JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

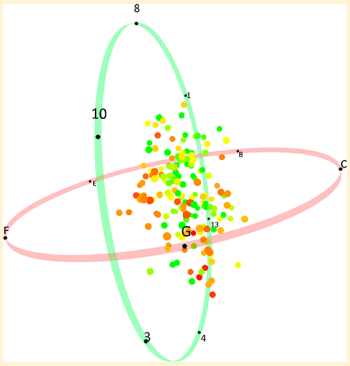
pubs.acs.org/jcim

Design of a Three-Dimensional Multitarget Activity Landscape

Antonio de la Vega de León 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

ABSTRACT: The design of activity landscape representations is challenging when compounds are active against multiple targets. Going beyond three or four targets, the complexity of underlying activity spaces is difficult to capture in conventional activity landscape views. Previous attempts to generate multitarget activity landscapes have predominantly utilized extensions of molecular network representations or plots of activity versus chemical similarity for pairs of active compounds. Herein, we introduce a three-dimensional multitarget activity landscape design that is based upon principles of radial coordinate visualization. Circular representations of multitarget activity and chemical reference space are combined to generate a spherical view into which compound sets are projected for interactive analysis. Interpretation of landscape content is facilitated by following three canonical views of activity, chemical, and combined activity/chemical space, respectively. These views focus on different planes of the underlying coordinate system. From the activity and combined views, compounds with well-defined target selectivity and structure−activity profile relationships can be extracted. In the activity landscape, such compounds display characteristic spatial arrangements and target activity patterns.



INTRODUCTION

The activity landscape concept has become increasingly popular in chemoinformatics and medicinal chemistry.[1,2] In general, an activity landscape is defined as any graphical representation that integrates molecular similarity and potency relationships between active compounds.[2] Reasons for the popularity of activity landscape modeling include that it provides access to structure−activity relationship (SAR) information contained in large and heterogeneous compound data sets and that it makes this information available in graphical form. As the size and complexity of compound data sets steadily increase, especially for high-priority therapeutic targets, methods for large-scale SAR analysis[2] and SAR visualization[3] are becoming important assets in medicinal chemistry that complement conventional SAR analysis strategies.

Activity landscape representations have mostly been designed for compounds with activity against individual targets, which represents the typical scenario for SAR analysis in medicinal chemistry. However, the activity landscape concept has also been extended for the exploration of compound selectivity and multitarget activities. Transforming a single-target activity landscape into a selectivity landscape is conceptually straightforward. The only requirement is the replacement of compound potency values with potency ratios for two targets.[4] Such selectivity landscapes have been generated on the basis of molecular network representations[4] and plots that report pairwise comparisons of chemical similarity and activity (difference) similarity of active compounds.[5] Compared to selectivity landscapes, it is considerably more challenging to generate multitarget activity landscapes. In this case, the underlying activity space is more complex and each compound has multiple activity annotations. A first multitarget activity landscape design utilized a network representation in which nodes were compounds annotated with multitarget activity codes and edges single- or multitarget activity cliffs.[6] In addition, pairwise potency differences have also been mapped on the axes of a Cartesian coordinate system generating a triple activity difference map in which data points represent compound pairs that are color-coded by calculated similarity.[7] This approach is only applicable to three targets. Another multitarget landscape design has been based on self-organizing maps in which structurally similar compounds were grouped together and annotated with pairwise target activity relationships.[8] This methodology also resolves activity profile differences between compounds at the level of target pairs and is hence only applicable to a limited number of targets (otherwise, too many pairs must be mapped). In addition, using plots that report pairwise compound and activity similarity, as discussed above, it has also been attempted to calculate activity similarity on the basis of vectors containing compound potency values for three targets.[9] In this case, the number of compound pairs and targets are also limiting factors for SAR information extraction. A common feature of multitarget activity landscape representations introduced thus far is that their interpretation is not straightforward if compounds are active against a larger number of targets, e.g., five or more. Recently, a first high-dimensional activity landscape representation has been introduced that captured compound activity against up to 100 targets, i.e., the ligand−target differentiation map.[10] The design of this simple two-dimensional (2D) representation differs from the ones discussed above in that high-dimensional bioactivity space is

Received: September 18, 2012
Published: October 31, 2012

first resolved by organizing compound pairs according to the number of targets they share and the number of targets against which they display a significant difference in potency. Then, structural similarity relationships between compounds in pairs are added to the representation.[10]

Herein we present a conceptually different design of a multitarget activity landscape that combines representations of chemical space and target space in a three-dimensional (3D) landscape view for interactive analysis. This activity landscape design has been applied to analyze compound data sets with activity against up to 15 targets. It reveals characteristic multitarget SAR patterns and enables the selection of key compounds containing selectivity information or SAR determinants.

MATERIALS AND METHODS

Design Principles. The design of the new 3D multitarget activity landscape was inspired by the radial coordinate visualization (RadViz) approach.[11] RadViz is a projection technique that maps objects from an m-dimensional to a 2D space. Following this approach, the original m dimensions are represented as dimensional anchors[12] that are arranged equidistantly around a circle (representing the 2D space). The position of an object inside the circle depends on its values for each anchor. The underlying principle for positioning of objects is that they are attracted by dimensional anchors. The attractive force correlates with the magnitude of the value of an object for each anchor (i.e., the value for each dimension); the larger the value is, the stronger becomes the force. The final position of an object results from the equilibrium of all forces (anchors).

