JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

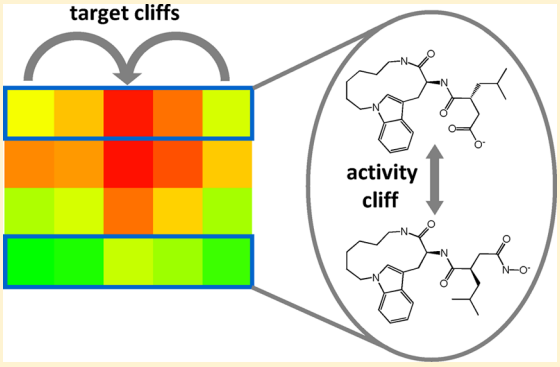
Introduction of Target Cliffs as a Concept To Identify and Describe Complex Molecular Selectivity Patterns

Ye Hu and Jürgen Bajorath*

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

Supporting Information

ABSTRACT: The study of target specificity or selectivity of small molecules is an important task in drug design. In an ideal situation, a compound would exclusively interact with an individual target and hence be target specific. However, such exclusive binding events are likely to be rare, as increasing evidence suggests. Because many compounds are active against more than one target, apparent selectivity often results from potency differences, i.e., a compound that is highly potent against a given target and weakly potent against one or more others displays target selectivity. In a simple case, a compound might have known activity against a pair of targets and be selective for one over the other. Then, selectivity is straightforward to rationalize. However, there are many more complex selectivity relationships associated with multi-target activities of compounds that are difficult to analyze and compare in a consistent manner. For this purpose, we introduce herein target cliffs as a concept to describe complex selectivity patterns. A target cliff is defined as a pair of targets against which at least one compound displays a large difference in potency. As such, target cliffs are distinct from activity cliffs. However, qualifying target pairs (target cliffs) and compound pairs (activity cliffs) can be systematically extracted from the same data structure termed target-compound matrices. Furthermore, these two types of cliffs can be compared to identify and prioritize compounds that are selective and reveal structure−activity relationship (SAR) information.



INTRODUCTION

Target specificity and selectivity continue to be the guiding principles of compound design in medicinal chemistry.[1] In addition, a high degree of target selectivity is essential for small molecular probes used in chemical biology.[2] At the same time, increasing evidence of multi-target activities of small molecules and ensuing polypharmacological effects[3] suggests that exclusive compound-target binding events are rare. Rather, apparent target selectivity is often a consequence of potency differences of a compound against multiple targets such as members of the same protein family. In these situations, target selectivity might often be difficult to rationalize. In a simple case, a compound might be selective for one target over another. However, selectivity patterns might quickly become complex and difficult to understand when a compound is active against more targets. For example, if a compound is highly potent against two targets and weakly potent against three other targets, how should one describe these target selectivity relationships?

Currently, no formalism exists for a consistent assessment of multi-target selectivity relationships. For this purpose, we introduce *target cliffs* as a concept to rationalize and compare selectivity patterns of active compounds in a systematic manner. Importantly, target cliffs are conceptually distinct from activity cliffs. Following the most general definition, *an activity cliff is formed by a pair of structurally similar compounds with a large potency difference against a given target.*[4,5] Accordingly, activity cliffs are of high interest in medicinal chemistry to identify chemical modifications of structural analogs that determine structure−activity relationships (SARs).[5]

By contrast, we define a target cliff as *a pair of targets against which at least one compound displays a large difference in potency.* Thus, different from the activity cliff concept and its extensions,[5] the focus changes here from *compound pairs to target pairs.* Thus, target cliffs define *compound-based target relationships.* Moreover, target cliffs represent *a measure for the target selectivity of individual compounds.* The latter aspect is of high relevance because target cliff distributions can be used to characterize and compare the selectivity patterns of individual compounds that are active against multiple targets. Herein, we systematically identify target cliffs across all current targets from publicly available compound activity data and characterize complex molecular selectivity patterns. The underlying data structure from which target cliffs are derived, a so-called target-compound matrix, also provides direct access to activity cliffs that are formed by compounds involved in target cliffs. We show that the information provided by these different types of cliffs can be considered in a complementary manner. Therefore, we also determine the propensity with which compounds displaying

