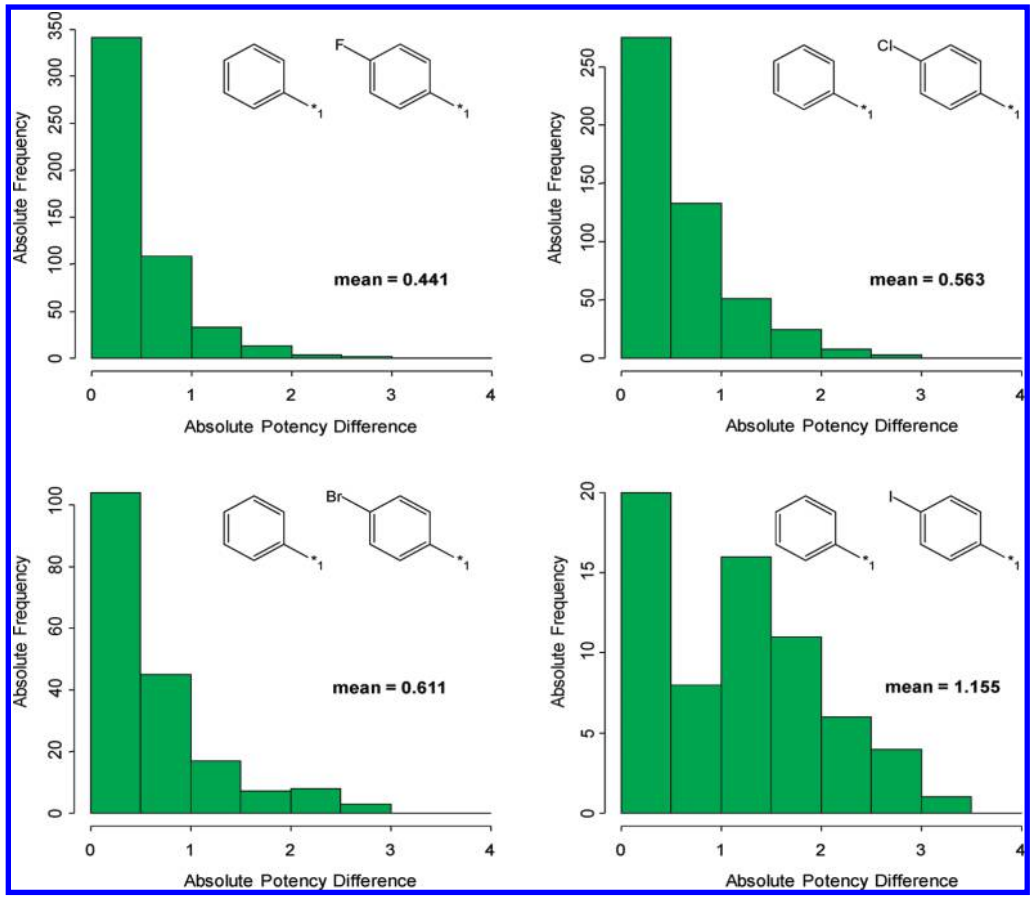

Figure 6. Potency distributions. For four similar transformations, distributions of absolute logarithmic potency differences between compounds in corresponding MMPs found in CDB ligand sets are shown, and arithmetic means of the potency differences are reported. The relative frequency of activity cliff formation of these transformations is given in Figure 5.

sets, targets$_i$ the number of targets for which transformation$_i$ was observed, and targetscliffs$_i$ is the number of targets for which transformation$_i$ introduced activity cliffs. Accordingly, transformations obtain high scores if they are frequently observed in ligand sets of different targets but rarely form activity cliffs for multiple targets.

RESULTS AND DISCUSSION

The major aim of our analysis has been to determine whether chemical changes exist that generally affect compound potency (i.e., across different compound classes and targets). If so, this information might be very helpful in order

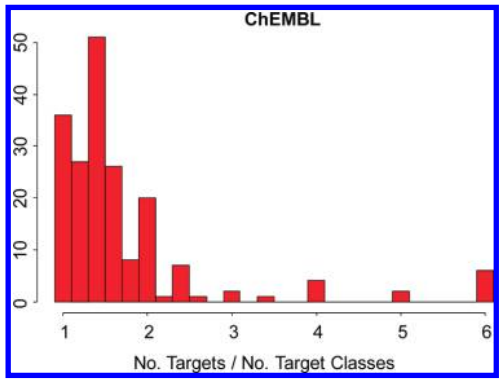

Figure 7. Target distributions of frequent cliff formers. For the 192 CDB cliff-forming transformations, the target-to-target group ratios are reported.

