JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

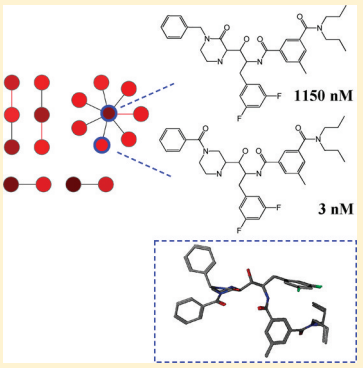
pubs.acs.org/jcim

Exploration of 3D Activity Cliffs on the Basis of Compound Binding Modes and Comparison of 2D and 3D Cliffs

Ye Hu 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

ABSTRACT: Activity cliffs are formed by pairs or groups of structurally similar compounds having large differences in potency and are focal points of structure−activity relationship (SAR) analysis. The choice of molecular representations is a critically important aspect of activity cliffs analysis. Thus far, activity cliffs have predominantly been defined on the basis of molecular graph or fingerprint representations. Herein we introduce 3D activity cliffs derived from comparisons of experimentally determined compound binding modes. The analysis of 3D activity cliffs is generally applicable to target proteins for which structures of multiple ligand complexes are available. For two popular targets, β-secretase 1 (BACE1) and factor Xa (FXa), public domain X-ray structures with bound inhibitors were collected. Crystallographic binding modes of inhibitors were systematically compared using a 3D similarity method taking conformational, positional, and atomic property differences into account. In addition, standard 2D similarity relationships were also determined. SAR information associated with individual compounds substantially changed when either bioactive conformations or 2D molecular graphs were used for similarity evaluation. 3D activity cliffs were identified for BACE1 and FXa inhibitor sets and systematically compared to 2D cliffs. It was found that less than 40% of 3D activity cliffs were conserved when 2D similarity was applied. The limited conservation of 3D and 2D cliffs provides further evidence for the strong molecule representation dependence of activity cliffs. Moreover, 3D cliffs represent a new class of activity cliffs that convey SAR information in ways that differ from graph-based similarity measures. In cases where sufficient structural information is available, the comparison of 3D and 2D cliffs is expected to aid in SAR analysis and mapping of critical binding determinants.



INTRODUCTION

Activity landscapes of compound data sets[1,2] and activity cliffs[3,4] have become popular concepts in structure−activity relationship (SAR) analysis. Activity cliffs are generally defined as pairs or groups of structurally similar compounds having large potency differences. They can be rationalized as a continuum of pairs of cliff-forming compounds with increasing potency differences or as discrete states.[4] In the latter case, potency differences required for cliff formation must be clearly defined. For example, for a variety of applications, an activity cliff definition has been useful that requires a potency difference of at least 2 orders of magnitude between cliff-forming compounds and, in addition, potency of one compound in the nanomolar range.[4] The most prominent activity cliffs in a data set represent centers of SAR discontinuity in activity landscapes, i.e., regions where structurally similar or analogous compounds have large potency differences. Such centers of discontinuity are usually rich in SAR information. However, a major variable in defining activity cliffs is the way compound similarity is assessed.[2,4] In addition, the choice of alternative experimental measurements also affects activity cliff distributions.[5] The strong influence of chosen molecular representations on activity cliff formation has been well-demonstrated in different investigations.[6,7] Activity cliffs have been determined on the basis of various 2D and 3D molecular descriptors,[7] also applying data fusion of different representations,[7] R-group decomposition,[8] matched molecular pair analysis,[9] and protein−ligand interaction fingerprints.[10] It has been found that activity cliffs are frequently formed in a coordinated manner in compound data sets, leading to the introduction of the activity ridge concept.[11] In addition to single-target activity cliffs, multitarget cliffs have also been described.[12]

Although activity cliffs have been studied using different types of descriptors and molecular representations (mostly 2D fingerprints), experimental structure information has thus far only been little considered in activity cliff analysis. In one of two previous studies, selected activity cliffs originally identified on the basis of 2D molecular graphs have been analyzed in light of ligand−target interactions seen in crystallographic complexes.[13] In addition to this qualitative interpretation of individual activity cliffs at the target structure level, another more recent investigation has calculated ligand−target interaction fingerprints from crystal structures as an alternative molecular representation for similarity calculation.[10] Apart from this indirect use of structural information, structural criteria have not been applied to define activity cliffs. In principle, however, the active conformation of a compound should provide a particularly informative representation for activity cliffs. Therefore, we have defined activity cliffs on the basis of

Received: January 18, 2012
Published: March 6, 2012

dx.doi.org/10.1021/ci300033e | J. Chem. Inf. Model. 2012, 52, 670−677

compound binding modes and the assessment of their 3D similarity. Accordingly, these activity cliffs are termed "3D cliffs". Experimental compound binding modes are obtained from X-ray structures of ligand−target complexes. In order to study activity cliff formation in a meaningful way, a sufficiently large number of compounds active against a given target must be available in complex X-ray structures. This generally limits the number of targets and active compounds that can be considered on the basis of currently available public domain structural data. NMR structures might also be used, but there currently are only few ligand−target complex structures available that were determined by NMR. In addition, binding modes might also be modeled, but we concentrate in our analysis on experimental data. Accordingly, we have assembled sets of crystallographically characterized inhibitors of BACE1 and FXa, respectively, for which well-defined potency data were also available and systematically studied 3D cliffs formed by these inhibitors. 3D cliffs were compared with conventional activity cliffs derived on the basis of 2D molecular graphs ("2D cliffs"). As expected, given the very different molecular representations and similarity measures, there was only limited conservation of 3D and 2D cliffs. Thus, 3D cliffs represent a new category of activity cliffs that complement and further extend the conventional analysis of activity cliffs on the basis of molecular graphs.