For multitarget activity landscape design, two multidimensional spaces must be considered, i.e., the activity space and the chemical space. Therefore, two RadViz-like circles are arranged orthogonally, yielding a 3D activity landscape view. One of these circles represents (multitarget) activity, and the other, chemical space. The orthogonally arranged circles define a sphere. Compounds are positioned inside the sphere based on their activity values and similarity relationships calculated from chemical structure.

Activity and Chemical Space Representations. Each target is considered as a different dimensional anchor. The potency of each compound against the corresponding target is used as the value of the anchor. The similarity of targets is assessed by global sequence alignment[13] using the BLOSUM65 substitution matrix.[14] From the alignment scores, a similarity matrix is generated and used to position the targets on the basis of pairwise sequence similarity relative to each other on the activity space circle (i.e., by first maximally separating the most dissimilar targets from each other).

For chemical space representation (i.e., the chemical reference space defined by a compound data set), pairwise compound Tanimoto similarity[15] using the MACCS structural key fingerprint[16] is calculated. Other fingerprints can also be used. A diverse subset of reference compounds is utilized as dimensional anchors. Initial test calculations have revealed that extracting internal reference compounds from a data set under investigation generally leads to better resolved landscape representations than external references or scaffolds isolated from data set compounds. For the applications reported herein, six reference compounds are consistently used as anchors and placed on the chemical space circle on the basis of pairwise fingerprint distances. Diverse subset selection from a given data set is initiated with the two most dissimilar compounds (i.e., compounds with largest fingerprint distances per set). In subsequent iterations, the compounds with largest average fingerprint distance to compounds already contained in the subset are selected. For consistent data display in all 3D landscapes, chemical similarity and potency values of compounds are normalized to a value range [0,1]. Therefore, chemical similarity and potency values are transformed into conventional Z-scores and cumulative probabilities are calculated assuming a normal distribution. Although similarity values fall into the interval [0,1], normalization ensures that potency and similarity value distributions are on an equivalent scale for compound projection. By contrast, sequence similarities are not normalized because they are not further utilized for compound projection after target anchors have been placed.

Following the generation of each circle, the circles are orthogonally arranged to generate a sphere representing the activity landscape, as illustrated in Figure 1.

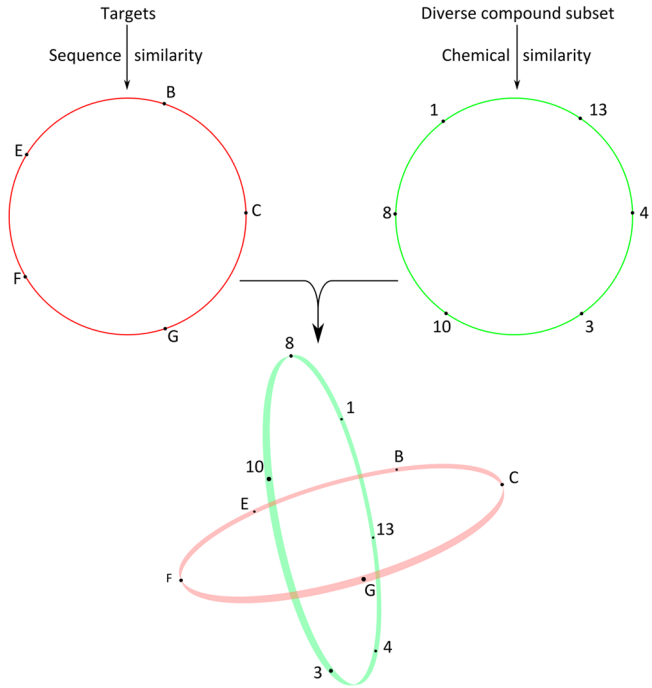


Figure 1. Initialization. The activity space is represented by a set of targets that are positioned on the red circle based on pairwise sequence similarity. Chemical reference space is delineated using a diverse subset of test compounds that are positioned on the green circle on the basis of pairwise fingerprint similarity. The two circles are arranged orthogonally to each other, hence defining a sphere.

Compound Projection. The position of each compound is calculated using the formula:

$$\vec{u}_i = \frac{\sum_{j \in A} k_{ij} \vec{A}_j}{|A|}$$

Here, u_i represents compound i, A_j, the dimensional anchor j, and k_{ij}, the normalized potency (target anchor) or similarity (compound anchor) value of compound i for anchor j. Depending on the distance distributions observed in a given data set, compound distances from the center of the landscape

dx.doi.org/10.1021/ci300444p | J. Chem. Inf. Model. 2012, 52, 2876−2883

sphere can also be scaled to further increase the resolution of the data set projections.

