JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

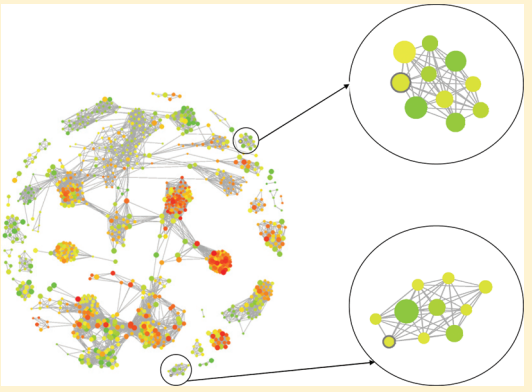
Rationalizing the Role of SAR Tolerance for Ligand-Based Virtual Screening

Peter Ripphausen, Britta Nisius, Mathias Wawer, and Jürgen Bajorath*

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

Supporting Information

ABSTRACT: It is well appreciated that the results of ligand-based virtual screening (LBVS) are much influenced by methodological details, given the generally strong compound class dependence of LBVS methods. It is less well understood to what extent structure−activity relationship (SAR) characteristics might influence the outcome of LBVS. We have assessed the hypothesis that the success of prospective LBVS depends on the SAR tolerance of screening targets, in addition to methodological aspects. In this context, SAR tolerance is rationalized as the ability of a target protein to specifically interact with series of structurally diverse active compounds. In compound data sets, SAR tolerance articulates itself as SAR continuity, i.e., the presence of structurally diverse compounds having similar potency. In order to analyze the role of SAR tolerance for LBVS, activity landscape representations of compounds active against 16 different target proteins were generated for which successful LBVS applications were reported. In all instances, the activity landscapes of known active compounds contained multiple regions of local SAR continuity. When analyzing the location of newly identified LBVS hits and their SAR environments, we found that these hits almost exclusively mapped to regions of distinct local SAR continuity. Taken together, these findings indicate the presence of a close link between SAR tolerance at the target level, SAR continuity at the ligand level, and the probability of LBVS success.



INTRODUCTION

The field of ligand-based virtual screening (LBVS) is characterized by a high degree of methodological diversity and a strong emphasis on benchmarking, while generally applicable community standards for method assessment are still lacking.[1] Hence, it continues to be difficult to compare different LBVS methods and judge their performance. However, while LBVS methods generally show strong compound class dependence,[1,2] methodological details are not the only factor that might influence LBVS performance, which is particularly relevant for prospective LBVS applications. Compound class dependence of LBVS methods, all of which essentially employ the concept of (global or local) molecular similarity, already implies that the nature of the screening targets might also play an important role. Of course, in LBVS, target and target−ligand interaction information is not taken into account. However, such information is, at least to some extent, implicitly contained in available active reference compounds (whose structural features and molecular properties must be complementary to a given binding site). In prospective LBVS applications, where the ultimate goal is to identify structurally diverse active compounds,[1] it is also frequently observed that methods of different design and complexity are comparably successful in identifying new hits, while in other cases many, if not all, methods fail.[2] These findings reflect the well-known compound class dependence of virtual screening methods. However, such observations have also given rise to the hypothesis that LBVS methods can only be successfully applied if their targets are capable of binding a different series of compounds, i.e., if they are structure−activity relationship (SAR) tolerant, as we term it here. This hypothesis might be intuitive, but it is currently lacking solid support. Importantly, in compound data sets, SAR tolerance is mirrored by the presence of so-called SAR continuity,[2] which is introduced by structurally similar and/or increasingly dissimilar compounds having similar potency. By contrast, the presence of SAR discontinuity, where only a few very similar molecules are active and where small structural changes diminish their activity, generally falls outside the applicability domain of LBVS approaches.[2] In such cases, the probability of finding a new active chemotype is typically very low. Thus, the study of SAR continuity/discontinuity in compound data sets enables the evaluation of the degree of SAR tolerance displayed by LBVS targets, albeit indirectly.