MATERIAL AND METHODS

Inhibitor Sets. BACE1 and FXa complex structures with nonpolymeric inhibitors with available K_i or IC_{50} values were collected from the Protein Data Bank (PDB).[14] Potency annotations of crystallographic inhibitors were manually curated. Approximate measurements indicated by ">", "<", or "~" relations were discarded. Only inhibitors were included in the analysis if potency values from different data sources or different types of potency values fell within the same order of magnitude. In this case, the geometric mean of available potency values was calculated to yield the final potency annotation of an inhibitor. By contrast, if different measurements and/or values differed by more than 10-fold, the inhibitor was not further taken into consideration.

3D Similarity Assessment. For 3D similarity evaluation, we applied a previously reported methodology that is based on the use of property density functions and takes binding conformations and orientations (positional differences) of crystallographic ligands into account.[15] This methodology was originally developed to systematically explore 3D similarity relationships among series of crystallographic ligands.[15]

Following this approach, α-carbon atoms of selected protein structures per set are first optimally superposed using the sequence/structure alignment function of the Molecular Operating Environment (MOE).[16] This superposition defined the position of all inhibitors for pairwise comparisons. A property density function is calculated for each compound, and the normalized overlap of the functions of two inhibitors is then calculated as a measure of 3D molecular similarity. On the basis of the coordinates of each compound, the property density function is calculated as follows: Each atom is represented by a spherically symmetric Gaussian density function that is centered at the position of the atom nucleus with a width determined by the van der Waals atomic radius. The density function is then defined for different atomic properties by the weighted sum of the density functions of all atoms in the molecule. The atomic Gaussians are weighted with respect to selected atom properties. The atomic property density weight is set to 1 if the atom has this property and to 0 if the property is absent. Four atomic properties are considered, including aromaticity, hydrogen-bond acceptor and donor potentials, and hydrophobic character. The overlap of the property density functions of two compounds is calculated as the sum of the individual property density functions, which also yields a Gaussian:

$$F(X, Y) = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{w_i^p w_j^p + w_i^q w_j^q + \ldots}{mn}\left(\frac{a^2}{2\pi(r_i^2 + r_j^2)}\right)^{3/2} \times \exp\left\{-\frac{a^2}{2(r_i^2 + r_j^2)}|x_i - y_j|^2\right\}$$

where F(X,Y) is the overlap of property density functions of conformations X and Y; X,Y are the matrices of spatial atom coordinates for the two molecules with dimension $3 \times m$ and $3 \times n$, respectively; m, n are the numbers of atoms in molecules X and Y, respectively; x_i is the vector of coordinates of atom i in conformation X; w_i^p is the weight of atom i with respect to property p; $w_i^p = 1$, if atom i has property p, otherwise $w_i^p = 0$; a is the scaling factor, which is set to 2 in our calculations; and r_i is the van der Waals radius of atom i.

Atomic properties were calculated with MOE applying the following definitions:

- $w_i^{\text{aromatic}} = 1$, if atom i is in a ring satisfying the Hueckel rule and is sp^2 hybridized
- $w_i^{\text{donor}} = 1$, if atom i is in pharmacophore class "donor" or in class "base"
- $w_i^{\text{acceptor}} = 1$, if atom i is in pharmacophore class "acceptor" or in class "acid"
- $w_i^{\text{hydrophobic}} = 1$, if atom i is in pharmacophore class "hydrophobe"

A normalization procedure is applied to obtain 3D similarity values between 0 and 1. Therefore, the overlap of the molecular property density functions was divided by the mean self-overlap of the binding conformations:

$$F^{\text{norm}}(X, Y) = \frac{F(X, Y)}{\frac{1}{2}F(X, X)F(Y, Y)}$$

2D Similarity Assessment and 3D Similarity Control. As a standard measure of 2D similarity, Tanimoto similarity[18] of MACCS structural keys[19] was calculated. As a control, Tanimoto similarity values were also calculated for a 3D fingerprint, on the basis of bioactive conformations and low-energy conformations of these inhibitors generated with MOE using pulse conformational search followed by energy minimization. Among a number of different 3D fingerprints tested, the pharmacophore atom triangle (piDAPH3) fingerprint implemented in MOE yielded the best correlation with 3D similarity values calculated on the basis of atomic property density functions. This fingerprint is a three-point phamacophore fingerprint that captures all tuples of the form (u, v, w, d, e, f) for u, v, w atom types and d, e, f interatomic distances.[16] A total of eight types are computed from three atomic properties including "in π system", "is a donor", and "is an acceptor".