Implementation. The activity landscape design has been implemented in Java. The landscape generation program uses BioJava[17] for sequence alignment, the OpenEye Java tools[18] for chemical structure handling and similarity calculations, and Java3D[19] for generating landscape components and the 3D environment. Compounds are represented as SMILES strings.[20]

Compound Sets. Compound data sets with multitarget annotations are rare in the public domain, in contrast to pharmaceutical environments where target profiling has almost become a routine procedure. We have extracted compound sets yielding a complete compound−target activity matrix for 5, 10, and 15 targets from a publicly available kinase inhibitor profiling data set.[21] This data set contains 1497 compounds with available structures and K_i measurements for varying numbers of 172 kinases. The extracted compound sets are reported in Table 1. Two of these compound sets with activity against five different kinases comprise 184 and 144 compounds, respectively. In addition, sets with activity against 10 and 15 kinases contain 63 and 33 compounds, respectively. The activity profile of a compound is formed by the K_i values for all targets the compound is active against.

Table 1. Compound Sets[a]

set	no. of cpds	targets
5T184	184	CLK2, CLK4, MAP4K4, DYRK1A,GSK3B
5T144	144	LCK, KDR, FLT1, CSF1R, FLT3
10T63	63	CLK2, CLK4, MAP4K4, DYRK1A, GSK3B, MAP4K2, PRKCN, CDK5, ROCK1, PRKACA
15T33	33	CSF1R, KDR, FLT1, FLT3, LCK, FGFR1, NTRK2, MAP4K4, MAP4K2, NTRK1, FYN, RPS6KA3, TAOK1, STK6, ROCK1

[a]For each set (i.e., a different subset of the kinase inhibitor profiling data set[21]), the number of compounds is reported that have K_i values for all listed targets. The set designation contains the number of compounds and targets. For example, "5T184" means that the subset contains 184 compounds with activity measurements against five targets. Kinases are named according to ref 21.

RESULTS AND DISCUSSION

Three-Dimensional Multitarget Landscape. Figure 2 shows an exemplary compound projection into the 3D multitarget activity landscape. The activity and chemical space circles are colored red and green, respectively, and the positions of dimensional anchors are marked. Compounds are individually placed in the spherical view and color-coded according to their normalized average potency against all targets using a continuous color code from green (value 0) over yellow to red (value 1). The two radial coordinate displays facilitate a dimension reduction of activity and chemical space. The orthogonal arrangement of the corresponding circles generates a canonical framework for activity landscape display and navigation. Dimensional anchors are positioned on two orthogonal planes, i.e., the *XY* plane (target anchors) and *YZ* plane (compound anchors). The *X* coordinate value of compounds projected into the activity landscape depends only on their potency, the *Z* value only on their chemical structure, and the *Y* value on both potency and structure. Accordingly, the *XY* plane is focused on (multitarget) activity,

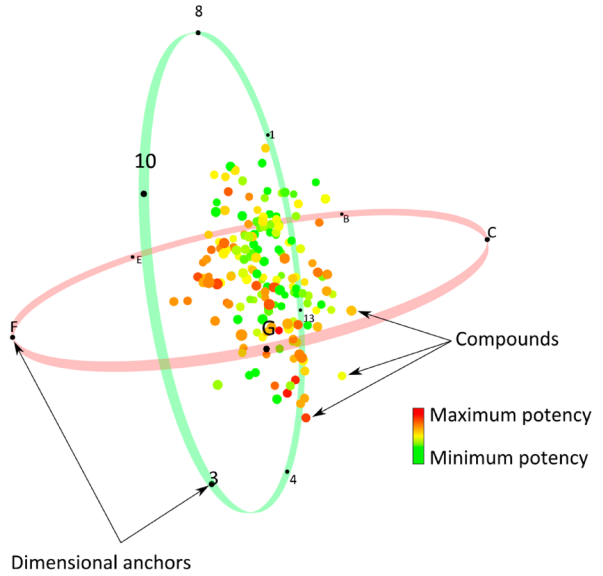


Figure 2. Dimensional anchors and compounds. On the activity and chemical space circles, dimensional anchors are shown as black spheres. Compounds are represented as spheres that are color-coded according to their normalized average potency values. The position of a compound depends on its potency against each target and its chemical similarity relative to the reference compounds.

and the *YZ* plane on chemical space. In addition, the *XZ* plane presents chemical and activity information in context.

Interpretation. This canonical landscape framework yields three well-defined viewing directions for landscape analysis, as illustrated in Figure 3A. Each view focuses on a combination of two orthogonal coordinate axes and the resulting plane. It can be generated in four different orientations because each of the two axes defining a view can be inspected along two opposite directions (i.e., from the "front" or "back" in Figure 3A). Each orientation provides equivalent information. Herein, we consistently utilize defined viewing directions. In each case, the camera is placed on the positive value side of the axis (that is not displayed) facing the center. The view is then oriented such that the positive values for the vertical axis are placed on the upper half (i.e., the *Y* axis of the activity view and the *Z* axis of the chemical and combined views).