Activity landscape representations integrate similarity and potency relationships between specifically active compounds and can be utilized to analyze the SAR character of compound data sets.[3] Landscape representations of a rather different design have been introduced to study both local and global SAR features

Received: February 9, 2011
Published: March 25, 2011

Table 1. Targets, Compound Sets, and Computational Methods[a]

target	ref	actives	hits	method
5-lipoxygenase (5LO)	16	967	2	2D fingerprint
acetylcholinesterase (AChE)	17	1394	1	3D pharmacophore
aromatase (CYP19)	18	966	3	3D pharmacophore
cannabinoid receptor 1 (CB1)	19	1799	27	3D pharmacophore
C−C chemokine receptor type 2 (CCR2)	20	656	3	3D pharmacophore
dopamine D3 receptor (D3)	21	1571	2	machine learning
dopamine transporter (DAT)	22	1202	1	3D pharmacophore
endothelin receptor ET-A(ETA)	23	1230	2	3D pharmacophore
glycogen synthase kinase-3, beta (GSK3B)	24	619	1	2D fingerprint
growth hormone secretagogue receptor (GHSR)	25	483	1	2D fingerprint
heat shock protein 90 (HSP90)	26	519	1	3D pharmacophore
histone deacetylase 1 (HDAC1)	27	1037	2	machine learning
serotonin receptor 2B (5HT2B)	28	308	3	machine learning
sphingosine 1-phosphate receptor 3(SIP3)	29	160	2	3D pharmacophore
urotensin II receptor (UTR2)	30	103	3	3D pharmacophore
voltage-gated potassium channel subunit Kv1.5 (KV15)	31	342	1	3D pharmacophore

[a] Targets of successful LBVS campaigns are reported and the number of known active compounds ("actives") in each data set that were utilized for activity landscape modeling. Abbreviations are provided that are used in the text. The LBVS studies are referenced and in each case, and the number of LBVS hits is given that were suitable for landscape modeling. Also reported are the successfully applied LBVS methods.

of compound data sets.[3] Provided sufficient numbers of active compounds are already available for a prospective virtual screening target, the corresponding activity landscape can be analyzed to examine whether this target might be SAR tolerant. If so, then the activity landscape of already known active compounds should contain regions of notable SAR continuity.

In this study, we have characterized the activity landscapes of compound data sets active against different targets for which scientifically sound prospective LBVS studies have been reported that have led to the identification of hits containing previously unobserved molecular scaffolds. Furthermore, we have analyzed which regions of these activity landscapes newly identified hits fell and characterized their SAR environment.

MATERIALS AND METHODS

Selection of Virtual Screening Hits. In a recent survey of published LBVS applications,[4] we have identified 78 studies using different LBVS methods that addressed a defined molecular target, utilized appropriate assays to confirm predicted hits, and reported valid scaffolds hops. In each case, we confirmed that new hits contained carbon skeletons that were not yet available in public domain active compounds. These 78 LBVS studies provided the starting point for our analysis. From them, we initially selected 58 studies that yielded hits with potency higher than 10 μM. Hits with lower potency were not further considered because we intended to focus our analysis on compounds for which the probability of false-positive activity assignments was low. Furthermore, for activity landscape analysis, as further described below, we required the availability of significant numbers of known active compounds, with at least one compound exceeding a predefined similarity threshold to newly reported hit(s). Therefore, we next collected for the proteins targeted in all 58 studies active compounds available in the most recent version of BindingDB[5] (that also includes ChEMBL[6] compounds). We identified those targets for which at least 50 active compounds were available, leading to the further consideration of 34 studies. Then, we compared each LBVS hit to all known active compounds using the "extended connectivity fingerprint with bond diameter 4" (ECFP4),[8] as implemented in Pipeline Pilot,[9] calculated the Tanimoto coefficient (Tc),[7] and identified hits that yielded a Tc value of greater than 0.4 relative to at least one database compound. This ECFP4 Tc threshold value was previously shown to be a reasonable similarity criterion for activity landscape analysis.[10] Hits that did not reach this level of similarity compared to any database compound had no structural neighbors and were thus not suitable for landscape modeling. Following similarity calculations, 16 studies with qualifying hits remained that targeted 16 different proteins, as summarized in Table 1. The analysis was carried out with in-house generated Perl and Pipeline Pilot scripts.