dx.doi.org/10.1021/ci300033e | J. Chem. Inf. Model. 2012, 52, 670−677

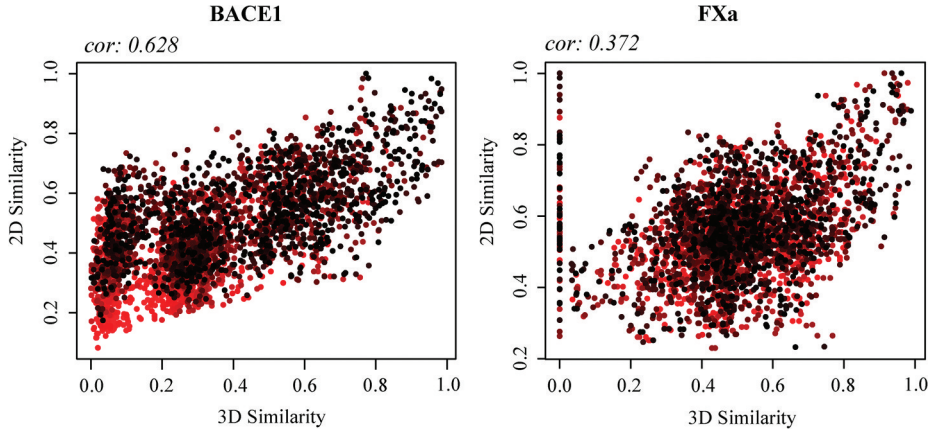


Figure 1. Comparison of 2D and 3D similarity values. Scatter plots report the 2D vs 3D pairwise comparison of 2D and 3D similarity values for BACE1 and FXa inhibitors. In these plots, each data point represents a ligand pair. The dots are color coded according to the potency difference of the compared inhibitors using a continuous spectrum from black (smallest potency difference within the data set) to red (largest potency difference). At the top of each graph, the correlation coefficient (cor) is reported.

Activity Cliff Definition. Activity cliffs were defined as alternative discrete states by considering 80, 85, 90, and 95% 2D or 3D similarity and at least either a 10-fold or 100-fold difference in potency as cliff criteria. Hence, eight categories of 2D and 3D activity cliffs were considered with increasing similarity thresholds and magnitude. Independently defined 2D and 3D cliffs were compared and conserved cliffs determined.

Discontinuity Scores. As a proxy of SAR information content, per-compound discontinuity scores were calculated as follows:[19]

$$\mathrm{disc}_{\mathrm{raw}}(i) = \frac{\sum_{\{i,j|\mathrm{sim}(i,j)>0.65,\mathrm{PotDiff}(i,j)>1,i>j\}} \mathrm{PotDiff}(i,j)*\mathrm{sim}(i,j)}{|\{i,j|\mathrm{sim}(i,j)>0.65,\ \mathrm{PotDiff}(i,j)>1,\ i>j\}|}$$

In this formulation, for molecule i, PotDiff(i,j) indicates the absolute potency difference between i and j if their structural similarity exceeds 65% and their potency differs by more than 1 order of magnitude, and sim(i,j) denotes the 2D or 3D similarity of i and j. The presence of more than 65% 2D structural similarity is required to exclude compound pairs from score calculations that are unable to make meaningful contributions to SAR discontinuity due to the lack of similarity. The raw scores were normalized with respect to all compound scores in the data set to obtain final discontinuity scores between 0 and 1.

Graphical Representation of Activity Cliffs. For activity cliffs formed by pairs of compounds with at least 80% 2D or 3D similarity and a 10- or 100-fold difference in potency, network representations were generated. For BACE1 and FXa inhibitors, all qualifying pairs of compounds were selected. In this activity cliff network, nodes represent compounds that are connected by an edge if they form a cliff (i.e., each edge represents an activity cliff). Black edges indicate activity cliffs that are only formed on the basis of 2D or 3D similarity (i.e., 2D or 3D cliffs) and red edges conserved 2D−3D cliffs. In addition, nodes are color coded according to potency values of inhibitors using a continuous spectrum from black (lowest potency within the data set) to red (highest potency). Network representations were drawn with Cytoscape.[20]

RESULTS AND DISCUSSION

Data Sets. From the PDB, 155 BACE1 and 109 FXa complex X-ray structures were initially retrieved. For the inhibitors contained in these structures, reported potency data were analyzed, and ultimately, 81 BACE1 and 79 FXa inhibitors with available bioactive conformations and well-defined potency values were obtained, hence providing compound data sets of reasonable size for activity cliff analysis.

Table 1. Similarity Values[a]

		similarity threshold			
		BACE1		FXa	
percentile		3D	2D	3D	2D
high similarity	95%	0.77	0.74	0.81	0.74
	90%	0.68	0.68	0.73	0.72
	85%	0.62	0.63	0.69	0.68
	80%	0.57	0.60	0.65	0.65
	75%	0.53	0.58	0.62	0.63
low similarity	5%	0.04	0.23	0.20	0.34
	10%	0.05	0.27	0.30	0.39
	15%	0.07	0.30	0.34	0.42
	20%	0.10	0.33	0.38	0.46
	25%	0.16	0.35	0.41	0.48

[a]2D and 3D similarity threshold values are reported for different percentiles. For example, for BACE1 inhibitors, 95% of all 3D similarity values fall below the threshold of 0.77.