The activity view is centered on the *XY* plane and the chemical view of the *YZ* plane, as shown in Figure 3B. Following these views, compound distributions in activity and chemical space are analyzed, respectively. The chemical view focuses on compound similarity relationships and is straightforward to interpret. Here, the distance between compounds inversely correlates with their similarity. Furthermore, for multitarget activities, the activity view can directly reveal compound selectivity patterns. Figure 4A illustrates how the position of a compound along the activity view correlates with its target selectivity. A compound near the center of the plane is not selective because it is essentially equally drawn to all targets. By contrast, a compound placed near the border of the activity space circle has selectivity for the nearest dimensionality anchor over other targets. Compounds located near the border and the *Y*-axis (representing the chemical space circle) might either be selective or nonselective. If compounds in these regions are nonselective, they must be highly similar to reference compounds on the chemical space circle (and are thus drawn to the boundary of the circle near the *Y*-axis along the activity

dx.doi.org/10.1021/ci300444p | J. Chem. Inf. Model. 2012, 52, 2876−2883

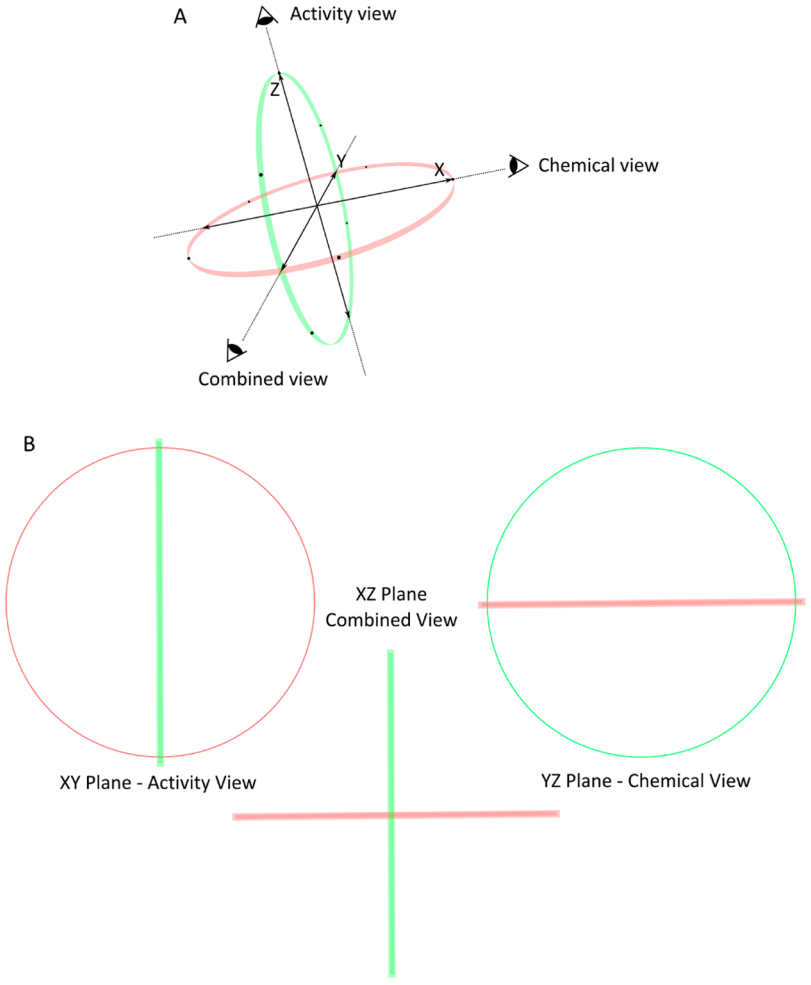


Figure 3. Landscape viewing. (A) The multitarget activity landscape can be viewed in three different ways including the activity, chemical, and combined view. The direction of each view focuses on a plane of an orthogonal coordinate system. (B) Schematic representations of each of the three views and the corresponding plane.

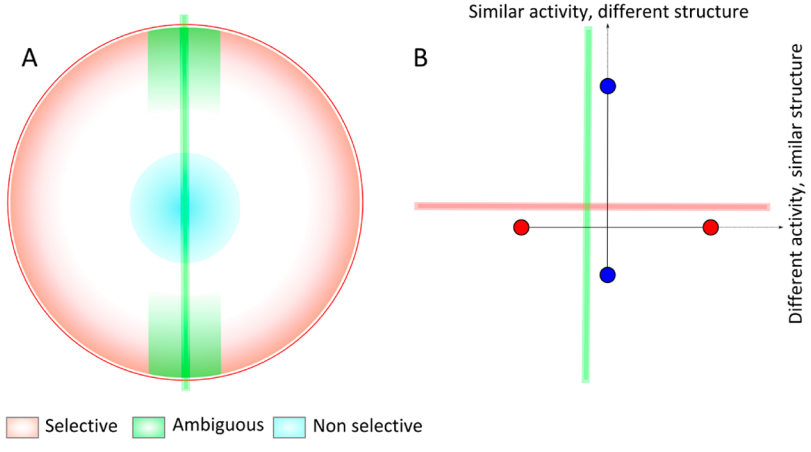


Figure 4. Analysis of the activity and the combined view. (A) Along the activity view, the position of a compound correlates with the degree of target selectivity. Compounds placed near the border of the activtiy circle are mostly selective for a given target, whereas compounds close to the center are nonselective (i.e., comparably active against multiple targets). Compounds mapping to the green areas might be selective or nonselective, as discussed in the text. (B) Along the combined view, pairs of compounds (blue) that that are vertically arranged, i.e., connected by a line that is parallel to the *Z*-axis (see Figure 3A), have a similar activity profile but different structures. By contrast, pairs of compounds (red) that are horizontally arranged, i.e., connected by a line parallel to the *X*-axis, have different activity profiles but similar structures.

view). The color code reflects against how many targets a compound has high potency. For example, dark green and red spheres represent compounds that have high potency against essentially none or all of the targets, respectively. It follows that these spheres can only be found near the center of the activity view.