Received: December 14, 2012

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

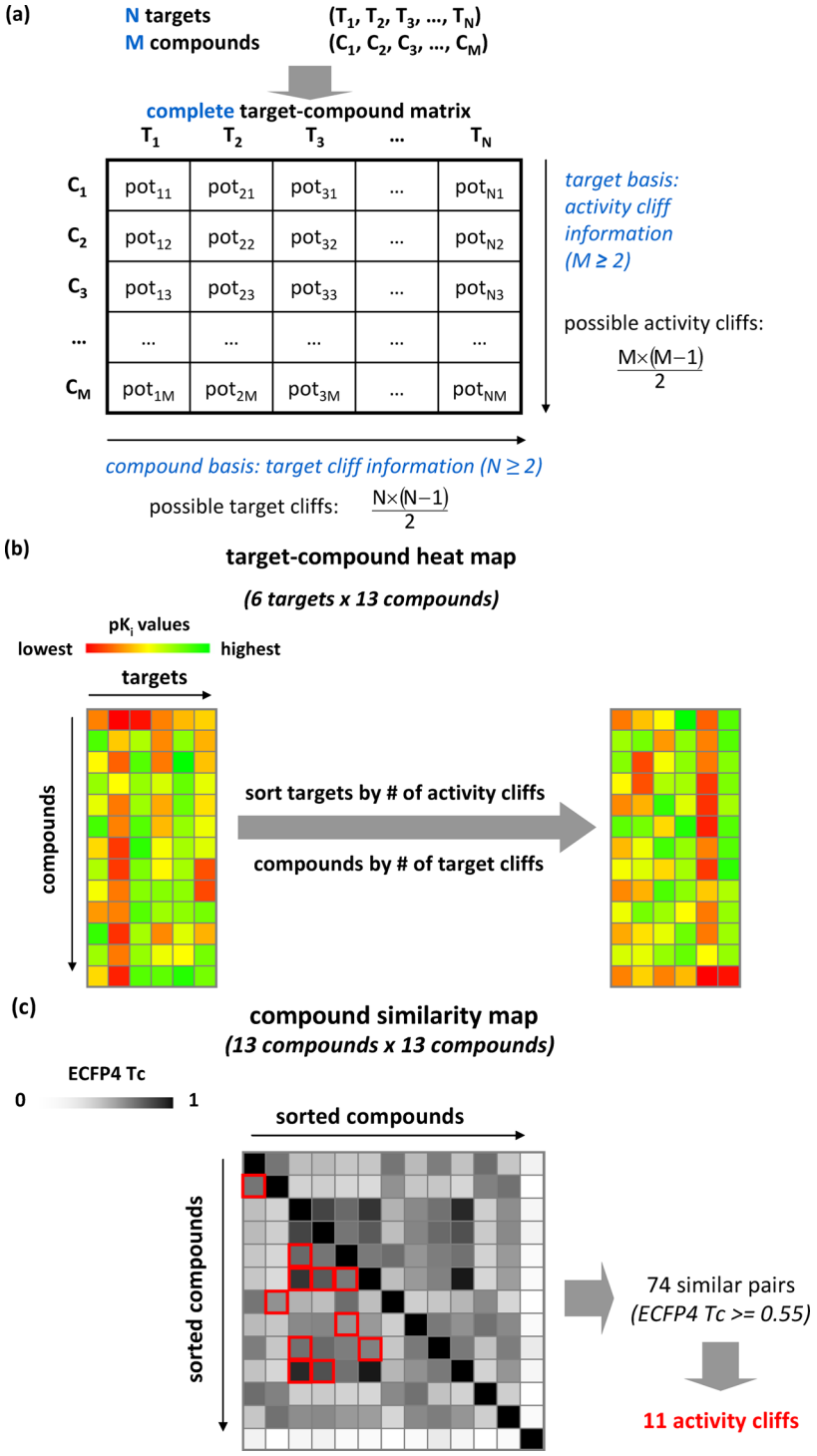


Figure 1. Target-compound matrix. (a) Schematic illustration of a target-compound matrix consisting of N targets and M compounds. The matrix is complete, i.e., each of the M compounds is active against all N targets. Such matrices are derived for all active compounds on the basis of their target annotations. If several compounds are active against the same targets, they form an individual target-compound matrix. The smallest matrix for selectivity analysis results from a single compound active against a pair of targets. From the target-compound matrix, both target cliffs and activity cliffs can be systematically extracted. For individual targets, activity cliffs are obtained by traversing the matrix vertically provided multiple compounds are available. In addition, by traversing the matrix horizontally, target cliffs are obtained for individual compounds provided multiple targets are available. The number of theoretically possible target cliffs and activity cliffs in a matrix is provided. (b) An exemplary target-compound heat map is shown (left) that represents a set of 13 compounds active against six carbonic anhydrases. In this heat map, columns represent targets and rows compounds. Each cell is color-coded according to the pK_i value of the corresponding compound for the respective target applying a continuous color spectrum from red (lowest potency) via yellow to green (highest). Furthermore, targets are sorted by the number of activity cliffs they are involved in and compounds by the number of target cliffs they form (in descending order). This reordered heat map is displayed on the right. (c) Shown is a compound similarity map derived from the same matrix in (b). In this symmetrical matrix, both rows and columns indicate compounds that are sorted according to the number of their target cliffs. Each cell is gray-scaled on the basis of pairwise compound ECFP4 Tanimoto similarity applying a continuous spectrum from white (lowest similarity) to black (highest). The number of unique compound pairs that meet the similarity threshold of 0.55 and the number of unique activity cliffs are provided. In addition, compound pairs that form activity cliffs are highlighted in the lower part of the matrix.

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. Target-Compound Matrices[a]

number of			
target families	targets	matrices	compounds
1	1	232	19,660
	2	167	6493
	3	96	3823
	4	57	1405
	5	29	419
	6	13	64
	7	8	31
	8	1	2
	9	4	17
	10	6	28
	11	4	63
	12	2	23
2	2	11	244
	3	9	49
	4	6	101
	5	1	2
	7	1	2
	8	2	4
	13	1	2
	20	1	2
total		651	32,434

[a]For target-compound matrices containing at least two compounds, the number of matrices involving different numbers of targets and target families is reported. In addition, the number of corresponding compounds is given.

Table 2. Target Cliff Statistics[a]

			number of
multi-target matrices			419
compounds			12,774
selective compounds			3717
target cliffs	total		514
	target family	single	468
		multiple	46
	directionality	uni-	376
		bi-	138

[a]For 419 matrices containing multiple targets, the total number of compounds and of selective compounds with an at least 100-fold potency difference for at least one target over one or more others is reported. In addition, the number of unique target cliffs is given. Also, the number of target cliffs with uni- or bi-directionality and cliffs belonging to single or multiple families are provided.

different selectivity patterns are involved in the formation of activity cliffs, leading to the prioritization of compounds that display target selectivity and, in addition, are rich in SAR information.