Compound Binding Modes. For the definition of 3D activity cliffs, as defined herein, the similarity of compound binding modes must be quantified in a pairwise manner. In order to compare compound binding modes and accurately calculate 3D similarity, it is required to take conformational, positional, and chemical differences into account. Therefore, we applied our property density function-based 3D similarity method, as detailed in the Material and Methods Section, following optimal superposition of target protein structures to provide a consistent reference frame for ligand comparison. Only a crude approximation of binding mode similarity would be obtained if bioactive conformations were extracted from X-ray structures and subsequently superposed. These potential caveats must be carefully considered when evaluating the similarity of compound binding modes.

Comparison of 3D and 2D Similarity. We systematically compared inhibitors in each data set in a pairwise manner and calculated 3D and 2D similarity values. The comparison of these values is presented in Figure 1. For BACE1 inhibitors,

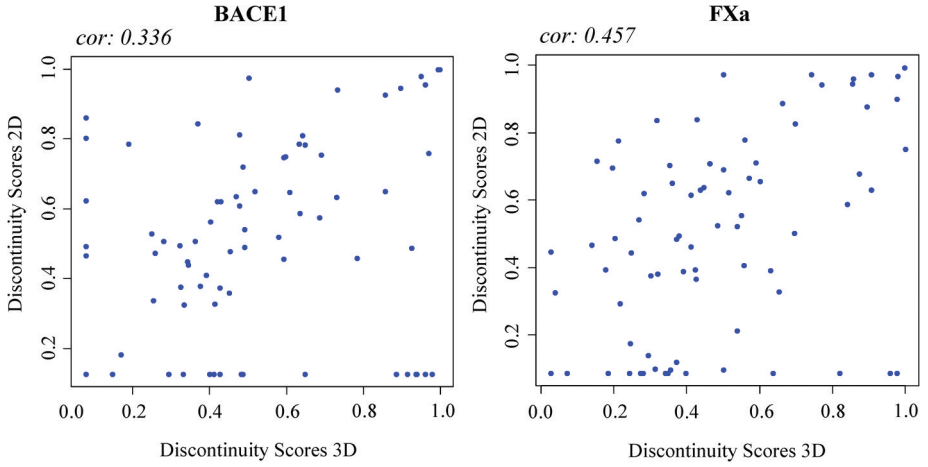


Figure 2. Comparison of discontinuity scores. Compound discontinuity scores calculated on the basis of 2D or 3D similarity are compared in scatter plots for BACE1 and FXa inhibitors. Here, each dot represents an inhibitor. At the top of each graph, the correlation coefficient (cor) is reported.

Table 2. Activity Cliff Statistics[a]

thresholds			number of activity cliffs		
similarity	target	potency difference	2D cliffs	3D cliffs	2D−3D cliffs
95%	BACE1	100-fold	1	1	0
		10-fold	1	6	0
	FXa	100-fold	0	1	0
		10-fold	7	6	3
90%	BACE1	100-fold	1	2	0
		10-fold	4	17	2
	FXa	100-fold	1	1	0
		10-fold	17	14	7
85%	BACE1	100-fold	5	4	1
		10-fold	13	24	4
	FXa	100-fold	2	4	0
		10-fold	26	31	12
80%	BACE1	100-fold	13	9	3
		10-fold	34	41	16
	FXa	100-fold	4	10	0
		10-fold	48	74	26

[a]For each activity cliff category, the numbers of 2D, 3D, and conserved 2D−3D activity cliffs are reported for BACE1 and FXa inhibitors. For example, the "95%" and "100-fold" category means that a pair of inhibitors must share at least 95% (2D or 3D) similarity and display a difference in potency of at least two orders of magnitude to qualify for the formation of a cliff. It follows that "95%" and "100-fold" represent the most stringent and "80%" and "10-fold" the most relaxed activity cliff criteria applied herein.

limited correlation between 3D and 2D similarity was observed (correlation coefficient 0.628). In this case, pairs of inhibitors with largest potency differences also displayed overall lowest 3D and 2D similarity. For FXa inhibitors, the correlation was even weaker (0.372), i.e., there essentially was no global correlation. Here, a number of pairs of inhibitors with very low 3D similarity but high 2D similarities was observed. In these instances, corresponding rings with partly similar heteroatom substitutions adopted rotated binding orientations, which essentially led to the absence of overlap between corresponding atomic property density functions.

In Table 1, corresponding 3D and 2D similarity values are reported for different percentiles of the similarity value distributions. For the high similarity range including the 75% percentile as the lower boundary, 3D and 2D similarity values were very similar for BACE1 and also FXa inhibitors. By contrast, for the low similarity range including the 25% percentile as the upper boundary, corresponding 3D and 2D similarity values differed in part significantly, especially for BACE1 inhibitors. Thus, correspondence between 3D and 2D similarity values was stronger for high similarity (the activity cliff-relevant range) than low similarity.