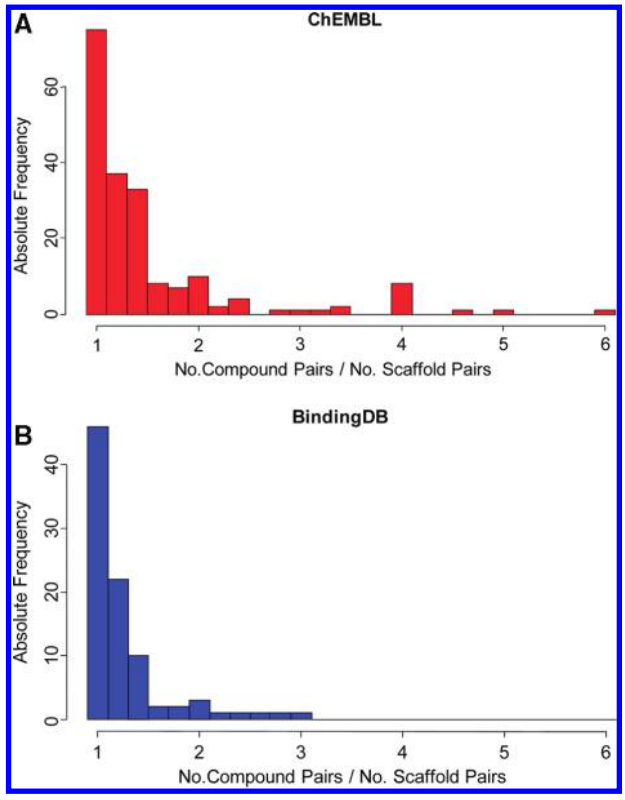

Figure 8. Chemotype distribution of frequent cliff formers. The MMP-to-scaffold pair ratios are reported for: (A) CDB and (B) BDB transformations.

to evaluate chemical modifications for compound optimization efforts. For the purpose of our analysis, we have applied the MMP concept to explore whether defined chemical changes exist that frequently introduce activity cliffs in different ligand sets.

Transformation and MMP Statistics. The initial set of MMPs defined approximately 1.9 and 1.2 million different chemical transformations for CDB and BDB compounds, respectively. Similar to the findings of Hussain and Rea,[8] when analyzing the National Institutes of Health (NIH) Molecular Libraries Small Molecule Repository,[19] the vast majority of these transformations only occurred in a single MMP. Consequently, our criterion that an accepted transformation had to occur in at least two compound pairs restricted the number of transformations to 323 075 and 199 527 for CDB and BDB, respectively. These transformations corresponded to 149 563 (29 565) and 82 213 (17 023) different MMPs (compounds in MMPs) for CDB and BDB, respectively. For 513 of our 523 CDB ligand sets and 307 of our 328 BDB sets, at least one MMP was identified.

Activity Cliff-Forming Substitutions. We considered transformations to form an activity cliff if compounds in corresponding MMPs displayed at least a 100-fold difference in potency. Transformations that introduced activity cliffs for multiple ligand sets (targets) and in different structural contexts (i.e., different compound pairs) were of most interest. Initially, 77 052 and 41 129 cliff-forming transformations were detected in CDB and BDB compounds, respectively. For all of these transformations, the relative frequency with which they introduced cliffs was calculated, i.e., the total number of cliffs they formed divided by their number of occurrences. Applying frequency-, target- and compound pair-oriented filters (see Methods Section), the number of transformations was substantially reduced to 559 and 177 "frequent cliff formers" for CDB and BDB, respectively. These filtered transformations ultimately yielded 192 (CDB) and 90 (BDB) nonredundant frequent cliff formers. SMIRKS[20] representations of these transformations, which provided the basis for our further analysis, are given in the Supporting Information, Tables S1 and S2. The frequency with which they formed activity cliffs ranged from 6.7 to 83.3% and from 6.7 to 70.6% for CDB and BDB ligand sets, respectively (the corresponding average frequencies were 18.1 and 15.0%). For individual transformations, Figure 3 reports the distribution of activity cliffs and of targets for which cliffs were formed. These transformations occurred in up to 51 and 44 cliff-forming MMPs and 40 to 28 CDB and BDB ligand sets, respectively. When analyzing cliff-forming transformations for individual targets, we found that the median cliff-to-occurrence ratio was 50% for both CDB and BDB and that on average 1.5 (CDB) and 1.4 (BDB) cliffs per target were formed. A single transformation formed a maximum of 61 and 48 cliffs for CDB and BDB compounds, respectively.