MATERIALS AND METHODS

Compound Data. From the current release of ChEMBL[6] (release 14), compounds directly interacting (i.e., target relationship type "D") with human targets at the highest confidence level (i.e., confidence score 9) were assembled. Only compounds annotated with explicit equilibrium constants (K_i values) were considered in order to ensure a high level of data confidence and consistency. Approximate potency annotations such as ">", "<", or "~" were discarded. For compounds with multiple available K_i measurements against the same target, the geometric mean of all values was calculated as the final potency annotation.

Target-Compound Matrix. A data structure termed target-compound matrix was utilized to organize target-compound interactions in a systematic manner. For each compound, all reported targets were assembled as a complete target annotation. All unique target annotations consisting of one or more targets were determined. For each of these target annotations, the corresponding compounds were grouped. Then, a complete target-compound matrix was created for a given target annotation consisting of N targets and M compounds. Thus, in this matrix, each of the M compounds was active against all N targets, as illustrated in Figure 1a. Also, each set of the M compounds was unique to the corresponding given matrix. Thus, each active compound exclusively occurred in a one matrix. This type of target-compound matrix was previously also used to define activity profiles of compounds.[7] From the target-compound matrix, target cliffs and activity cliffs were obtained

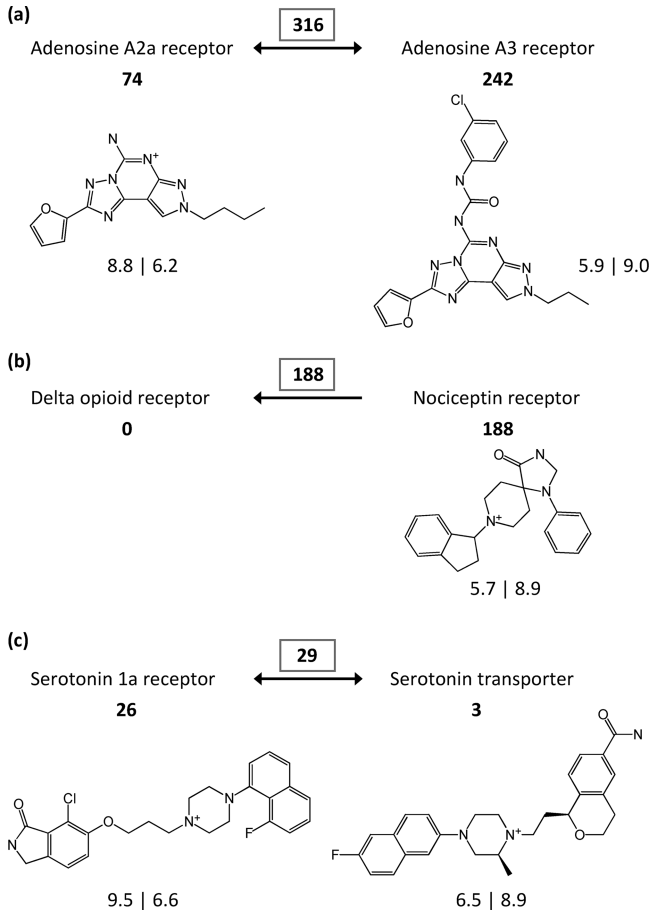


Figure 2. Exemplary target cliffs. Three different types of target cliffs are illustrated. (a) Bi-directional target cliff involving the adenosine A2a and A3 receptors. (b) Uni-directional target cliff involving the delta opioid and nociceptin receptors. (c) Bi-directional target cliff involving two targets from different (yet functionally related) families, i.e., the serotonin 1a receptor and the serotonin transporter. For each target cliff, the total number of selective compounds is reported in a gray rectangle. In addition, the number of compounds that are selective for individual targets is provided below each target. If available, a representative selective compound is shown. Potency (pK_i) values of a compound for two targets forming a cliff are also given (separated by a vertical line).

