MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **80** (2018) 283-294

ISSN 0340 - 6253

Activity Cliffs and Structural Cliffs for Categorical Responses

Roberto Todeschini[1], Cecile Valsecchi[2]

[1]*Milano Chemometrics & QSAR Research Group, University of Milano-Bicocca, Dept. of Earth and Environmental Sciences, 20126, Milano, Italy*
roberto.todeschini@unimib.it

[2]*IRCCS - Istituto di Ricerche Farmacologiche "Mario Negri", Laboratory of Environmental Chemistry and Toxicology, Via La Masa 19, 20156 Milano, Italy*

(Received February 8, 2018)

Abstract

In the last ten years, the concept of activity cliff gained a lot of interest by medicinal chemists and QSAR modellers. The reason of this interest is its capacity to highlight problematic as well as interesting situations in analysing the activity/property of pairs of compounds, thus focusing the research on the reasons why two similar compounds show different activities. In the literature, activity cliffs as well as their dual aspect, i.e. structural cliffs, are calculated from activities represented by a continuous response.
In this paper, the proposal is to estimate the activity cliffs of categorical responses, i.e. for compounds partitioned into different classes.

1 Introduction

The Activity Cliffs (*AC*) have been defined as the ratios of the difference in activity of two compounds to their dissimilarity in a given chemical space, i.e. pairs of structurally similar compounds having a significant difference in their activities [1]. Structure-Activity Relationships (SAR) information is directly related to activity cliffs, they being centres of discontinuity in activity landscapes of compound data sets and focal points in SAR analysis [2].

The first quantitative expression to calculate activity cliffs was proposed by Guha and van Drie in 2008 [3], defined as:

$$AC_{st} = \frac{|A_s - A_t|}{1.01 - S_{st}} \qquad 0 \leq S_{st} \leq 1 \tag{1}$$

where A are the activities of the compounds s and t and S_{st} their similarity.

The definition of activity cliffs leaves two problems open: (a) the selection of a similarity measure and (b) the criterion defining a meaningful activity difference [4]. The distribution of activity cliffs often misses of invariance, being significantly dependent on the molecular representations and similarity measures used [1, 5, 6]. The similarity between two molecules can be quantified by comparing their binary fingerprints using a similarity metric such as the Tanimoto coefficient or other binary similarity measures [3, 7, 8]. In other cases, the similarity is calculated by using continuous molecular descriptors through distance measures such as Euclidean or Manhattan distances [8, 9]. Anyway, the availability of many different molecular representations adds further source of variance to the results [10].

In several cases, for the activity cliffs exploration in medicinal chemistry, intuitive measures of inter-compound distance based on shared sub-structures were preferred as similarity criteria [11-14].

Furthermore, the considered activity measurement and the magnitude of activity difference influence the activity cliff identification [10]. As a heuristic criterion, assay-independent measurement and at least 100-fold difference in potency are generally preferred [4, 10].

Structural cliffs (*SC*) are the dual concept of activity cliffs. They highlight the pairs of compounds that are dissimilar molecular structures but show similar activity. The presence of such cases can be related to two different problems: (a) the need to modify the molecule description because inadequate or redundant, i.e. change the selected molecular descriptors, and/or (b) the underlying models need to account also for non-linearity [1]. The term “structural cliff” was suggested some years ago by our group to Jerry Maggiora (private communication) and used by him replacing the term “scaffold hopping”.

The interest for the activity cliffs is highlighted by the number of papers dealing with this topic in the last 10 years. From the first concerns about the lack of neighborhood invariance in QSAR modelling [1], several studies were later performed by both medicinal chemists and QSAR modellers to quantitatively analyse the nature of Structure-Activity Relationships [2, 15] and, in particular, of *AC*s [3, 16-20].

Studies were performed to provide graphical representations of the activity cliffs, such as Structural-Activity Maps, Network-like Similarity Graphs, and Activity Cliff Networks, which better point out the SAR topology and/or the activity cliffs features [21-24].

The activity landscape has been largely explored, also with the aid of geostatistical techniques [9,25,26], and its main attributes, such as cliff types, have been well characterized [15, 27].