Consensus Transformations. A total of 37 nonredundant cliff-forming transformations comprising 44 different value fragments were found in both CDB and BDB, as reported in the Supporting Information, Table S3. A representative subset of 16 transformations is shown in Figure 4. The 37 frequent cliff formers included both R-group substitutions and core substructure exchanges. For seven of these transformations, the MMP sets largely overlapped between both databases; for the others, there was little, if any, overlap. The 37 fragment pairs can roughly be divided into two groups, i.e., fragments of similar size that differ in their charge or hydrogen (H)-bonding properties and nonpolar fragments that differ in size.

Transformation Characteristics. For both the CDB and BDB sets of cliff-forming transformations, an enrichment of fragment pairs was observed where one of the exchanged fragments was negatively charged. Among these, transformations involving a carboxyl group were very prevalent. We found that the exchange of a carboxyl group with a variety of other groups (e.g., halogens, methyl, methylether, phenyl, or cyano groups) frequently introduced activity cliffs. For the larger set of CDB cliff-forming transformations, corresponding but much weaker trends were also detected for transformations involving positively charged nitrogens in different structural environments. Other substructures that

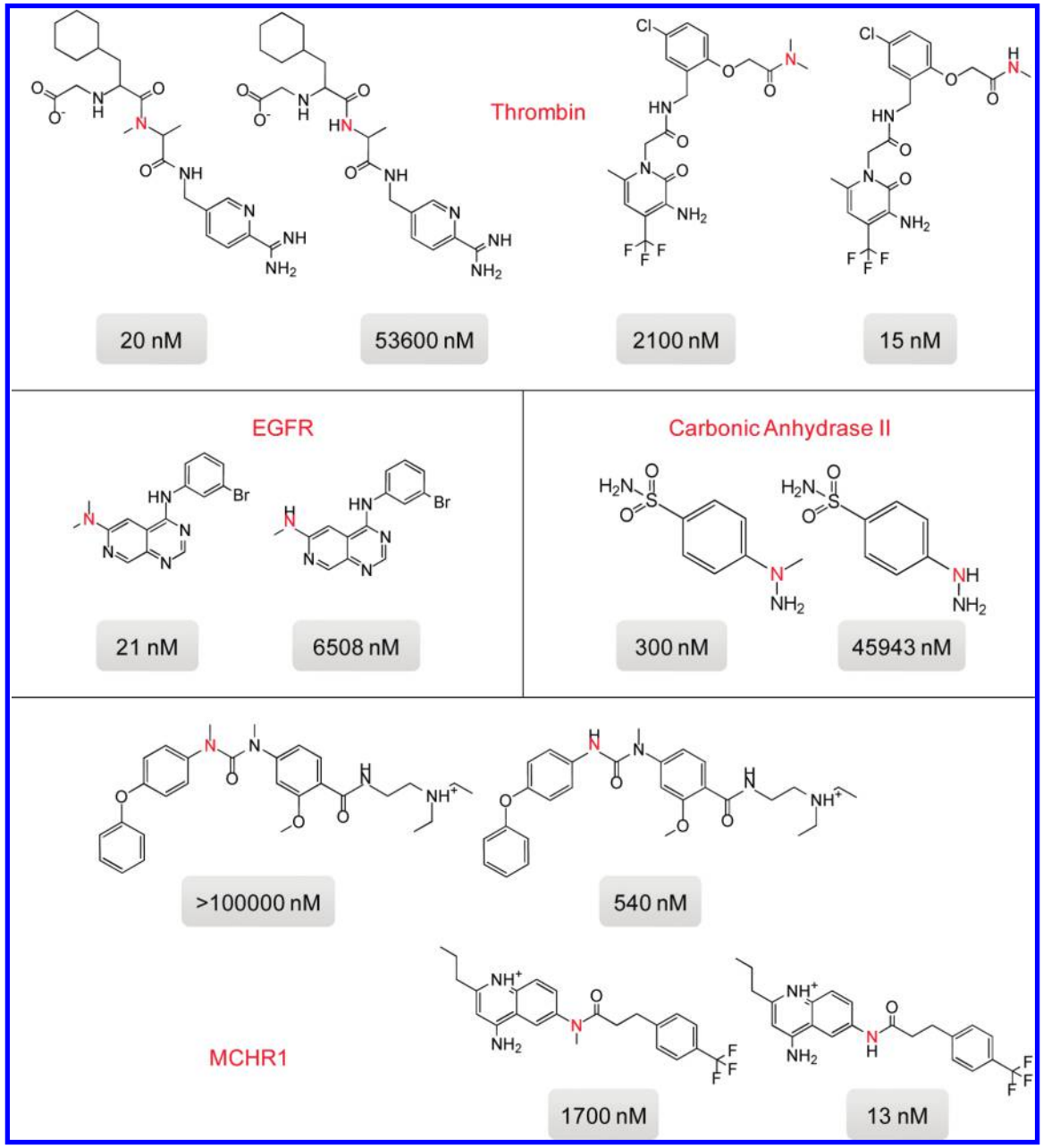

Figure 9. Structurally diverse MMPs. Examples of MMPs are shown that are defined by the same transformation (i.e., the exchange of a secondary against a tertiary amine), contain different scaffolds, and form activity cliffs for different targets. All compounds are annotated with potency values taken from CDB.

were often involved in the formation of activity cliffs included the carbonyl group and amines. By contrast, although changes in ortho, meta, and para substituent positions of phenyl rings were frequently observed, these positional changes rarely introduced activity cliffs. Moreover, a number of very similar transformations were found that displayed