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

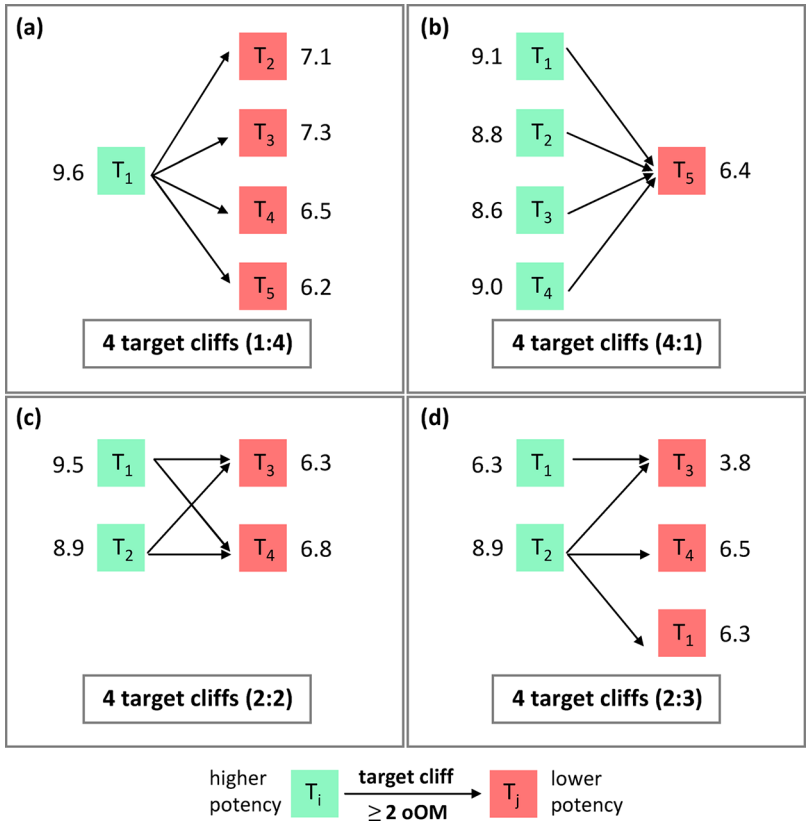


Figure 3. Target cliff patterns. For a compound involved in multiple target cliffs, three alternative patterns are illustrated. (a) The compound is selective for one target over several others or (b) for several targets over a single target. In addition, in (c) and (d), patterns are illustrated where a compound is selective for targets in multiple and in part overlapping pairs. In each case, targets for which the compound has high potency are shown on a green background, while targets for which the compound has low potency are shown on a red background (for illustration, hypothetical potency values are used). For a given compound, a target cliff pattern combines the number of targets against which the compound has high or low potency. For example, "1:4" indicates that the compound is highly potent against one and lowly potent against four other targets.

by traversing the matrix in different directions, as illustrated in Figure 1a.

Target Cliff Criterion. For the formation of a target cliff, a potency difference of at least 2 orders of magnitude of a compound against a pair of targets was required. Therefore, for a matrix containing at least two targets (i.e., $N \geq 2$), there were $(N \times (N-1))/2$ possible target cliffs (Figure 1a).

Target Similarity. Pairs of targets forming cliffs were evaluated in the presence or absence of a target similarity criterion. For our current analysis, family membership was used as a similarity criterion. Hence, two targets were considered similar if they belonged to the same family. Therefore, targets of all qualifying compounds were classified into target families following the protein classification hierarchy of ChEMBL and the family annotations of UniProt.[8] Only target families that consisted of at least four targets were further considered.

Activity Cliff Criteria. As the similarity criterion for activity cliff formation, a pair of compounds had to yield a Tanimoto coefficient (Tc)[9] value of at least 0.55 using the extended connectivity fingerprint with bond diameter of four (ECFP4)[10] as a molecular representation, which is often applied to study activity cliffs.[5] In addition, as the potency difference criterion, two structurally similar compounds were required to differ in their potency against a given target by at least 2 orders of magnitude. Therefore, for a matrix containing at least two compounds (i.e., $M \geq 2$), there were $(M \times (M-1))/2$ possible activity cliffs (Figure 1a).

Table 3. Prioritized Target Cliff Patterns[a]

target cliff pattern	# selective compounds
1:1	2225
1:2	679
1:3	197
1:4	20
2:1	256
2:2	88
2:3	23
2:4	6
3:1	77
3:2	12
3:3	12
4:1	9
4:2	9
4:5	8
10:2	5

[a]Fifteen most frequently occurring target cliff patterns represented by at least five selective compounds are shown. These target cliff patterns involve different numbers of high- and low-potency targets. The number of selective compounds representing each pattern is reported. The complete list of target cliff patterns is provided in Table S1 of the Supporting Information.

Heat Map. Furthermore, targets were ordered by the number of activity cliffs they were involved in and compounds by the number of target cliffs they formed (in descending order).

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

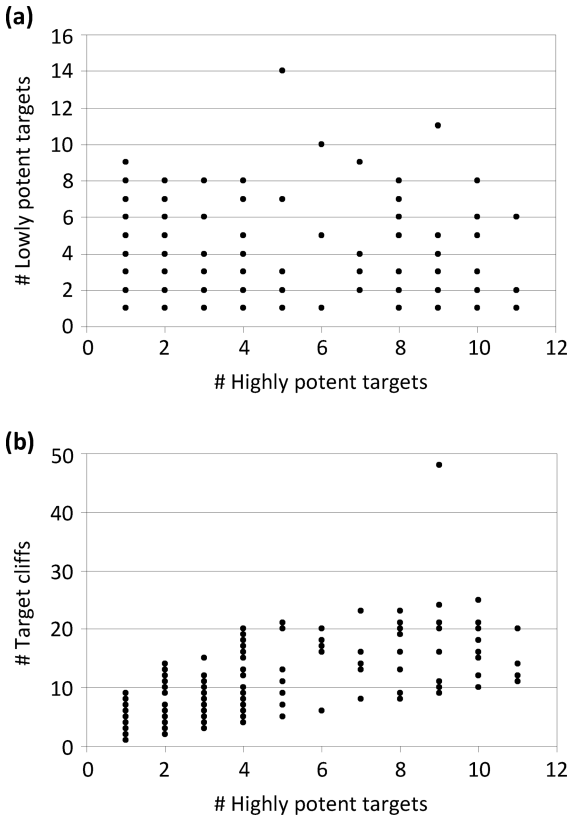


Figure 4. Observed target cliff patterns. (a) For target cliff patterns isolated from ChEMBL, the distribution of targets against which individual compounds are highly and lowly potent is shown. Each dot represents a target cliff pattern. (b) For selective compounds, the distribution of the number of target cliffs they are involved in is shown as a function of the number of targets against which they are highly potent. Here, each dot represents a selective compound.