dx.doi.org/10.1021/ci300444p | *J. Chem. Inf. Model.* 2012, 52, 2876−2883

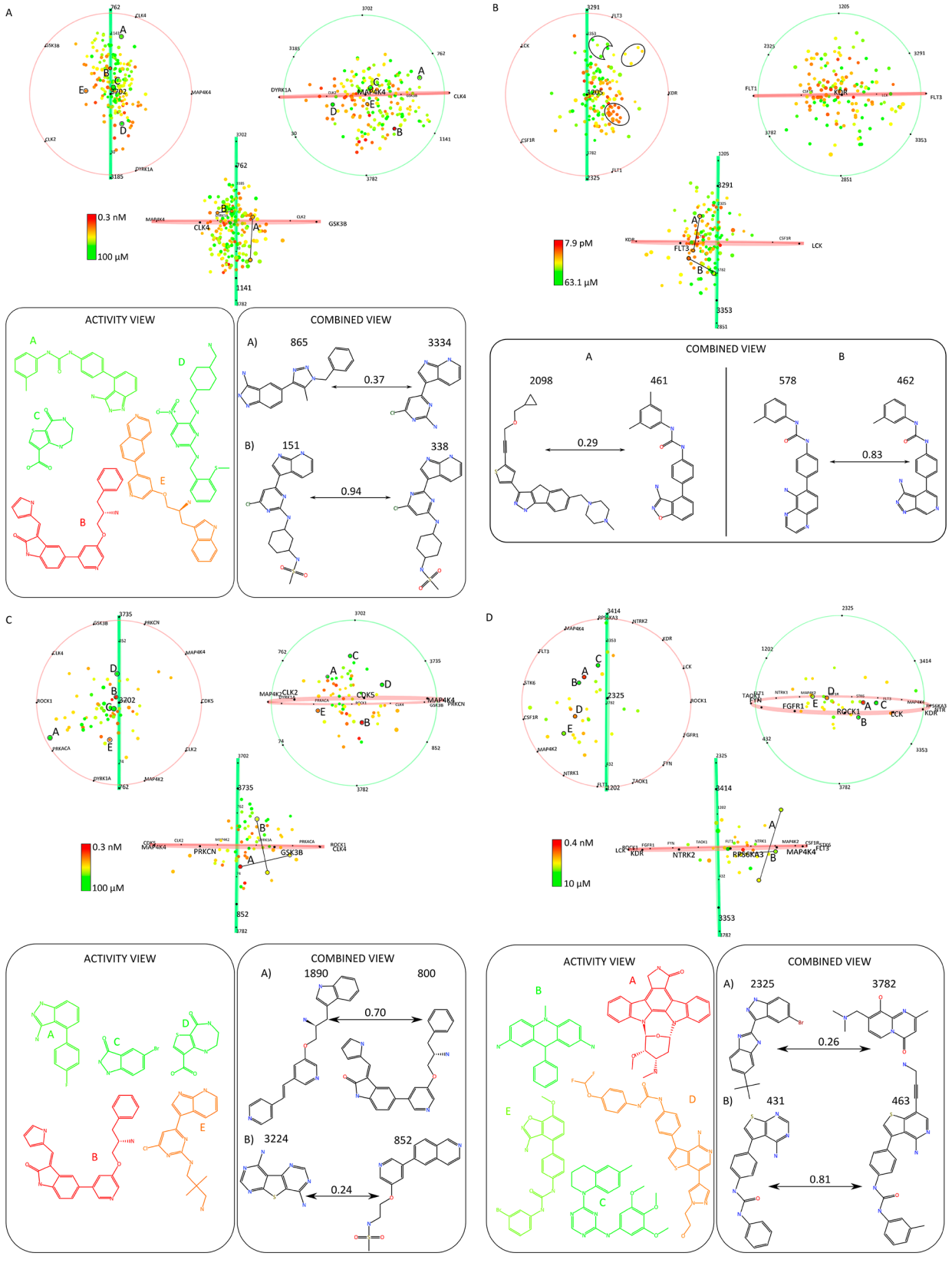


Figure 5. Exemplary applications. Results from the analysis of four sets of kinase inhibitors according to Table 1 are presented. At the top, the compound distribution along the three landscape views is diplayed. At the bottom, exemplary compounds selected from the activity and/or combined view are shown. Compounds taken from the activtiy view are individually labeled and color-coded according to their average potency. The positions of these compounds are also shown in the chemical view. Compound pairs extracted from the combined view are labeled. Compound data set IDs and MACCS Tanimoto similarity values are provided for each pair. Sets (A) 5T184, (B) 5T144, (C) 10T63, and (D) 15T33 are shown. In the activtiy view of B, groups of compounds having similar activtiy profiles are encircled.

dx.doi.org/10.1021/ci300444p | J. Chem. Inf. Model. 2012, 52, 2876−2883

Table 2. Activity Values of Selected Compounds[a]