The discontinuities highlighted by the activity cliffs can provide useful information for medicinal chemists in drug discovery, but on the other side, they represent anomalies, which may affect computational methods [28].

In order to overcome the lack of invariance of the chemical space from the chosen molecular descriptors, consensus strategies, which consider different structural representations and similarity values, were applied [6, 29].

Other studies are focused on multiple target activity, extending the *AC*s concept to pairs of structural analogues that are active against different numbers of targets, highlighting the existence of target promiscuity [21, 30, 31]. Another source of SAR information was provided considering also inactive compounds [28, 32].

Furthermore, global attempts were made to find common patterns in *AC*s from which derive the concepts of “activity ridges” and “coordinated activity cliffs” [8, 12, 33, 34]. Recently, studies on activity cliffs have ranged from molecular modelling, such as docking and virtual ligand screening [29, 35], to machine learning models, like Support Vector Machine and Deep Neural Networks [8, 36, 37]. These machine learning approaches were employed to predict which molecules are most likely activity cliffs generators and to smooth the activity landscape by transforming input descriptors into higher dimensional spaces [8, 37]. Undoubtedly, *AC*s add useful information to the understanding of SARs especially when related to their dual complementary, i.e. structural cliffs, even though there is still evidence of under-utilization of ACs in practice of medicinal chemistry [10, 17,20].

Disregarding the different approaches to activity cliffs, all the studies have been focused on continuous activity values and only few of them considered explicitly structural cliffs in the ACs assessment [27].

In this paper, we simply tried to extend the concept of activity and structural cliffs from continuous responses (activities, properties) to categorical responses, where the compounds are partitioned into two or more classes. Thus, activity and structural cliffs are here related to classification models instead of regression models. Classification models are actually largely used in QSAR (perhaps more than regression models) and the evaluation of activity and structural cliffs should be very useful in checking the class partition of the compounds as well

as the model performances. It can be noted that the uncertainties of the experimental measures might be reduced by a clusterization of the compounds.

2 Theory

Given *G* classes, a comparison between two objects *s* and *t* allows only the two following cases:

$$\Delta_{st} = \begin{cases} 1 & if\ C_s \neq C_t \\ 0 & if\ C_s = C_t \end{cases} \tag{2}$$

where C_s and C_t indicate the class of the objects *s* and *t*, respectively, and $\Delta_{st} = 1$ indicates that the two compounds belong to different classes.

Given a similarity measure S_{st} between the two objects *s* and *t* in the range [0, 1], the measures for activity and structural cliffs can be defined as the following:

$$AC_{st} = \Delta_{st} \cdot \left[e^{S_{st}} - 1 \right] \qquad 0 \le AC_{st} \le 1.718 \tag{3}$$

$$SC_{st} = \left(1 - \Delta_{st}\right) \cdot \left[e^{(1-S_{st})} - 1 \right] \qquad 0 \le SC_{st} \le 1.718 \tag{4}$$

The two quantities can be also straightforwardly scaled between [0, 1] as:

$$AC^*_{st} = \frac{AC_{st}}{1.718} \qquad 0 \le AC^*_{st} \le 1 \tag{5}$$

$$SC^*_{st} = \frac{SC_{st}}{1.718} \qquad 0 \le SC^*_{st} \le 1 \tag{6}$$

The unscaled definition will be used from now on.

In the case where $\Delta_{st} = 0$, i.e. when the classes of the two objects coincide, the activity cliff between *s* and *t* is always zero, i.e.

$$\Delta_{st} = 0 \quad \Rightarrow \quad AC_{st} = 0$$

In this case, the structural cliff only depends on the degree of similarity between the two objects. Considering the two extreme cases, if $S_{st} = 1$ the structural cliff is equal to zero, i.e.

$$SC_{st} = \left(1 - \Delta_{st}\right) \cdot \left[e^{(1-S_{st})} - 1 \right] = (1-0) \cdot [1-1] = 0$$

while if $S_{st} = 0$ the structural cliff reaches its maximum value equal to 1.718, i.e.