substantially different propensities to introduce activity cliffs, as shown in Figure 5. For example, in the CDB ligands sets, introduction of a fluorine atom in para position of a phenyl ring led to an activity cliff in only 5 of 501 cases. However, the relative frequency of cliff formation was found to increase with the size of the halogen substituent, with iodine substitutions introducing cliffs with a nearly 17% frequency. Furthermore, the introduction of a nonterminal hydroxyl group showed a lower cliff frequency than that of a carbonyl oxygen, another rather unexpected finding. In addition, the exchange of a carbonyl group and a carboxyl ester group displayed a significantly lower tendency to cause large potency changes than the exchange of a carbonyl group and an amide. We also observed that distributions of compound potency differences in MMPs shifted toward high values for transformations that displayed an increasing propensity to introduce activity cliffs, as illustrated in Figure 6. Thus, transformations that frequently introduced activity cliffs also displayed the tendency to introduce cliffs of large magnitude.

Target Distribution of Transformations. We also analyzed whether the 192 CDB transformations preferentially introduced activity cliffs for related targets or different target classes. For this purpose, the protein target classification hierarchy of ChEMBL was adopted and slightly extended to group all targets covered by our ligand sets. The target classification used herein is reported in the Supporting Information, Table S4. For each of the 192 transformations,

the number of targets for which it produced activity cliffs was determined and divided by the number of different groups to which these targets belong. The distribution of all target-to-target group ratios is reported in Figure 7. For the majority of transformations that produced MMPs active against multiple targets, the targets mostly belonged to different groups, with a median ratio of 1.33 targets per class. Thus, there was no apparent target group bias of activity cliff-forming transformations.