		Figure 5A								
		activity view					combined view			
compound	label	A	B	C	D	E	A	A	B	B
	ID	**524**	**800**	**3582**	**3541**	**876**	**865**	**3334**	**151**	**338**
activity	**MAP4K4**	5.8	8.3	4.5	6.5	5.3	6	6.1	6.7	7.8
	CLK4	7.1	9.5	5.4	5.5	9	8	7.7	8.5	7.2
	GSK3B	5.4	8.6	4.3	5.7	6.8	7.2	6.9	6.1	7.8
	CLK2	5.7	9.1	4.7	5.7	8.6	7	7.1	7.8	5.7
	DYRK1A	5.3	9.4	5.2	5.5	8.5	7	6.9	7.3	7.3

		Figure 5B			
		activity view (N/A)	combined view		
compound	label		A	A	B	B
	ID		**2098**	**461**	**578**	**462**
activity	**KDR**		8	8.7	7.2	8.9
	FLT3		8.8	8.6	7.3	9.4
	LCK		6.7	6.9	5.9	6.5
	CSF1R		8.5	8.4	8.2	9.2
	FLT1		9.4	9.6	7.6	10.3

		Figure 5C								
		activity view					combined view			
compound	label	A	B	C	D	E	A	A	B	B
	ID	**2671**	**800**	**763**	**3582**	**192**	**1890**	**800**	**3224**	**852**
activity	**CDK5**	4.7	9.2	5.5	4.9	8.9	5.3	9.2	6.0	5.8
	MAP4K4	4.7	8.3	4.7	4.5	7.3	6.4	8.3	6.1	5.9
	PRKCN	4.1	7.8	5.5	4.5	6.8	6.1	7.8	6.1	6.1
	GSK3B	4.7	8.6	4.7	4.3	5.9	5.9	8.6	5.8	6.2
	CLK4	7.4	9.5	5.9	5.4	8.2	8.5	9.5	7.9	8.3
	ROCK1	6.3	8.4	4.8	4.2	8.6	7.6	8.4	6.5	6.6
	PRKACA	7.1	8.9	4.7	4.3	8.4	7.5	8.9	5.8	5.9
	DYRK1A	4.6	9.4	5.8	5.2	7.4	7.4	9.4	7.1	7.3
	MAP4K2	4.5	7.8	4.5	4.4	7.3	5.4	7.8	5.8	5.4
	CLK2	5.7	9.1	4.8	4.7	7.3	7.6	9.1	6.7	7.3

		Figure 5D								
		activity view					combined view			
compound	label	A	B	C	D	E	A	A	B	B
	ID	**3629**	**3353**	**3048**	**506**	**526**	**2325**	**3782**	**431**	**463**
activity	**ROCK1**	9.1	6.1	5.5	6.7	5.6	6.3	6.8	5.7	6.4
	LCK	7.6	5.3	5.3	8.9	6.5	6.2	6.5	6.9	6.8
	KDR	8.2	5.7	6.1	8.9	6.9	6.6	6.3	6.8	8.1
	NTRK2	9.1	5.9	6.1	8.4	6.4	6.9	6.5	5.8	7.1
	RPS6KA3	8.8	6	5.6	6.1	6.2	6.1	6.9	5.5	6.2
	MAP4K4	8.8	6.3	5.9	7.6	5.9	6.5	6.8	5.8	6.1
	FLT3	7.9	6.4	6	7.9	7.3	8.1	7.2	8.5	8.6
	STK6	8.6	6.3	5.8	8.8	6.9	6.9	7.4	8.3	7.2
	CSF1R	8.1	5.9	5.7	8	7.1	6.5	6.9	6.7	8.3
	MAP4K2	8.3	6.6	5.8	7.7	6.2	7.4	7.1	6.6	6.3
	NTRK1	8.2	6.2	6.1	7.8	6	7.2	7	5.8	6.7
	FLT1	8.3	6.4	5.5	8.6	8	7.1	7.1	7.3	9.4
	TAOK1	8	5.9	6.1	6.1	5.9	5.4	6.4	5.6	6.4
	FYN	7.3	5.7	5.3	7.2	6.5	6.6	6.6	6.3	6.3
	FGFR1	7.4	6	5.4	7.7	5.9	6.3	7.2	6.4	7.6

[a]Activity values (pK_i) of compounds highlighted in Figure 5 are given.

Different from the chemical and activity views, the combined view focuses on the *XZ* plane, as also shown in Figure 3B. Importantly, along the combined view, SAR information becomes apparent, if available. In the *XZ* plane, well-defined structure−activity profile relationships between compounds are detected, as illustrated in Figure 4B. If two compounds are vertically arranged parallel to the *Z*-axis, they have similar activity profiles for multiple targets but are structurally distinct. By contrast, if two compounds are horizontally arranged

dx.doi.org/10.1021/ci300444p | J. Chem. Inf. Model. 2012, 52, 2876−2883

parallel to the *X*-axis, they are structurally similar but have different activity profiles

Kinase Inhibitor Analysis. As an exemplary application of our 3D multitarget activity landscape method, four sets of kinase inhibitors with activity against five, 10, or 15 different kinases are analyzed. These compound sets are reported in Table 1. In Figure 5, the 3D activity landscapes formed by these compound sets are shown. Exemplary compounds are highlighted and their activity values are provided in Table 2.