As a control, we also calculated theoretical 3D similarity using alternative fingerprint representations implemented in MOE. When experimental bioactive conformations were used as input, the best-performing 3D fingerprint, piDAPH3, yielded a correlation coefficient for directly calculated 3D similarity (i.e., using the property density function approach) and fingerprint similarity values of 0.702 and 0.452 for BACE1 and FXa inhibitors, respectively. When modeled low-energy conformations were used as input for piDAPH3 calculations, correlation coefficients of only 0.530 (BACE1) and 0.420 (FXa) were obtained. The latter calculation is typically carried out when no experimental binding conformations are available. Thus, calculation of 3D fingerprint similarity did not offer any advantage compared to 2D similarity calculations. Even when known bioactive conformations were used for 3D fingerprint calculations, only limited (BACE1) or poor (FXa) correlation between 3D similarity and 3D fingerprint similarity values was observed. Hence, in the following, only directly calculated 3D and 2D similarities are considered.

SAR Discontinuity. We next calculated compound discontinuity scores for inhibitors in both sets on the basis of 3D or 2D similarity values. These scores account for the contribution of individual compounds to introduce local SAR discontinuity in the activity landscape of a compound data set. A molecule makes a large contribution to SAR discontinuity if it has a potency value very different from its immediate structural neighbors. Thus, compound discontinuity scores reflect local topological features of activity landscapes and serve as a proxy for SAR information. The presence of high compound discontinuity scores is an indicator of activity cliff formation. The comparison of discontinuity scores is shown in Figure 2. For BACE1 and FXa inhibitors, very little correlation, if any, was observed between discontinuity scores calculated on the basis of 3D and 2D similarity values, with correlation coefficients of 0.34 and 0.46, respectively. These findings indicated that SAR information contained in the two inhibitor