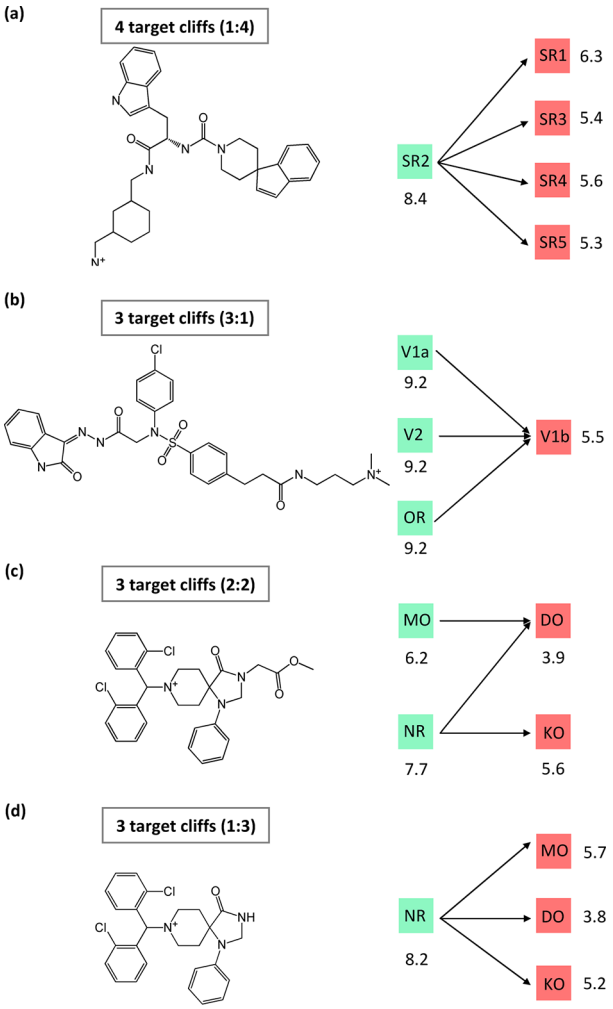


Figure 5. Compounds with different target cliff patterns. In (a)−(d), four exemplary compounds forming different target cliff patterns are shown. For each compound, the number of target cliffs and the corresponding pattern are reported in a gray box. The representation is according to Figure 3. Target abbreviations: SR, somatostatin receptor; V1a, vasopressin V1a receptor; V1b, vasopressin V1b receptor; V2, vasopressin V2 receptor; OR, oxytocin receptor; MO, mu opioid receptor; DO, delta opioid receptor; KO, kappa opioid receptor; and NR, nociceptin receptor.

In addition, cells were color-coded according to compound potency. The target-compound matrix was rearranged accordingly, yielding an ordered heat map, as illustrated in Figure 1b.

Similarity Matrix. In addition, to complement a heat map with molecular similarity information, a compound similarity map reporting pairwise similarity values calculated with ECFP4 was generated (Figure 1c).

RESULTS AND DISCUSSION

Compound Activity Data and Target-Compound Matrices. On the basis of the stringent compound selection criteria we applied, 32,968 bioactive compounds active against 416 human targets were assembled that yielded at total of 58,463 K_i measurements. These 416 targets belonged to 29 different families. In addition, 1185 target annotations were identified that produced unique compound sets. A total of 543 of these annotations (~45%) were only associated with single compounds and not further considered. The remaining 651 target annotations were associated with multiple compounds and subjected to further analysis. These qualifying annotations yielded 651 unique target-compound matrices, which were further classified by the number of their targets and families the targets belonged to, as reported in Table 1. A total of 232 (~35.6%), 387 (~59.4%), and 32 (~4.9%) matrices were found to contain single targets, 2−12 targets from the same family, and 2−20 targets from two different families, respectively. Accordingly, the majority of the matrices involved compounds active against multiple targets belonging to the same family. Compound activity across different target families was rarely observed.

Target Cliff Concept. The introduction of target cliffs was motivated by our evaluation of compound selectivity behavior across multiple targets. We define a target cliff as a pair of targets against which an individual compound has significantly different potency. Thus, target cliffs represent compound-based pairwise target relationships, with a focus on differential binding affinity. Because compounds are often active against multiple targets, target cliffs not only represent compound-based target relationships but also selectivity characteristics of individual compounds across different targets. Accordingly, target cliffs can be analyzed to describe, organize, and compare compound selectivity patterns and further prioritize compounds with desired selectivity profiles. This represents another key aspect of the target cliff concept, as introduced herein. It should be noted that our general target cliff definition does not include a target

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

similarity criterion. Hence, all possible combinations of targets are a priori considered, regardless of their degree of relatedness. This represents a departure from the activity cliff concept where compounds forming a cliff must by definition be structurally similar. In analogy, if one would like to limit the analysis to closely related targets, which might often be meaningful, a target similarity measure can be applied (for which many different types of measure can be considered). Herein, we have used shared family membership as a target similarity criterion, which ensures that targets forming cliffs are related to each other. Applying a target similarity measure might in principle be expected to limit the number of target cliffs for further consideration because compounds having high potency against a given target might have similar potency against closely related ones, but might be weakly potent or inactive against distantly or unrelated targets. On the other hand, as further discussed below, we have found that the majority of unique target cliffs we identified exclusively involved targets from the same family.