Chemotype Distribution of Transformations. We were also interested in the question to what extent frequent cliff-forming transformations introduced activity cliffs in different chemotypes. Therefore, for all MMPs that were defined by the same transformation and represented activity cliffs, compound scaffolds were generated. Then the ratio of the number of cliff-forming MMPs and of MMPs representing different scaffold pairs was determined for each transformation. The results for frequent CDB and BDB cliff formers are shown in Figure 8. The MMP-to-scaffold pair ratios were usually close to one, thus indicating that most transformations introduced activity cliffs in variable structural environments. Figure 9 shows examples of MMPs that were defined by the same transformation and contained rather different scaffolds. In each of these compound pairs, a secondary and tertiary amine was exchanged. These MMPs formed activity cliffs against different targets, including thrombin, melanin-concentrating hormone receptor 1, carbonic anhydrase II, and epidermal growth factor receptor. For thrombin, the replacement of the secondary amine by the tertiary amine increased potency by more than three orders of magnitude for the compound pair on the left in Figure 9 and decreased potency by about two orders of magnitude for the compound pair on the right.

In light of these findings, we further investigated the direction of potency changes by transformations inducing multiple activity cliffs for at least one target. We found that 68 (BDB) and 71% (CDB) of these transformations displayed only one potency direction for each target. However, as shown in the Supporting Information, Table S6, there was a clear dependency on the number of different scaffolds in a set of compound pairs. A transformation that occurred in different chemotypes had a considerably higher probability to cause large potency changes in different directions.

Transformations Lacking Cliff-Forming Potential. We also identified a number of transformations that did not display a notable tendency to form activity cliffs. A set of 50 nonredundant transformations that obtained the best sum of ranks for CDB and BDB according to our scoring function (eq 1) is reported in the Supporting Information, Table S5, and the 10 top-ranked transformations are shown in Figure 10. Transformations that did not cause notable potency changes in MMPs included the introduction of small halogens, especially as substituents of aromatic rings, the introduction of methylether groups, and the exchange of aliphatic moieties and rings. In part, these transformations clearly differed from frequent cliff formers that often involved the exchange of polar groups or the replacement of small fragments by larger ones. For example, none of the transformations reported in the Supporting Information, Table S6, exchanges a negatively charged substructure, which was one of the characteristic features of frequent cliff formers. Among the transformations shown in Figure 10, the chlorine

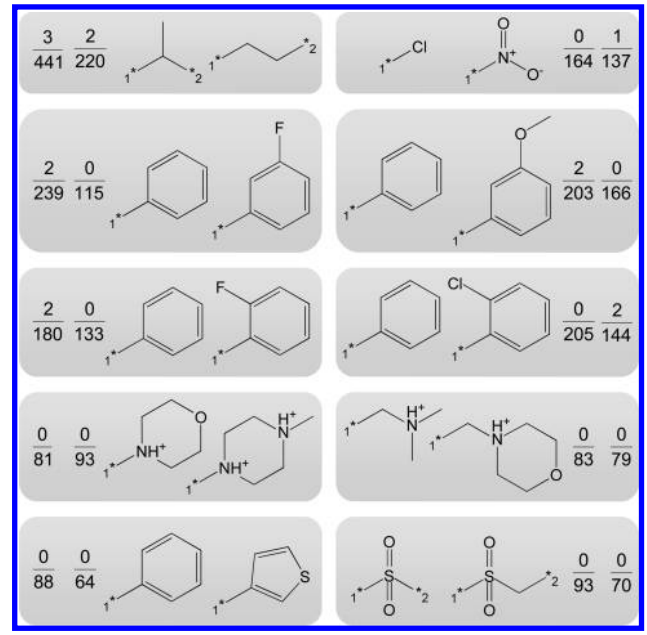

Figure 10. Transformations with lowest cliff-forming potential. Displayed are 10 transformations with the smallest sum of ranks for CDB and BDB (according to eq 1). Transformations are annotated with their cliff-to-occurrence ratios for BDB (on the right) and CDB ligand sets (on the left).

substitution in ortho position at the phenyl ring had essentially no cliff-forming potential, whereas chlorine substitutions at other ring positions introduced activity cliffs at varying frequency. Hence, in some instances, subtle differences between transformations notably altered their tendency to form activity cliffs.