$$SC_{st} = (1-\Delta_{st}) \cdot \left[e^{(1-S_{st})} - 1 \right] = (1-0) \cdot [2.718-1] = 1 \cdot 1.718 = 1.718$$

In the case where $\Delta_{st} = 1$, i.e. when the classes of the two objects not coincide, the structural cliff between *s* and *t* is always zero, i.e.

$$\Delta_{st} = 1 \quad \Rightarrow \quad SC_{st} = 0$$

In this case, the activity cliff only depends on the degree of similarity between the two objects. Considering the two extreme cases, if $S_{st} = 0$ the activity cliff is equal to zero, i.e.

$$AC_{st} = \Delta_{st} \cdot \left[e^{S_{st}} - 1 \right] = 1 \cdot [1-1] = 0$$

while if $S_{st} = 1$ the activity cliff reaches its maximum value equal to 1.718, i.e.

$$AC_{st} = \Delta_{st} \cdot \left[e^{S_{st}} - 1 \right] = 1 \cdot [2.718-1] = 1.718$$

In Table 1, values of activity and structural cliffs are reported for some similarity values for the two cases of Δ_{st}.

Table 1. *AC* and *SC* unscaled values for different similarity values and the two cases of Δ_{st}.

similarity	$\Delta_{st} = 0$		$\Delta_{st} = 1$	
	AC	*SC*	*AC*	*SC*
0	0	1.718	0.000	0
0.1	0	1.460	0.105	0
0.2	0	1.226	0.221	0
0.3	0	1.014	0.350	0
0.4	0	0.822	0.492	0
0.5	0	0.649	0.649	0
0.6	0	0.492	0.822	0
0.7	0	0.350	1.014	0
0.8	0	0.221	1.226	0
0.9	0	0.105	1.460	0
1	0	0.000	1.718	0

2.1 Extension to ordered classes

In several cases, the classes are obtained by the discretization of a continuous response. For example, given toxicity measures of a set of compounds, they can be distributed into *G* classes ($G > 2$) based on some arbitrarily selected $G - 1$ thresholds. This provides ordered classes, such as, for example, non-toxic (C_1), toxic (C_2), very toxic (C_3).

In this case, a pair of compounds *s* and *t* belonging to adjacent classes are less different from two compounds one belonging to C_1 and the other to C_3.

Thus, the numerical difference between the classes of the two objects *s* and *t* is meaningful and can be defined as:

$$\delta_{st} = |C_s - C_t| \tag{7}$$

The quantity Δ_{st} is obtained by the following simple expression:

$$\Delta_{st} = \frac{\delta_{st}}{G-1} \tag{8}$$

The activity and structural cliffs are then obtained by the same formulas (3) and (4), respectively.

For example, according to formulas (7) and (8), for a 3-class case (G = 3), $\delta_{st} = \{0, 1, 2\}$ and $\Delta_t = \{0, 0.50, 1\}$ and for a 4-class case (G = 4), $\delta_{st} = \{0, 1, 2, 3\}$ and $\Delta_t = \{0, 0.33, 0.67, 1\}$.

For the cases where $\Delta_{st} = 0$ or $\Delta_{st} = 1$, activity and structural cliffs assume the same values already reported in Table 1. The values of activity and structural cliffs for the cases $\Delta_{st} = \{0.33, 0.50, 0.67\}$ are collected in Table 2, for different similarity values.

Table 2. *AC* and *SC* unscaled values for different values of similarities and different values of Δ_{st}.

Similarity	$\Delta_{st} = 0.33$		$\Delta_{st} = 0.50$		$\Delta_{st} = 0.67$	
	AC	*SC*	*AC*	*SC*	*AC*	*SC*
0	0	1.146	0	0.859	0	0.572
0.1	0.035	0.974	0.053	0.730	0.070	0.486
0.2	0.074	0.817	0.111	0.613	0.148	0.408
0.3	0.117	0.676	0.175	0.507	0.233	0.338
0.4	0.164	0.548	0.246	0.411	0.328	0.274
0.5	0.216	0.433	0.324	0.324	0.433	0.216
0.6	0.274	0.328	0.411	0.246	0.548	0.164
0.7	0.338	0.233	0.507	0.175	0.676	0.117
0.8	0.408	0.148	0.613	0.111	0.817	0.074
0.9	0.486	0.070	0.730	0.053	0.974	0.035
1	0.572	0.000	0.859	0.000	1.146	0.000