Activity Landscape Models. We utilized network-like similarity graphs (NSGs)[11] to represent the activity landscapes of compound data sets. An NSG is an annotated graph representation that captures similarity and potency relationships in a data set. Nodes represent compounds that are connected by an edge if their pairwise similarity exceeds the ECFP4 Tc. Nodes are color coded according to the potency range present in a compound set using a continuous gradient from green (lowest potency) via yellow to red (highest potency). In addition, nodes are scaled in size according to the SAR index (SARI)[12] per compound discontinuity score[11] that accounts for the contribution of each individual compound to local SAR discontinuity. Accordingly, large red and green nodes linked by edges represent activity cliffs in a data set, the extreme form of SAR discontinuity. For NSG display, the Fruchterman−Reingold[13] graphical layout algorithm is applied. This algorithm places densely interconnected compounds in close proximity and separates weakly or unconnected groups of compounds from each other. NSGs were drawn using the igraph[14] package of R.[15] NSGs for the 16 data sets were calculated after adding the LBVS hits to the corresponding active database compounds.

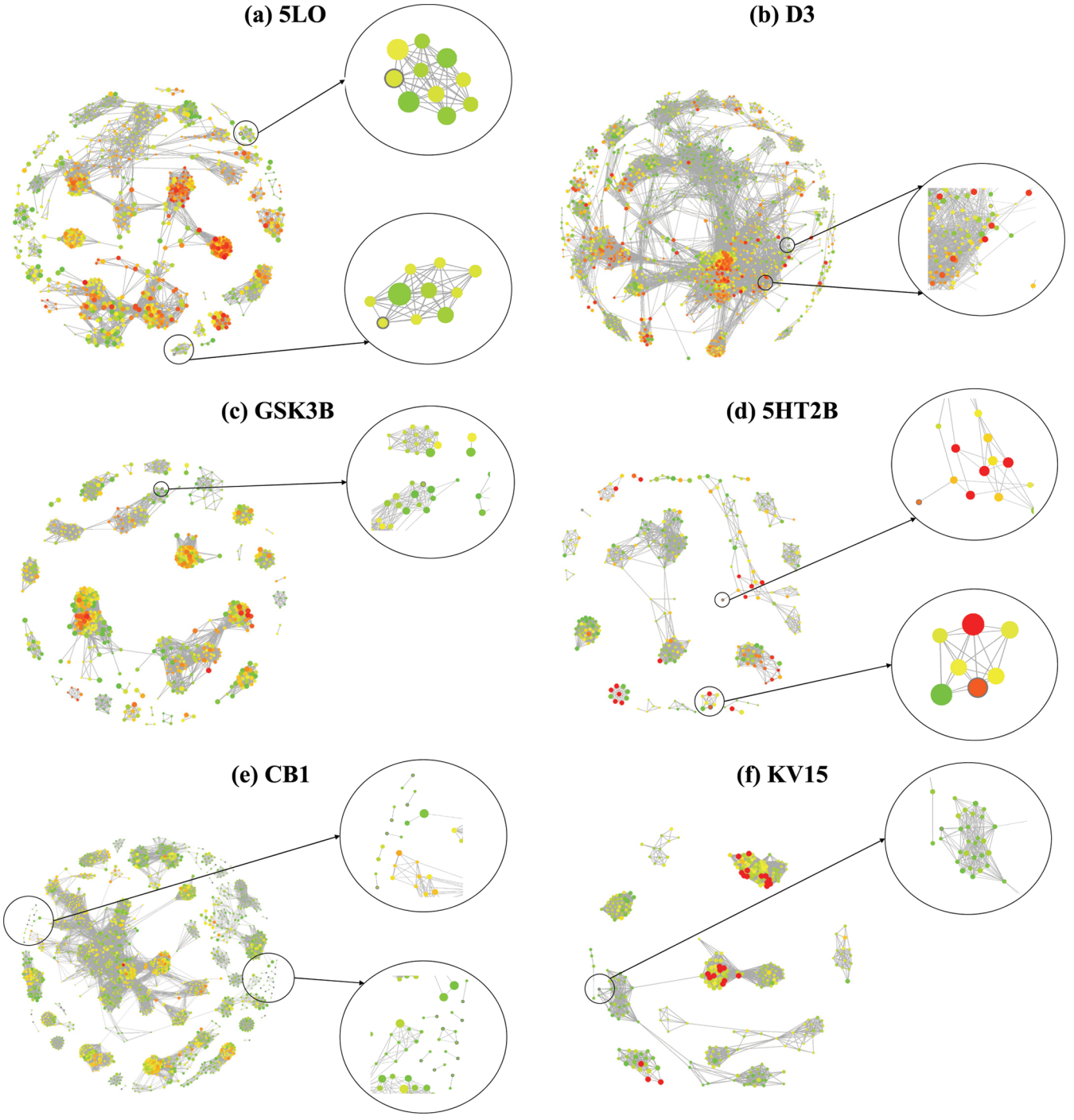


Figure 1. Activity landscapes. NSG representations are shown for six data sets. Nodes are compounds and are color coded according to potency and scaled in size according to per compound discontinuity scores (i.e., the contribution of individual compounds to local SAR discontinuity). Edges indicate pairwise similarity relationships. Combinations of large red and green nodes indicate activity cliff regions and combinations of small yellow/orange/green nodes regions of SAR continuity. Nodes representing newly identified LBVS hits are displayed with a border, and NSG regions containing these hits are encircled and zoomed to reveal further details: (a) 5LO, (b) D3, (c) GSK3B, (d) 5HT2B, (e) CB1, and (f) KV15.

RESULTS AND DISCUSSION

Study Rationale. LBVS ultimately aims at the identification of new active chemotypes, and it is intuitive that promising LBVS targets should display a significant degree of ligand permissiveness, or promiscuity, and SAR tolerance. It should be noted that SAR tolerance of targets goes beyond commonly understood ligand promiscuity; it implies that a target binds to different series of compounds with well-defined chemical changes representing different SARs, i.e., it is an expression of SAR continuity at the target level. However, in the context of LBVS, which does not take target structure or binding site information into account, the hypothesis that SAR tolerance is a prerequisite for successful virtual screening and scaffold hopping cannot be directly