CONCLUSIONS

We have systematically analyzed chemical transformations in public domain compound data sets that defined matched molecular pairs and determined the potential of these substitutions to introduce activity cliffs. More than 200 nonredundant transformations were identified that displayed a strong tendency to introduce cliffs in different chemotypes and across different targets. This general tendency was not necessarily expected. However, clear trends emerged for specific chemical substitutions to globally introduce activity cliffs. By contrast, in other instances, closely related substitutions displayed only little, if any, cliff potential. These findings complement and further extend previous observations that certain molecular scaffolds had a high propensity of activity cliff formation against different targets. The picture emerges that subsets of currently available core structures and chemical transformations are preferred to yield diverse compounds with significant potency variations against different biological targets. Hence, concerning the formation of activity cliffs, privileged scaffolds and substitutions exist. These results have implications for screening library design and compound optimization efforts. We provide a compendium of transformations found in matched molecular pairs having strong activity cliff-forming potential. The notion of such substitution patterns should be helpful in analog design.

ACKNOWLEDGMENT

We wish to thank Lisa Peltason for discussions and Eugen Lounkine for help with the Scientific Vector Language.

Supporting Information Available: Supplementary Tables S1−S6 provide CDB and BDB transformations that predominantly form activity cliffs across different targets, cliff-forming transformations common to both databases, a target classification scheme, potency directionality of cliff-forming transformations, and transformations with low cliff-forming potential. This information is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Topliss, J. G. Utilization of Operational Schemes for Analog Synthesis in Drug Design. *J. Med. Chem.* **1972**, *15*, 1006–1011.
(2) Sheridan, R. P. The Most Common Chemical Replacements in Drug-Like Compounds. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 103–108.
(3) *MDL Drug Data Report (MDDR)*; Symyx Software: San Ramon, CA, 2005.
(4) Haubertin, D. Y.; Bruneau, P. A Database of Historically-Observed Chemical Replacements. *J. Chem. Inf. Model.* **2007**, *47*, 1294–1302.
(5) Raymond, J. W.; Watson, I. A.; Mahoui, A. Rationalizing Lead Optimization by Associating Quantitative Relevance with Molecular Structure Modification. *J. Chem. Inf. Model.* **2009**, *49*, 1952–1962.
(6) Hajduk, P. J.; Sauer, D. R. Statistical Analysis of the Effects of Common Chemical Substituents on Ligand Potency. *J. Med. Chem.* **2008**, *51*, 553–564.
(7) Kenny, P. W.; Sadowski, J. Structure Modification in Chemical Databases. In *Chemoinformatics in Drug Discovery*; Oprea, T. I., Ed.; Wiley-VCH: Weinheim, Germany, 2004; pp 271−285.
(8) Hussain, J.; Rea, C. Computationally Efficient Algorithm to Identify Matched Molecular Pairs (MMPs) in Large Data Sets. *J. Chem. Inf. Model.* **2010**, *50*, 339–348.
(9) Maggiora, G. M. On Outliers and Activity Cliffs - Why QSAR Often Disappoints. *J. Chem. Inf. Model.* **2006**, *46*, 1535.
(10) Bajorath, J.; Peltason, L.; Wawer, M.; Guha, R.; Lajiness, M. S.; Van Drie, J. H. Navigating Structure-Activity Landscapes. *Drug Discovery Today* **2009**, *14*, 698–705.
(11) Hu, Y.; Bajorath, J. Molecular Scaffolds with High Propensity to Form Multi-Target Activity Cliffs. *J. Chem. Inf. Model.* **2010**, *50*, 500–510.
(12) ChemblDB; EMBL-EBI (European Bioinformatics Institute: Cambridge, U.K.; http://www.ebi.ac.uk/chembl/. Accessed January 2, 2010.
(13) Chen, X.; Lin, Y.; Liu, M.; Gilson, M. K. The Binding Database: Data Management and Interface Design. *Bioinformatics* **2002**, *18*, 130–139.
(14) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.
(15) PubChem BioAssay; National Center for Biotechnology Information: Bethesda, MD; http://pubchem.ncbi.nlm.nih.gov/. Accessed January 2, 2010.
(16) *SVL*, Molecular Operating Environment; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2007.
(17) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893.
(18) *Scitegic Pipeline Pilot*, Student Edition, Version 6.1; Accelrys, Inc.: San Diego, CA, 2007.
(19) NIH Molecular Libraries Small Molecule Repository; BioFocus: Saffron Walden, U.K.; http://mlsmr.glpg.com. Accessed January 2, 2010.
(20) *SMIRKS*; Daylight Chemical Information Systems, Inc: Aliso Viejo, CA, 2008.

CI1001845