3 Simulated example

A simulated data set of 15 objects (Table 3) was prepared as the following. The objects are described by 2 variables (x, y) and were generated in such a way that objects are partitioned into 3 clusters, each constituted by 5 objects; cluster 1 is constituted by objects {1 – 5}, cluster

2 by {6 – 10}, and cluster 3 by {11 – 15}. The ordered classes (column *A*) were assigned as the following: for each cluster, the last two objects are assigned to the other two classes; for example, the first 5 objects belonging to cluster 1 are assigned to the classes as {1 1 1 2 3}. The chosen class partition allows a quick detection of activity/structural cliffs of the pairs.

Table 3. The simulated data set of 15 objects distributed into 3 classes (A) and 2 classes (B).

id	*x*	*y*	*A*	*B*
1	1	2	1	1
2	3	1	1	1
3	1.5	4	1	1
4	2.5	3.5	2	1
5	4	1	3	1
6	11	13	2	1
7	12	14	2	1
8	13	12	2	1
9	15	14	1	2
10	14	12	3	2
11	21	25	3	2
12	24	22	3	2
13	23	23	3	2
14	22	23	1	2
15	25	23	2	2

The similarity matrix between all the pairs of objects was calculated as the following: (1) the two variables *x* and *y* are range scaled, (2) the average Euclidean distance matrix **D** is calculated, (3) the similarity matrix **S** is obtained as $\mathbf{S}=1-\mathbf{D}$.

The obtained maximum similarity between pairs of different objects is 0.979, while the minimum similarity is 0.336. The *AC* and *SC* values of the 105 pairs are shown in Figure 1.

All the objects belonging to the same class are located on the horizontal axis in correspondence to an activity cliff equal to zero by definition (equations (2) and (3)) and to the corresponding values of structural cliffs (equation (4)). On the right side of the structural cliff axis (set A in Figure 1) are located the objects that, although belonging to the same class, show a low similarity. This happens for all the pairs of objects belonging to the same class but located in different clusters. For example, the pairs 5 – 11, 5 – 12, 5 – 13, 1 – 14, 2 – 14 and 3 – 14, which belong to opposite clusters, i.e. cluster 1 and cluster 3, with *SC* values around 0.85. An analogous behaviour of the structural cliffs is also observed, but in a lower measure, for the

pairs belonging to the same class but located in the adjacent clusters $C_1 – C_2$ and $C_2 – C_3$. For example, the pairs 1 – 9, 10 – 11, 10 – 13, and so on.

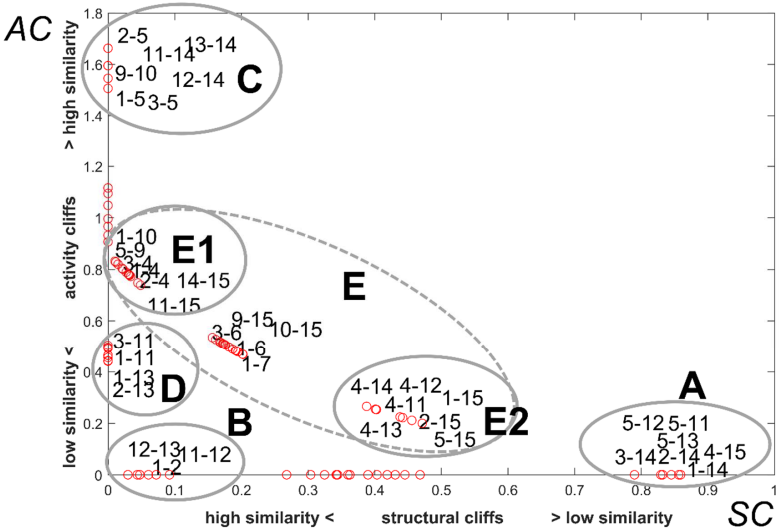

Figure 1. The scatter plot of all the 105 pairs of the 15 objects of the data shown in Table 3, classified according to column *A* (unscaled results). Some pairs are manually shifted from the bulk of pairs and others are deleted. The meaning of the sets A – E are discussed in the text.