Activity Cliffs. From 651 matrices containing at least two compounds, activity cliffs were systematically extracted. A total of 259,468 compound pairs met the structural similarity criterion (i.e., Tc values of ECFP4 $\geq$ 0.55). However, only of these 16,848 pairs (~6.5%) originating from 307 matrices also met the potency difference criterion and qualified as activity cliffs. These activity cliffs involved 9944 compounds, i.e., ~31% of all compounds that were analyzed. Most activity cliffs were single-target cliffs; only 1383 multi-target cliffs were identified. These multi-target cliffs only involved targets from the same families. These findings obtained on the basis of target-compound matrices were consistent with those of general activity cliff surveys.[11,12]

Target Cliff Statistics. In 419 matrices containing multiple targets, target cliffs were identified, as reported in Table 2. A total of 3717 compounds showed at least a 100-fold difference in potency against at least one pair of targets and were thus classified as selective. These compounds formed a total of 514 target cliffs involving one or more selective compounds (i.e., representing unique target combinations). Most of these target cliffs (i.e., 468) were formed by pairs of targets belonging to the same family, i.e., closely related targets. The remaining 46 target cliffs were formed by targets from different families. It should be noted that compounds representing the same target cliff might display selectivity for one or the other target forming a pair (i.e., selectivity for target A over target B or for B over A). These relationships defined the directionality of a target cliff, which was further analyzed. The majority of the target cliffs (i.e., 376) were unidirectional, indicating that compounds were consistently selective for one target of the pair over the other. The remaining 138 target cliffs were bi-directional, capturing inverse selectivity relationships. Figure 2 shows three different categories of target cliffs we observed including uni-directional cliffs as well as bi-directional cliffs formed by targets from the same or two different (yet functionally related) families.

Target Cliff Patterns. The simplest form of observed compound selectivity, i.e., the compound is only active against two targets and selective for A over B, leads to the formation of a single target cliff. However, more complex selectivity relationships lead to multiple target cliffs. Among 3717 selective compounds we detected, 1492 compounds (~40%) were found to be involved in multiple target cliffs. Different target cliff patterns represented various types of selectivity relationships, as illustrated in Figure 3. These patterns characterized compounds selective for one target over several others, for several targets over a single target, or for multiple targets over several other distinct or overlapping targets. As a possible representation of these target cliff patterns, a compound was assigned the number of targets against which it displayed high or low potency. For example, "1:4" indicates that the compound was highly potent against one and lowly potent against four other targets (Figure 3a). In addition, as illustrated in Figure 3a–d, compounds that were involved in the same number of target cliffs displayed different selectivity patterns, which reflected distinct selectivity relationships. Thus, this numerical representation scheme accounts for target cliff patterns and helps to organize and differentiate increasingly complex selectivity patterns in a systematic manner.

Using this representation scheme, a total of 64 unique target cliff patterns were identified. Their distribution including the simplest 1:1 pattern is reported in Figure 4a. Many target cliff patterns were sparsely populated. However, there also were a number of target cliff patterns involving many high- and/or low-potency targets. The distribution of target cliffs relative to the number of targets against which the corresponding compounds were highly potent is shown in Figure 4b. No apparent correlation was observed, but compounds that were highly potent against multiple targets were generally also involved in multiple target cliffs. In addition, the distribution of selective compounds over target cliff patterns was analyzed. Table 3 reports 15 most frequently occurring target cliff patterns represented by at least five selective compounds. In addition to the simplest 1:1 pattern, a large number of compounds formed patterns 1:2, 1:3, 2:1, 2:2, and 3:1. The complete list of target cliff patterns is provided in Table S1 of the Supporting Information. Four exemplary compounds representing prevalent target cliff patterns are shown in Figure 5. For example, Figure 5c and d show two selective compounds from the same matrix that

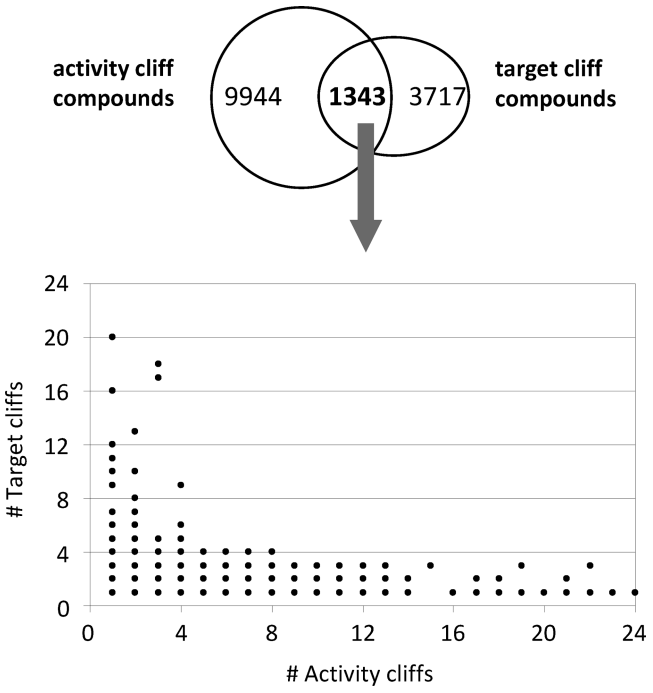


Figure 6. Target cliffs versus activity cliffs. The subset of 9944 activity cliff compounds is compared to 3717 target cliff compounds. A total of 1343 compounds are common to both subsets. For these target/activity cliff compounds, the distribution of the number of target cliffs they are involved in is shown as a function of the number of activity cliffs they form. Each dot represents a target/activity cliff compound.