dx.doi.org/10.1021/ci300033e | J. Chem. Inf. Model. 2012, 52, 670−677

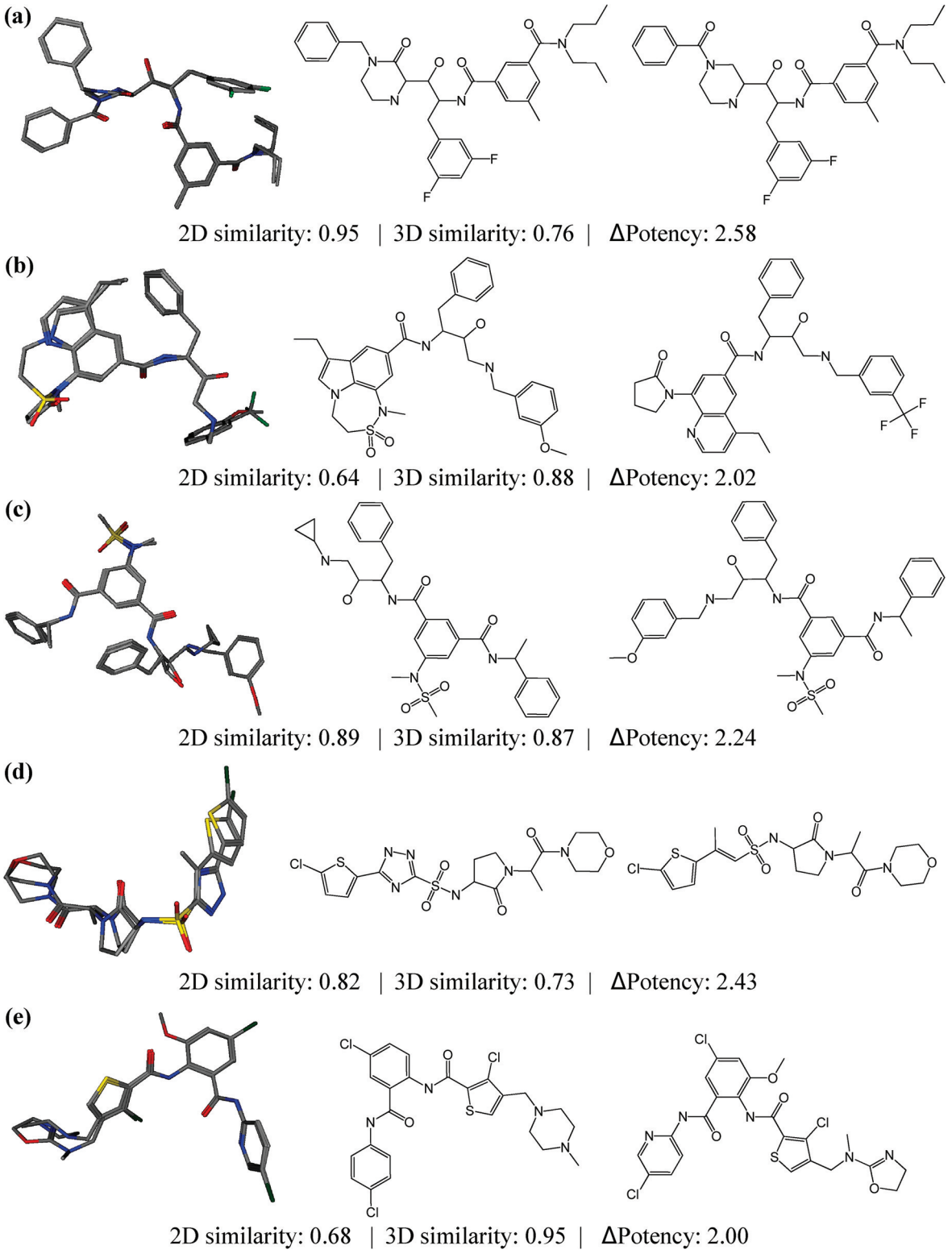


Figure 3. Activity cliffs. Three representative activity cliffs are shown for BACE1 inhibitors including a (a) 2D, (b) 3D, and (c) 2D−3D cliff. In addition, for FXa inhibitors, (d) 2D and (e) 3D cliffs are shown. For each cliff, experimental compound binding modes are compared on the left, and the corresponding 2D structures are depicted on the right. Furthermore, 2D and 3D similarity values and absolute logarithmic potency differences between inhibitors in each pair are reported.

sets generally changed with the use of alternative 2D or 3D similarity measures. Accordingly, changes in activity cliff distributions were also anticipated.

Activity Cliff Distribution. Eight categories of 3D and 2D activity cliffs were defined using different combinations of four similarity threshold values (i.e., 80, 85, 90, or 95% 3D or 2D similarity) and two potency differences (i.e., at least 10- or 100-fold). For each category, the numbers of 3D and 2D cliffs were determined for both inhibitor sets and compared. The results are reported in Table 2. In both data sets, 3D and 2D cliffs were