dx.doi.org/10.1021/ci200064c |J. Chem. Inf. Model. 2011, 51, 837–842

assessed. However, information about target-dependent SAR tolerance is implicitly encoded in available active compounds in the form of SAR continuity. Therefore, the SAR information contained in such data sets might reveal how permissive a given target is for alternative binding events involving structurally diverse compounds. Hence, activity landscapes of promising LBVS targets should reveal regions of distinct SAR continuity. Furthermore, if there would indeed be a firm relationship between SAR tolerance at the target level and SAR continuity at the ligand level, then newly identified LBVS hits should also have a much higher probability to fall into continuous SAR regions, as opposed to discontinuous ones. These questions have been addressed in our analysis.

Targets and LBVS Methods. Table 1 lists the LBVS targets for which known ligand sets and newly identified hits were analyzed as well as the LBVS methods that were applied. These targets include a variety of functionally distinct enzymes, receptors, and transporters that interact with very different classes of compounds. Thus, there were no apparent target family or compound class preferences. Furthermore, different types of LBVS methods were successfully utilized in these investigations including two-dimensional (2D) similarity and 3D pharmacophore searching and machine learning techniques. Hence, the success of these LBVS campaigns was also not dependent on the application of a single class of methods, as to be expected.

Activity Landscape Analysis. In Figure 1, representative NSGs are shown for six compound data sets. The NSGs of the remaining 10 sets are provided in Figure S1 of the Supporting Information. Data sets from which activity landscapes were generated contained between 103 and 1799 compounds. Thus, these compound sets were generally large enough to yield informative landscape views. The structures of all hits that were mapped in Figure 1 are presented in Figure S2 of the Supporting Information. With the exception of target CB1, for which an unusually large number of 27 confirmed LBVS hits was found, between one and three qualifying hits were available for mapping in each case.