All the pairs belonging to the same class and to the same cluster are located on the left side of the horizontal axis (set B in Figure 1), i.e. these pairs belong to the same class and are very similar (1 – 2; 6 – 7, 11 – 12, etc.). These pairs do not have neither structural cliffs nor activity cliffs.

All the objects located on the vertical axis having structural cliffs equal to zero (by definition, by expressions (2) and (4)) are pairs of objects belonging to different extreme classes (in this example, belonging to class 1 and 3, $C_1 – C_3$). High values of activity cliffs (set C in Figure 1) highlight pairs having high similarity (they belong to the same cluster) although located in different extreme classes. These are the pairs 1 – 5, 2 – 5, 3 – 5, 11 – 14, 12 – 14, 13 – 14, with *AC* values around 1.6. A second set of activity cliffs, with *AC* values between [1.0 – 1.2] are the pairs still belonging to different extreme classes but located in adjacent clusters (1 – 10, 2 – 10, 3 – 10, 1 – 11, 2 – 11, and so on). Low values of the activity cliffs (set D in Figure 1) are obtained for pairs belonging to different extreme classes but also located in very different clusters, namely clusters 1 and 3, such as, for example, 2 – 13.

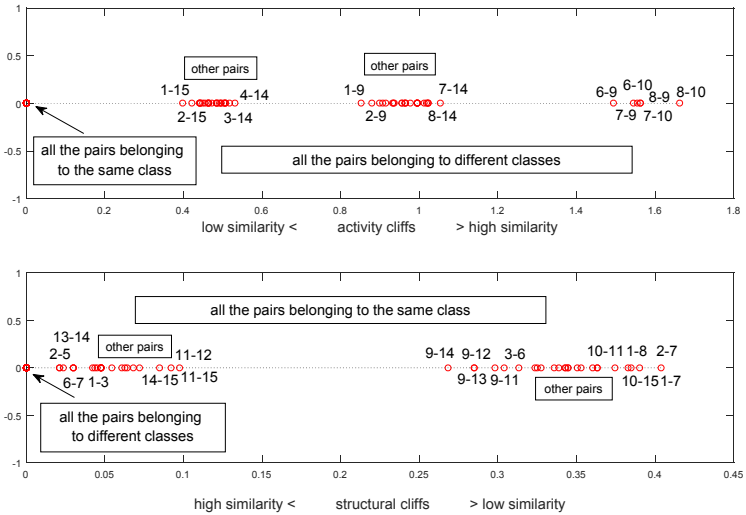

Figure 2. Plots of the activity cliffs (on the top) and structural cliffs (on the bottom) from data of Table 1 for a 2-class case, i.e. using a class partition as defined by column *B* (unscaled results).

Pairs of objects belonging to different but adjacent classes ($C_1 - C_2$ and $C_2 - C_3$) show values greater than zero both for *AC* and *SC* (set E in Figure 1). In particular, the pairs belonging to the subset E1 are those pairs belonging to adjacent clusters presenting quite high similarities, while the pairs belonging to the subset E2 are those presenting quite low similarities.

If the class partition is performed using column *B* of Table 1, i.e. partitioning the data in two classes, the scatter plot of the pairs (Figure 2) is constituted only by pairs projected along the two axes (one for activity and one for structural cliffs).

It can be easily noted that the pairs on the right side of the activity cliff plot are the pairs having high similarity, i.e. objects belonging to the same cluster but assigned to different classes (e.g. 8 – 10); these pairs show quite high activity cliff values (~ 1.6) . In the second plot, structural cliffs, on the right side are located all the pairs belonging to the same class but to different clusters (e.g. 2 – 7), thus having low similarity.