dx.doi.org/10.1021/ci300033e | J. Chem. Inf. Model. 2012, 52, 670−677

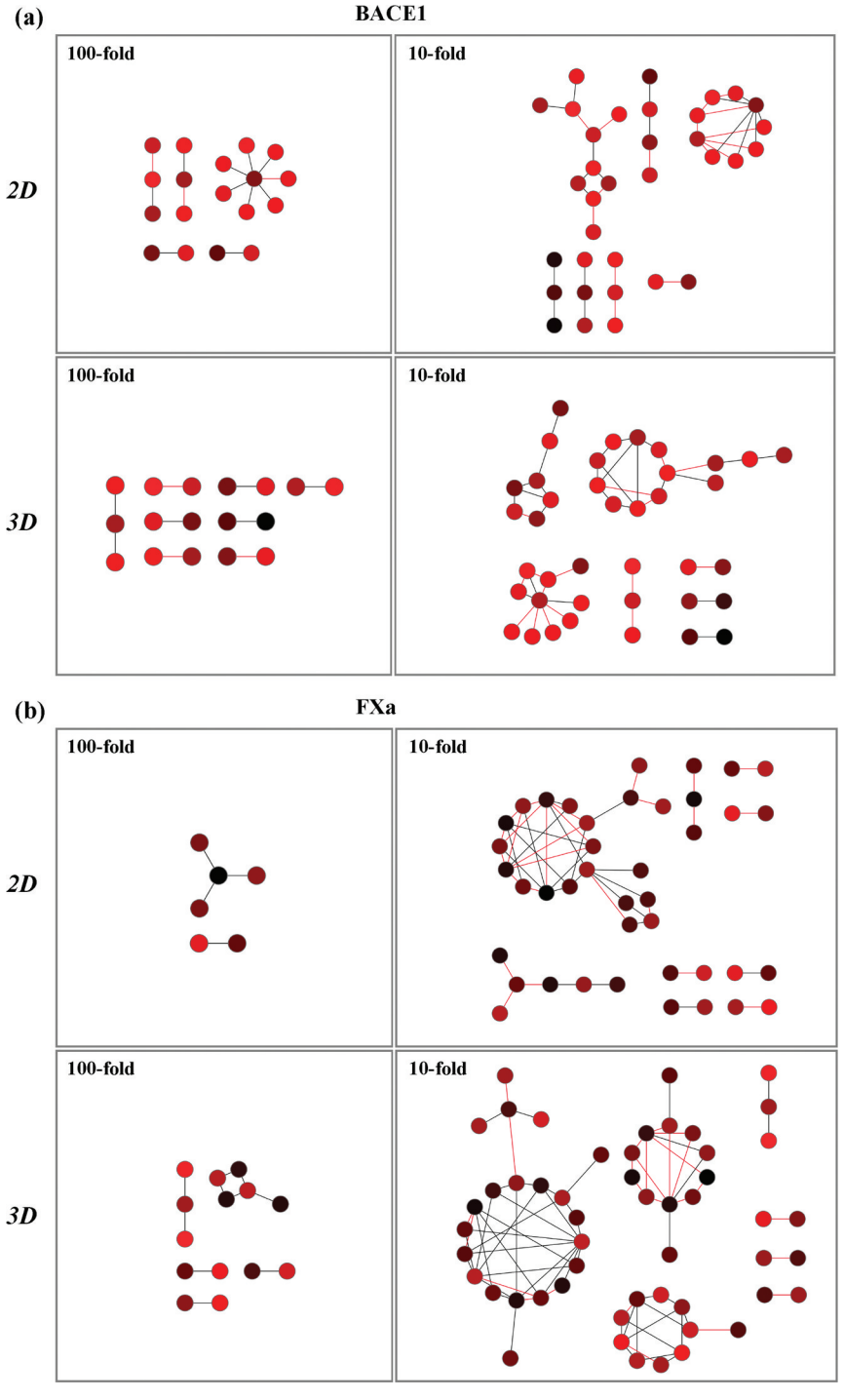


Figure 4. Activity cliff networks. As criteria for the formation of activity cliffs, the presence of 80% 2D (top) or 3D (bottom) similarity and a 10-fold (left) or 100-fold (right) difference in potency were applied. Shown is a comparison of 2D and 3D activity cliffs for (a) BACE1 and (b) FXa inhibitors. Activity cliffs are visualized in a network representation containing all pairs of compounds meeting the cliff criteria. In each network, nodes represent ligands that are connected by an edge if they form an activity cliff. Edges indicating conserved 2D−3D activity cliffs are colored red. In addition, nodes are color coded according to potency values of the inhibitors using a continuous spectrum from black (lowest potency within the data set) to red (highest potency).

detected at increasing levels of stringency. However, the conservation of 3D and 2D cliffs was generally limited. For the 80%/100-fold category, BACE1 inhibitors formed 13 2D and 9 3D cliffs, only 3 of which were conserved. For the 85%/100-fold category, only 5 2D, 4 3D, and 1 2D−3D cliffs were found. FXa inhibitors consistently formed fewer cliffs than BACE1 inhibitors for 100-fold potency differences and more for 10-fold differences. For the 80%/100-fold category, 4 2D and 10 3D cliffs were found for FXa inhibitors and for the 85%/100-fold category, 2 2D and 4 3D cliffs. However, there were no conserved 2D and 3D cliffs among FXa inhibitors with at least 100-fold potency differences; 2D-3D cliffs were only found for 10-fold differences.

dx.doi.org/10.1021/ci300033e | J. Chem. Inf. Model. 2012, 52, 670−677

For the 80%/10-fold category, which includes all cliffs falling into the other seven categories, there were a total 34 2D and 41 3D cliffs for BACE1 and 48 2D and 74 3D cliffs for FXa inhibitors. Of these, 16 and 26 cliffs were conserved for BACE1 and FXa, respectively. Thus, of the 3D cliffs formed by BACE1 and FXa inhibitors, only ∼39% and ∼35% were conserved, respectively, when activity cliffs were determined on the basis of 2D similarity. Figure 3 shows exemplary 2D, 3D, and 2D−3D activity cliffs.

Activity Cliff Networks. Cliff distributions have been compared in network representations. Figure 4 shows networks of activity cliffs with at least 80% (2D or 3D) similarity and 10- or 100-fold differences in potency. For both BACE1 (Figure 4a) and FXa inhibitors (Figure 4b), it was observed that activity cliffs were often not formed in isolation (i.e., by compound pairs without immediate structural neighbors) but in a coordinated manner, consistent with earlier observations made for conventional 2D cliffs.[11] Coordinated cliff formation means that individual compounds were frequently involved in the formation of multiple cliffs with structural neighbors. However, several isolated cliffs were also found. For both sets of inhibitors, there was a strong tendency of coordinated (3D or 2D) cliff formation when required potency differences were lowered to 10-fold (Figure 4). Hence, in both cases, a series of structurally similar compounds predominantly had moderate differences in potency. However, for the more sparsely distributed large-magnitude activity cliffs, differences between 3D and 2D cliffs were also observed. For example, 3D cliffs formed by BACE1 inhibitors at the 100-fold level were (with one exception) isolated, including three 2D-3D cliffs, whereas the corresponding 2D cliffs were mostly formed in a coordinated manner, also including the three conserved cliffs. This observation also implied that chemical differences between inhibitors were resolved at higher resolution at the level of binding modes than 2D graph comparison, as might be expected.

Concluding Remarks. Herein we have introduced the concept of 3D activity cliffs that are defined on the basis of experimentally observed compound binding modes. For this purpose, careful quantification of binding mode similarity is required, taking conformational, positional, and chemical criteria into account. Despite the very rapid growth of the PDB, limited availability of ligand−target complex structures might make the evaluation of 3D activity cliffs difficult in many cases. Of course, it would also be possible to conduct 3D activity cliff analysis on the basis of hypothetical binding modes (e.g., docked complexes). However, because activity cliffs and SAR analyses greatly benefit from a high level of rigor, we currently do not make use of hypothetical ligand−target complexes, given their approximate and error-prone nature. For the two enzyme inhibitor sets studied here, a number of 3D activity cliffs were identified and compared with conventional 2D cliffs (derived on the basis of molecular graphs) that were also determined. Only limited conservation of 3D and 2D cliffs was observed, which highlights the strong molecular representation dependence of activity cliff formation. The presence of limited correlation between 3D and 2D similarity values, which leads to nonconservation of cliffs, is anticipated (and consistent with previous findings of independent studies). Given the discrepancy between 2D and 3D similarity assessment, 3D cliffs add to SAR information that can be obtained on the basis of 2D representations. Importantly, 3D cliffs, as defined herein, represent a new category of activity cliffs that capture conformational, positional and chemical information. Although experimental binding modes should principally contain more activity-relevant information than 2D molecular graphs of active compounds, the evaluation of 3D activity cliffs is also not free of approximations and potential inaccuracies. However, given the limited conservation of 3D and 2D cliffs observed in our analysis, the study of 3D activity cliffs is highly recommended in cases where sufficient structural information is available. This is the case because 3D activity cliffs present SAR information in a different way compared to other activity cliff representations. Especially for the derivation of pharmacophore hypotheses or queries, the analysis of 3D activity cliffs in compound data sets should be very helpful. In summary, the introduction of 3D cliffs further extends the activity cliff concept and provides an opportunity to analyze available experimental ligand structure information from an immediate SAR perspective. In addition, the comparison of 3D and 2D cliffs is expected to often provide complementary information.