The activity landscape of the 5T184 set is presented in Figure 5A, which shows the compound distribution along the three different views. In this set, compounds are active against five targets. Exemplary compounds are marked and displayed that reveal characteristic information, as discussed above. For example, along the activity view, compound A is located near the border and is selective for kinase CLK4, i.e., it has high potency against this kinase but low potency values for the other four targets (and thus a relatively low average potency). By contrast, compounds B and C are located near the center of the circle and are nonselective at a high and low average potency level, respectively. In addition, the activity profile of compound D is similar to C, but D is removed from the center due to high similarity to a reference compound. Furthermore, compound E has high potency against all targets, except MAP4K4. Thus, its position is opposite to the MAP4K4 anchor. Along the combined view in Figure 5A, the compounds in pair A have limited structural similarity (Tanimoto similarity of 0.37) but very similar activity profiles, consistent with their nearly parallel arrangement relative to the vertical axis. Their average potency difference is less than 0.2 pK_i units for each target. By contrast, the compounds in pair B are very similar (Tanimoto similarity of 0.94) but display potency differences of approximately 2 orders of magnitude against four of five targets, consistent with their nearly parallel arrangement relative to the horizontal axis.

Figure 5B shows the activity landscape of set 5T144, in which compounds are also active against five kinases. In the activity view, three clusters of compounds with similar activity profiles are marked. The cluster with the green spheres contains compounds selective for FLT3. The cluster with yellow spheres corresponds to compounds with high potency against FLT3, FLT1, and KDR. Their potency against FLT3 is higher than for the other two kinases. Accordingly, the cluster is located closer to the FLT3 anchor than the others. Furthermore, the cluster with red spheres contains compounds with high potency against all targets, except LCK. The overall uneven distribution of compounds along the activity view correlates with a general difference in the distribution of potency values for the five targets. Most compounds in this set have higher potency against FLT1, FLT3, and KDR than against CSF1R and LCK. Hence, the left half of the activity view is only sparsely populated. From the combined view, two pairs of compounds are also selected that display low structural and high activity profile similarity (pair A) and high structural and low activity profile similarity (B), corresponding to the relationships discussed above.

In the activity view of set 10T63 in Figure 5C, five compounds have been marked, one of which (compound A) is selective for two of 10 kinases (CLK4 and PRKACA), whereas three others (B, C, and D) are nonselective. Compound A is located close to the border of the activity circle, whereas B and C are near the center, as expected. Compound D is further removed from the center because of high structural similarity to reference compound 3735. In addition, compound E is highly potent against eight of 10 targets (except GSK3B and PRKCN) and thus positioned opposite to those targets (anchors). From the combined view in Figure 5C, two compound pairs with inverse structure−activity profile relationships are extracted.

Figure 5D represents the activity landscape for the 15T33 set that contains only 33 compounds. However, all of these compounds are active against 15 kinases. In the activity view, there are only very few compounds located near the center. Thus, most compounds vary in their potency against at least a few targets. Given this variance, compounds with high (A) or low (B) average potency are more removed from the center when compared to compounds with high and low average potency in lower-dimensional activity spaces, as discussed above. Compound C has consistently low potency against all 15 kinases but is highly similar to reference compound 3414, which alters its position compared to other consistently lowly potent compounds. Furthermore, compound D has high potency against 12 of 15 kinases (except RPS6KA3, TAOK1, and ROCK1), but E is only highly potent against FLT1, FLT3, and CSF1R. In the combined view in Figure 5D, two compound pairs have been marked. Compounds in pair A have low similarity but an overall similar activity profile. In this case, compound potency against only three of 15 targets varies by approximately an order of magnitude. In contrast, compounds in pair B are structurally highly similar (and thus in parallel arrangement relative to the horizontal axis) but compound 463 is consistently more potent against all kinases than compound 431. Hence, the activity profiles of these two compounds are distinct.

Scope and Limitations. Our 3D activity landscape representation is designed to deconvolute bioactivity space for multiple targets. In our experience, 10−20 different targets can be well covered. If the number of targets is larger than 20, the interpretation of SAR and selectivity patterns becomes more challenging. In this case, it is difficult to analyze the influence of specific targets on the position of a molecule. By contrast, the number of compounds has only little influence on interpretability and is not considered a limiting factor. Multitarget activity landscape analysis as presented herein is applicable to both compound optimization and screening data (which might result, for example, from compound profiling experiments on families of targets). In the latter case, inactive compounds would be excluded.

It should also be noted that the chemical reference space utilized in our multitarget activity landscape is a local space because the reference compounds used to define the space are obtained from the data set that is analyzed. Therefore, several data sets can currently not be compared within the same representations (which also falls outside the scope of the methodology). However, this might be facilitated using "external" references such as is done, for example, in the ChemGPS representation by Oprea and Gottfries.[22] ChemGPS is a global three-dimensional chemical space viewer in which the axes are principal components calculated from a variety of chemical descriptors. Thus, ChemGPS is not an activity landscape representation.

■ CONCLUSIONS

Herein we have introduced a new multitarget activity landscape design that utilizes principles of radial coordinate visualization and dimensionality reduction. Characteristic features of this multitarget activity landscape include its three-dimensional design and its resolution at the level of individual compounds,

rather than compound pairs. This multitarget activity landscape model is primarily designed for analyzing the information contained in complete compound-target profiling matrices. The orthogonal arrangement of nonlinear projections of activity and chemical space yields a spherical multitarget landscape view that integrates molecular and activity similarity relationships. In this activity landscape, compound positions and positional relationships are associated with specific SAR information content. The landscape is best analyzed on the basis of three canonical views accounting for chemical and activity space and, in addition, combined chemical/activity space. The chemical and activity views reveal structural relationships and target selectivity information of active compounds, respectively. Moreover, the combined view reveals different types of structure−activity profile relationships. The 3D multitarget activity landscape introduced herein conceptually differs from previous activity landscape representations and further expands the currently still limited spectrum of activity landscape models for navigating high-dimensional bioactivity spaces.