All six landscapes in Figure 1 have distinctly different topology and composition. In addition, with the exception of the D3 landscape, all landscapes contain a relatively large number of individual densely connected clusters, i.e., compound series without notable similarity to other ones, which is a general indicator of structural diversity within an activity landscape. Furthermore, in all landscapes, multiple continuous and discontinuous local SAR regions can be seen, although their relative proportion in part substantially varies. Discontinuous local SAR regions are characterized by the presence of large nodes and activity cliffs (i.e., connected pairs of large red and green nodes), whereas continuous SAR regions are characterized by the prevalence of small nodes of low (green) to moderate (yellow) potency or, alternatively, equally low or high (red) potency. Within compound clusters, SAR continuity results from the presence of compounds whose structural neighbors have similar potency. Importantly, the presence of multiple such clusters means that many structurally diverse compounds with similar potency are available, which is indicative of a high degree of SAR continuity and scaffold hopping potential (it should be noted that interclusters distances in NSGs are determined by the graphical layout algorithm and have no chemical meaning). Thus, activity landscapes having such features would be characteristic of promising LBVS targets.

Table 2. SARI Discontinuity Scores of Virtual Screening Hits[a]

target	discontinuity score
(a) 5-lipoxygenase (5-LO)	hit 1: 0.51
	hit 2: 0.15
(b) dopamine D3 receptor (D3 receptor)	hit 1: 0.44
	hit 2: 0.37
(c) glycogen synthase kinase-3, beta (GSK3B)	hit: 0.25
(d) serotonin receptor 2B (5HT2B)	hit 1: 0.28
	hit 2: 0.81
(e) cannabinoid receptor 1 (CB1)	hits (average value): 0.03 (SD: 0.02)
(f) voltage-gated potassium channel subunit Kv1.5 (KV15)	hit: 0.16

[a] For all virtual screening hits in Figure 1, SARI per compound discontinuity scores are reported. Under "target", a–f refers to the corresponding panels in Figure 1. Multiple hits are numbered here from the top (hit 1) to bottom (hit 2) according to their location in Figure 1. For CB1 (27 hits), the average per hit discontinuity score and the standard deviation (SD) are reported instead of individual scores.

The 5LO landscape in Figure 1a contains many different continuous regions and only little local SAR discontinuity. In this case, two moderately potent hits (yellow nodes) are available that map to different continuous SAR regions that are formed by relatively few compounds. Thus, both hits have only a limited number of structural neighbors. Different from 5LO, the D3 landscape in Figure 1b, representing a large compound set, contains a large and densely connected central graph component with many activity cliffs. However, also in this case, the two LBVS hits are found within the largely continuous periphery of the central graph component. The GSK3B landscape in Figure 1c presents yet another phenotype. Here many different compound clusters exist that represent continuous or discontinuous local SARs. Thus, this activity landscape is characterized by a high degree of SAR heterogeneity. Again, the new LBVS hit is found in a region of high continuity. A similar degree of SAR heterogeneity is observed in the activity landscape of the 5HT2B data set in Figure 1d. In this case, three hits are available. Two of them have only one structural neighbor and hence do not contribute much SAR information, but the third (relatively potent) hit maps to a small discontinuous cluster where it forms a moderately sized activity cliff with another compound. Thus, this hit introduces notable local SAR discontinuity and represents an exceptional case in our analysis. The CB1 landscape in Figure 1e represents the largest data set and contains many continuous and moderately discontinuous regions. In this case, 27 hits are available, many of which are found to have only one or a few structural neighbors. Thus, despite the large number of available active compounds, this LBVS application has added considerable structural novelty to the exiting spectrum of CB1 ligands. Here we also find that LBVS hits having multiple structural neighbors map to continuous SAR regions. In contrast to CB1, the activity landscape of the KV15 data set in Figure 1f is largely dominated by three regions of strong SAR discontinuity that contain many activity cliffs. However, coexisting regions of local SAR continuity are also observed, and the single available LBVS hits map to one of these continuous regions.

Table 2 reports the per compound SARI discontinuity scores[10,11] for all hits shown in Figure 1. With one exception

(5HT2B hit 2 in Figure 1d (bottom); as discussed), the scores reflect the presence of a low degree of SAR discontinuity, consistent with the conclusions drawn from graphical analysis. Thus, in summary, the activity landscapes studied here contain multiple continuous SAR regions, and nearly all of the newly identified LBVS hits are found in regions of distinct SAR continuity.