4 Conclusions

In this paper the classical approach to activity/structural cliffs based on a continuous property is modified to deal with categorical responses, restricted to ordered classes when the classes are

more than two. The kind of information that can be extracted by the activity/structural cliffs approach here proposed is closely similar to that of the classical approach based on continuous responses and then it can be extended to the literature applications.

In multi-class cases, but when the classes are not ordered, the same approach could be applied comparing separately all the pairs of classes.

Finally, the function (4) for the calculation of the structural cliffs can be also used to check the final results provided by any clustering method. Indeed, for all the pairs of objects belonging to the same cluster, the pairs showing high or relatively high values of structural cliffs are pairs having low similarity although located into the same cluster.

References

[1] G. M. Maggiora, On outliers and activity cliffs - why qsar often disappoints, *J. Chem. Inf. Model.* **46** (2006) 1535-1535.

[2] L. Peltason, J. Bajorath, SAR index: quantifying the nature of structure-activity relationships, *J. Med. Chem.* **50** (2007) 5571-5578.

[3] R. Guha, J. H. Van Drie, Structure-activity landscape index: identifying and quantifying activity cliffs, *J. Chem. Inf. Model.* **48** (2008) 646-658.

[4] D. Stumpfe, A. de la Vega de Leon, D. Dimova, J. Bajorath, Advancing the activity cliff concept, part II, *F1000 Res.* **3** (2016) 1-8.

[5] H. Eckert, J. Bajorath, Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches, *Drug Discov. Today* **12** (2007) 225-233.

[6] J. Medina-Franco, K. Martinez-Mayorga, A. Bender, R. M. Marin, M. A. Giulianotti, C. Pinilla, R. A. Houghten, Characterization of activity landscapes using 2D and 3D similarity methods: consensus activity cliffs, *J. Chem. Inf. Model.* **49** (2009) 477-491.

[7] V. Namasivayam, J. Bajorath, Searching for coordinated activity cliffs using particle swarm optimization, *J. Chem. Inf. Model.* **52** (2012) 927-934.

[8] K. Klimenko, R-based tool for a pairwise structure-activity relationship analysis, *Mol. Inform.* **36** (2017) 1-5.

[9] L. Peltason, P. Iyer, J. Bajorath, Rationalizing three-dimensional activity landscapes and the influence of molecular representations on landscape topology and the formation of activity cliffs, *J. Chem. Inf. Model.* **50** (2010) 1021-1033.

[10] Y. Hu, D. Stumpfe, J. Bajorath, Advancing the activity cliff concept, *F1000 Res.* **2** (2013) 1-11.

[11] A. M. Wassermann, J. Bajorath, Chemical substitutions that introduce activity cliffs across different compound classes and biological targets, *J. Chem. Inf. Model.* **50** (2010) 1248-1256.

[12] M. Vogt, Y. Huang, J. Bajorath, From activity cliffs to activity ridges: informative data structures for SAR analysis, *J. Chem. Inf. Model.* **51** (2011) 1848-1856.

[13] D. Dimova, J. Bajorath, Extraction of SAR information from activity cliff clusters via matching molecular series, *Eur. J. Med. Chem.* **87** (2014) 454-460.

[14] D. Dimova, J. Bajorath, Advances in activity cliff research, *Mol. Inform.* **35** (2016) 181-191.

[15] Y. Hu, J. Bajorath, Extending the activity cliff concept: structural categorization of activity cliffs and systematic identifcation of different types of cliffs in the ChEMBL database, *J. Chem. Inf. Model.* **52** (2012) 1806-1811.

[16] R. Guha, J. H. Van Drie, Assessing how well a modeling protocol captures a structure-activity landscape, *J. Chem. Inf. Model*. **48** (2008) 1716-1728.

[17] D. Dimova, K. Heikamp, D. Stumpfe, J. Bajorath, Do medicinal chemists learn from activity cliffs? A systematic evaluation of cliff progression in evolving compound data sets, *J. Med. Chem.* **56** (2013) 3339-3345.

[18] J. Medina-Franco, Activity cliffs: facts or artifacts? *Chem. Biol. Drug Des.* **81** (2013) 553-556.