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX–XXX

contained four short peptide G-protein coupled receptors (GPCRs). These two compounds differed by a methyl ester group and were involved in three target cliffs. However, different target cliff patterns emerged. The compound shown in Figure 5d (1:3) had more clearly defined selectivity for the nociceptin receptor over three opioid receptors than the compound in Figure 5c (2:2). Thus, for compounds sharing activity against the same targets, it was possible to detect and further distinguish different selectivity patterns.

Target Cliffs versus Activity Cliffs. Although target cliffs and activity cliffs conceptually differ, they can be considered in a complementary manner. We compared the subsets of 3717 compounds involved in target cliffs and 9944 activity cliff-forming compounds we isolated from systematically derived target-compound matrices. A total of 1343 compounds were found to be present in both subsets. For these compounds, the distribution of the number of target cliffs they were involved in is shown in Figure 6 relative to the number of activity cliffs they formed. For the majority of these compounds, more activity cliffs than target cliffs were observed. Interestingly, the number of activity cliffs and target cliffs the selective compounds were involved in was often inversely correlated. Exemplary compounds of different structural complexity and the associated target and activity cliffs are shown in Figures 7 and 8. In Figure 7a,

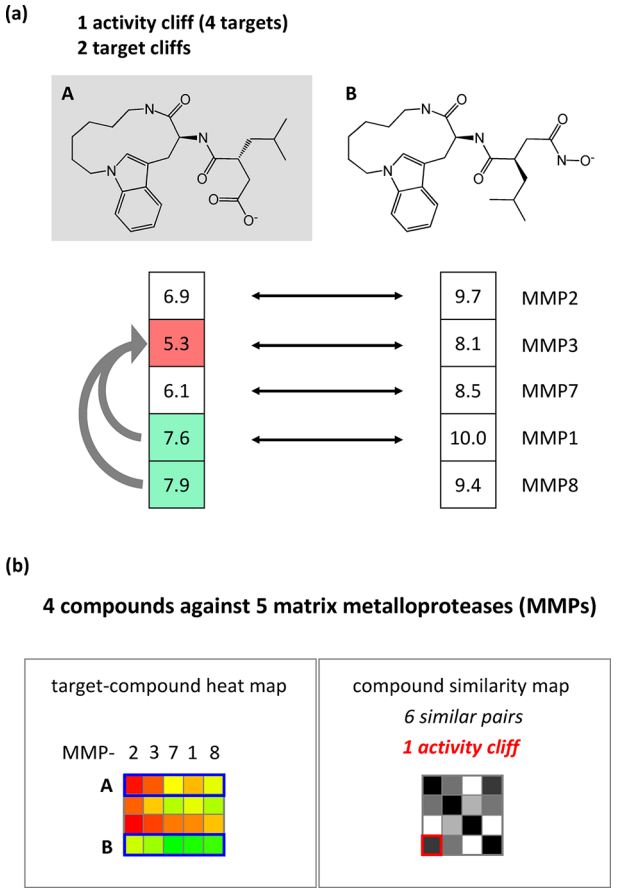


Figure 7. Representative target/activity cliff compounds for MMPs. In (a), an exemplary dual-cliff compound (A) is shown (on a gray background). An activity cliff partner (B) is shown on the right. Both compounds are active against five MMPs. For each compound, the potency profile is given (consisting of potency values for corresponding targets) and color-coded according to Figure 3. Targets involved in the formation of activity cliffs are indicated by black horizontal arrows. Target cliffs are represented by gray arrows that are directed from the high- to the low-potency target. Target abbreviation: MMP, matrix metalloprotease. In (b), the ordered heat map for the same set of targets and compounds A and B (highlighted in blue) is shown on the left. In addition, the corresponding compound similarity map is shown on the right. The representation is according to Figure 1b and c.