■ AUTHOR INFORMATION

Corresponding Author

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

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) 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.

(2) Wassermann, A. M.; Wawer, M.; Bajorath, J. Activity Landscape Representations for Structure-Activity Relationship Analysis. *J. Med. Chem.* **2010**, *53*, 8209−8223.

(3) Maggiora, G. M. On Outliers and Activity Cliffs − Why QSAR often Disappoints. *J. Chem. Inf. Model.* **2006**, *46*, 1535−1535.

(4) Stumpfe, D.; Bajorath, J. Exploring Activity Cliffs in Medicinal Chemistry. *J. Med. Chem.* **2012**, 55, DOI: 10.1021/jm201706b.

(5) Stumpfe, D.; Bajorath, J. Assessing the Confidence Level of Public Domain Compound Activity Data and the Impact of Alternative Potency Measurements on SAR Analysis. *J. Chem. Inf. Model.* **2011**, *51*, 3131−3137.

(6) Peltason, L.; Iyer, P.; Bajorath, J. Rationalizing Three-Dimensional Activity Landscapes and the Influence of Molecular Representations on Landscape Topology and the Formation of Activity Cliffs. *J. Chem. Inf. Model.* **2010**, *50*, 1021−1033.

(7) Medina-Franco, J. L.; Martínez-Mayorga, K.; Bender, A.; Marín, R. M.; Giulianotti, M. A.; Pinilla, C.; Houghten, R. A. Characterization of Activity Landscapes using 2D and 3D Similarity Methods: Consensus Activity Cliffs. *J. Chem. Inf. Model.* **2009**, *49*, 477−491.

(8) Agrafiotis, D. K.; Wiener, J. J. M.; Skalkin, A.; Kolpak, J. Single R-Group Polymorphisms (SRPs) and R-Cliffs: An Intuitive Framework for Analyzing and Visualizing Activity Cliffs in a Single Analog Series. *J. Chem. Inf. Model.* **2011**, *51*, 1122−1132.

(9) Wassermann, A. M.; Bajorath, J. Chemical Substitutions That Introduce Activity Cliffs across Different Compound Classes and Biological Targets. *J. Chem. Inf. Model.* **2010**, *50*, 1248−1256.

(10) Seebeck, B.; Wagener, M.; Rarey, M. From Activity Cliffs to Target-specific Scoring Models and Pharmacophore Hypotheses. *ChemMedChem* **2011**, *6*, 1630−1639.

(11) Vogt, M.; Huang, Y.; Bajorath, J. From Activity Cliffs to Activity Ridges: Informative Data Structures for SAR Analysis. *J. Chem. Inf. Model.* **2011**, *51*, 1848−1856.

(12) Wassermann, A. M.; Dimova, D.; Bajorath, J. Comprehensive Analysis of Single- and Multi-Target Activity Cliffs Formed by

Currently Available Bioactive Compounds. *Chem. Biol. Drug Des.* **2011**, *78*, 224−228.
(13) Sisay, M. T.; Peltason, L.; Bajorath, J. Structural Interpretation of Activity Cliffs Revealed by Systematic Analysis of Structure-Activity Relationships in Analog Series. *J. Chem. Inf. Model.* **2009**, *49*, 2179−2189.
(14) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(15) Peltason, L.; Bajorath, J. Molecular Similarity Analysis Uncovers the Presence of Heterogeneous Structure-Activity Relationships and Variable Activity Landscapes within Active Sites. *Chem. Biol.* **2007**, *14*, 489−497.
(16) *MOE (Molecular Operating Environment)*; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2007.
(17) Willett, P. Searching Techniques for Databases of Two- and Three-Dimensional Structures. *J. Med. Chem.* **2005**, *48*, 4183−4199.
(18) *MACCS Structural Keys*; Symyx Software: San Ramon, CA, 2005.
(19) Wawer, M.; Peltason, L.; Weskamp, N.; Teckentrup, A.; Bajorath, J. Structure-Activity Relationship Anatomy by Network-like Similarity Graphs and Local Structure-Activity Relationship Indices. *J. Med. Chem.* **2008**, *51*, 6075−6084.
(20) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498−2504.