[19] R. Guha, J. Medina-Franco, On the validity versus utility of activity landscapes: are all activity cliffs statistically significant? *J. Cheminf.* **6** (2014) 1-9.

[20] D. Stumpfe, Y. Hu, D. Dimova, J. Bajorath, Recent progress in understanding activity cliffs and their utility in medicinal chemistry, *J. Med. Chem.* **57** (2014) 18-28.

[21] M. Wawer, L. Peltason, N. Weskamp, A. Teckentrup, J. Bajorath, Structure-activity relationship anatomy by network-like similarity graphs and local structure-activity relationship indices, *J. Med. Chem.* **51** (2008) 6075-6084.

[22] D. Dimova, M. Waver, A. M. Wassermann, J. Bajorath, Design of multitarget activity landscapes that capture hierarchical activity cliffs distributions, *J. Chem. Inf. Model*. **51** (2011) 258-266.

[23] D. Stumpfe, J. Bajorath, Activity cliff networks for medicinal chemistry, *Drug Develop. Res*. **75** (2014) 291-298.

[24] J. Bajorath, Representation and identification of activity cliffs, *Expert Opin. Drug Discov*. **12** (2017) 879-883.

[25] J. Bajorath, L. Peltason, M. Waver, R. Guha, M. S. Lajiness, J. H. Van Drie, Navigating structure-activity landscapes, *Drug Discov. Today* **14** (2009) 698-705.

[26] A. M. Wassermann, M. Waver, J. Bajorath, Activity landscape representations for structure-activity relationship analysis, *J. Med. Chem.* **53** (2010) 8209-8223.

[27] P. Iyer, D. Stumpfe, M. Vogt, J. Bajorath, G. M. Maggiora, Activity landscapes, information theory, and structure – activity relationships, *Mol. Inform.* (2013) 421-430.

[28] M. Cruz-Monteagudo, J. Medina-Franco, Y. Perez-Castillo, O. Nicolotti, M. N. D. S. Cordeiro, F. Borges, Activity cliffs in drug discovery: Dr Jekyll or Mr Hyde?, *Drug Discov. Today* **19** (2014) 1069-1080.

[29] O. Méndez-Lucio, J. Pérez-Villanueva, R. Castillo, J. Medina-Franco, Identifying activity cliff generators of PPAR ligands using SAS maps, *Mol. Inform.* **31** (2012) 837-846.

[30] J. Bajorath, From activity cliffs to promiscuity cliffs, *Future Sci. OA* **3** (2017).

[31] D. Dimova, J. Bajorath, Rationalizing promiscuity cliffs, *Chem. Med. Chem.* **12** (2017) 1-6.

[32] Y. Hu, G. M. Maggiora, J. Bajorath, Activity cliffs in PubChem confrmatory bioassays taking inactive compounds into account, *J. Comp. Aided Mol. Des*. **27** (2013) 115-124.

[33] D. Dimova, D. Stumpfe, Y. Hu, J. Bajorath, Activity cliff clusters as a source of structure–activity relationship information, *Expert Opin. Drug Discov*. **10** (2015) 441-447.

[34] D. Dimova, D. Stumpfe, J. Bajorath, Systematic assessment of coordinated activity cliffs formed by kinase inhibitors and detailed characterization of ctivity cliff clusters and associated SAR information, *Eur. J. Med. Chem.* **90** (2015) 414-427.

[35] J. Husby, G. Bottegoni, I. Kufareva, R. Abagyan, A. Cavalli, Structure-based predictions of activity clffs, *J. Chem. Inf. Model*. **55** (2015) 1062-1076.

[36] D. Horvath, G. Marcou, A. Varnek, S. Kayasha, A. de la Vega de Leon, J. Bajorath, Prediction of activity cliffs using condensed graphs of reaction representations, descriptor recombination, support vector machine classifcation, and support vector regression, *J. Chem. Inf. Model.* **56** (2016) 1631-1640.

[37] D. A. Winkler, T. C. Le, Performance of deep and shallow neural networks, the universal approximation theorem, activity cliffs, and QSAR, *Mol. Inform.* **36** (2016) 1-6.