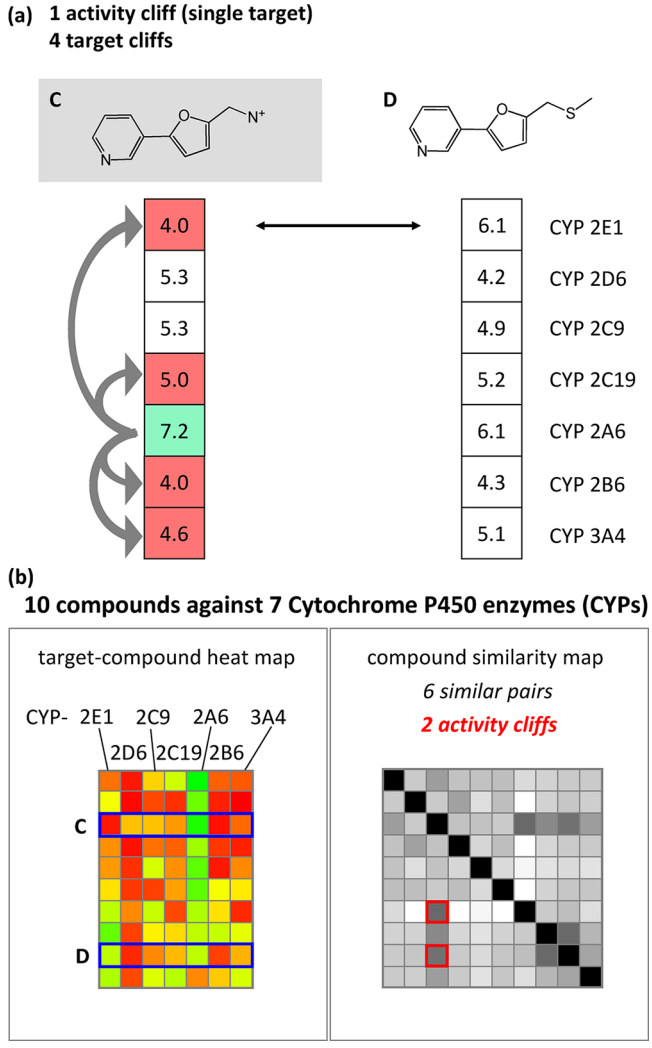


Figure 8. Representative target/activity cliff compounds for CYPs. In (a), one exemplary dual-cliff compound (C) is shown (on a gray background). An activity cliff partner (D) is shown on the right. Both compounds are active against seven cytochrome P450 enzymes. The representation is according to Figure 7a. Target abbreviation: CYP, cytochrome P450. In (b), the corresponding ordered heat map and compound similarity map are shown.

two structurally related compounds (A and B) active against five matrix metalloproteases (MMPs) formed activity cliffs for four of these targets. However, only one of these compounds was involved in the formation of target cliffs. It was found to be selective for MMP1 and MMP8 over MMP3, yielding two target cliffs with a 2:1 cliff pattern. The corresponding heat map representing the complete matrix and the compound similarity map are shown in Figure 7b. Compound A and B formed the only activity cliff in this matrix. In addition, Figure 8a shows two structurally similar compounds (C and D) that were active against seven cytochrome P450 enzymes. This compound pair formed a single activity cliff for one of these targets. However, one of these two compounds was also involved in four target cliffs

dx.doi.org/10.1021/ci300602m | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(1:4), displaying selectivity for cytochrome P450 2A6 over the other four closely related targets. The corresponding heat map and the compound similarity map are shown in Figure 8b. Thus, compounds that were involved in the formation of both types of cliffs revealed different selectivity patterns, based on their involvement in varying numbers of target cliffs, as well as SAR information.

CONCLUSIONS

Herein, we have introduced and explored the concept of target cliffs. The target cliff concept can be applied to establish compound-based target relationships and, in addition, analyze and represent complex compound selectivity patterns involving multiple targets. Target cliffs are derived from target-compound matrices. In addition, the target-compound matrix data structure also provides direct access to activity cliffs. Using ChEMBL compounds with high-confidence potency measurements and target annotations, target cliffs were systematically identified and organized. Approximately 3700 compounds were found to be involved in the formation of target cliffs. A total of more than 500 unique target cliffs with different frequency of occurrence were identified, which provide a large knowledge base for the analysis of complex compound selectivity relationships. Nearly 1500 of the 3700 target cliff compounds participated in multiple target cliffs that represented selectivity patterns of different complexity. A numerical representation scheme was applied to systematically organize and differentiate increasingly complex selectivity patterns. More than 1300 target cliff compounds were also found to form varying numbers of activity cliffs. Albeit conceptually distinct, target cliffs and activity cliffs can be considered in a complementary manner, with a simultaneous focus on selectivity and SAR information. Given the dual role of target cliffs as indicators of compound-based target relationships and signatures of compound selectivity patterns, they combine target and ligand selectivity information and can be studied from different points of view. From a target perspective, relationships are established on the basis of shared ligands with differential activity, reflecting differences in the way targets respond to active compounds. This information might be utilized, for example, in the design of target differentiation networks. From a compound-centric perspective, target cliffs are particularly attractive for the identification and representation of multi-target selectivity relationships, which are in general difficult to capture. Compound selectivity patterns established on the basis of target cliffs can be directly compared, enable detailed selectivity analysis, and might reveal selectivity determinants for closely related compounds. Thus, taken together, the target cliff concept provides an advanced basis for the study of molecular selectivity relationships across multiple targets.

ASSOCIATED CONTENT

Supporting Information

Table S1 reports the complete list of target cliff patterns. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

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

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Wermuth, C. G. *The Practice of Medicinal Chemistry*, 3rd ed.; Academic Press: Burlington, MA, 2008.
(2) Castanero, A. B.; Eggert, U. S. Small molecular probes of cellular pathways and networks. *ACS Chem. Biol.* **2011**, *6*, 86−94.
(3) Paolini, G. V.; Shapland, R. H. B.; van Hoorn, W. P.; Mason, J. S.; Hopkins, A. L. Global mapping of pharmacological space. *Nat. Biotechnol.* **2006**, *24*, 805−815.
(4) Maggiora, G. M. On outliers and activity cliffs − Why QSAR often disappoints. *J. Chem. Inf. Model.* **2006**, *46*, 1535−1535.
(5) Stumpfe, D.; Bajorath, J. Exploring activity cliffs in medicinal chemistry. *J. Med. Chem.* **2012**, *55*, 2932−2942.
(6) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−D1107.
(7) Hu, Y.; Bajorath, J. Chemical transformations that yield compounds with distinct activity profiles. *ACS Med. Chem. Lett.* **2011**, *2*, 523−527.
(8) UniProtConsortium.. Reorganizing the protein space at the universal protein resource (UniProt). *Nucleic Acids Res.* **2012**, *40*, D142−D148.
(9) Willett, P. Searching techniques for databases of two- and three-dimensional structures. *J. Med. Chem.* **2005**, *48*, 4183−4199.
(10) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742−754.
(11) Stumpfe, D.; Bajorath, J. Frequency of occurrence and potency range distribution of activity cliffs in bioactive compounds. *J. Chem. Inf. Model.* **2012**, *52*, 2348−2353.
(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.