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) Bajorath, J.; Maggiora, G.; Lajiness, M., organizers. The Emerging Concepts of Activity Landscapes and Activity Cliffs and their Role in Drug Research. *240th National Meeting of the American Chemical Society, Divisions of Chemical Information and Computers in Chemistry*, Boston, MA, August 22−26, 2010.

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

(3) Stumpfe, D.; Bajorath, J. Methods for SAR Visualization. *RSC Adv.* **2012**, *2*, 369−378.

(4) Peltason, L.; Hu, Y.; Bajorath, J. From Structure-Activity to Structure-Selectivity Relationships: Quantitative Assessment, Selectivity Cliffs, and Key Compounds. *ChemMedChem* **2009**, *4*, 1864−1873.

(5) Perez-Villanueva, J.; Santos, R.; Hernandez-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. Structure-Activity Relationships of Benzimidazole Derivatives as Anti-Parasitic Agents: Dual-Activity Difference (DAD) Maps. *Med. Chem. Commun.* **2011**, *2*, 44−49.

(6) Dimova, D.; Wawer, M.; Wassermann, A. M.; Bajorath, J. Design of Multi-target Activity Landscapes that Capture Hierarchical Activity Cliff Distributions. *J. Chem. Inf. Model.* **2011**, *51*, 256−288.

(7) Medina-Franco, J. L.; Yongye, A. B.; Perez-Villanueva, J.; Houghten, R. A.; Martinez-Mayorga, K. Multitarget Structure-Activity Relationships Characterized by Activity-Difference Maps and Consensus Similarity Measures. *J. Chem. Inf. Model.* **2011**, *51*, 2427−2439.

(8) Iyer, P.; Bajorath, J. Representation of Multi-Target Activity Landscapes through Target Pair-Based Compound Encoding in Self-Organizing Maps. *Chem. Biol. Drug Des.* **2011**, *78*, 778−786.

(9) Waddell, J.; Medina-Franco, J. L. Bioactivity Landscape Modeling: Chemoinformatic Characterization of Structure-Activity Relationships of Compounds Tested across Multiple Targets. *Bioorg. Med. Chem.* **2012**, *20*, 5443−5452.

(10) Iyer, P.; Dimova, D.; Vogt, M.; Bajorath, J. Navigating High-Dimensional Activity Landscapes: Design and Application of the Ligand-Target Differentiation Map. *J. Chem. Inf. Model.* **2012**, *52*, 1962−1969.

(11) Hoffman, P.; Grinstein, G.; Marx, K. A.; Grosse, I.; Stanley, E. DNA Visual and Analytic Data Mining. In *Proceedings of the 8th Conference on Visualization '97*, Phoenix, AZ, USA, October 18−24; IEEE Computer Society Press: Los Alamitos, CA, USA, 1997; pp 437−441.

(12) Hoffman, P.; Grinstein, G.; Pinkney, D. Dimensional Anchors: A Graphic Primitive for Multidimensional Multivariate Information Visualizations. In *Workshop on New Paradigms in Information Visualization and Manipulation (NPIVM '99), in conjunction with the Eighth ACM International Conference on Information and Knowledge Management (CIKM '99)*, Kansas City, MO, USA, November 2−6; ACM: New York, NY, USA, 1999; pp 9−16.

(13) Needleman, S. B.; Wunsch, C. D. A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. *J. Mol. Biol.* **1970**, *48*, 443−453.

(14) Henikoff, S.; Henikoff, J. G. Amino Acid Substitution Matrices from Protein Blocks. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10915−10919.

(15) Willett, P.; Barnard, J. M.; Downs, G. M. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983−996.

(16) *MACCS Structural keys*; Symyx Software, San Ramon, CA, USA, 2005.

(17) Holland, R. C. G.; Down, T. A.; Pocock, M.; Prlić, A.; Huen, D.; James, K.; Foisy, S.; Dräger, A.; Yates, A.; Heuer, M.; Schreiber, M. J. BioJava: An Open-Source Framework for Bioinformatics. *Bioinformatics* **2008**, *24*, 2096−2097.

(18) *OEChem TK*, version 1.7.4.3; OpenEye Scientific Software Inc.: Santa Fe, NM, 2010.

(19) *Java3D*, version 1.5; Sun Microsystems, 2008; java.net/projects/java3d (accessed August 23, 2012).

(20) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.

(21) Metz, J. T.; Johnson, E. F.; Soni, N. B.; Merta, P. J.; Kifle, L.; Hajduk, P. J. Navigating the Kinome. *Nature Chem. Biol.* **2011**, *7*, 200−202.

(22) Oprea, T. I.; Gottfries, J. Chemography: The Art of Navigating in Chemical Space. *J. Comb. Chem.* **2001**, *3*, 157−166.