CONCLUDING REMARKS

Herein we have assessed the as of yet unsupported assumption that LBVS success should depend on the SAR tolerance of screening targets. In the context of LBVS, this assumption cannot be directly assessed. However, in ligands sets, SAR tolerance is reflected by the presence of SAR continuity, which can be analyzed in activity landscape representations. Activity landscapes of targets that are amenable to LBVS should contain extensive regions of local SAR continuity. Furthermore, if there is a link between SAR tolerance and LBVS success, then it also follows that LBVS hits should have a much higher probability to map to continuous regions of target-specific activity landscapes than to discontinuous ones. In regions of SAR discontinuity, molecular similarity does not correlate with activity similarity and different chemotypes with similar activity are difficult to find. In order to address these questions, we have generated activity landscapes for different targets for which successful LBVS campaigns have been reported. These activity landscapes were found to contain multiple regions of SAR continuity, although their topology and the balance between SAR continuity and discontinuity in part substantially differed. Newly identified LBVS hits were almost always found to map to continuous SAR regions. Thus, taken together, our findings provide support for the proposed relationship between SAR tolerance, SAR continuity, and the likelihood of success in prospective LBVS applications.

ASSOCIATED CONTENT

Supporting Information. Figure S1 shows NSG representations according to Figure 1 for all remaining data sets in Table 1. Figure S2 shows the structures of all virtual screening hits that were mapped in Figure 1. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: bajorath@bit.uni-bonn.de.

ACKNOWLEDGMENT

The authors thank Anne Mai Wassermann for many helpful discussions. P.R. is supported by a post graduate program of the Deutsche Forschungsgemeinschaft (GRK 804) and M.W. by Boehringer Ingelheim Pharma GmbH & Co. KG.

REFERENCES

(1) Geppert, H.; Vogt, M.; Bajorath, J. Current Trends in Ligand-based Virtual Screening: Molecular Representations, Data Mining Methods, New Application Areas, and Performance Evaluation. *J. Chem. Inf. Model.* **2010**, *50*, 205–216.

(2) Eckert, H.; Bajorath, J. Molecular Similarity Analysis in Virtual Screening: Foundations, Limitations and Novel Approaches. *Drug Discovery Today* **2007**, *12*, 225–233.

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

(4) Ripphausen, P.; Nisius, B.; Bajorath, J. State-of-the-Art in Ligand-based Virtual Screening. *Drug Discovery Today*, in press, doi:10.1016/j.drudis.2011.02.011.

(5) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.

(6) *ChEMBL*; European Bioinformatics Institute (EBI): Cambridge, U.K., 2010; http://www.ebi.ac.uk/chembl/. (Accessed January 10, 2011).

(7) Willett, P. Searching Techniques for Databases of Two- and Three-dimensional Structures. *J. Med. Chem.* **2005**, *48*, 1–17.

(8) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.

(9) *Scitegic Pipeline Pilot*, student ed.; Accelrys, Inc.: San Diego, CA, 2010.

(10) Wawer, M.; Bajorath, J. Extracting SAR Information from a Large Collection of Anti-Malarial Screening Hits by NSG-SPT Analysis. *ACS Med. Chem. Lett.* **2011**, *2*, 201–206.

(11) Wawer, M.; Peltason, L.; Weskamp, N.; Teckentrup, A.; Bajorath, J. Structure-Activity Relationship Anatomy by Network-like Similarity Graphs and Local Structure-Activity Relationship Indices. *J. Med. Chem.* **2008**, *51*, 6075–6084.

(12) Peltason, L.; Bajorath, J. SAR Index: Quantifying the Nature of Structure-Activity Relationships. *J. Med. Chem.* **2007**, *50*, 5571–5578.

(13) Fruchterman, T. M. J.; Reingold, E. M. Graph Drawing by Force-Directed Placement. *Software — Practice and Experience* **1991**, *21*, 1129–1164.

(14) Csardi, G. *igraph library*, version 0.5.5; XXX: Budapest, Hungary, 2009; http://igraph.sourceforge.net/. (Accessed December 16, 2010).

(15) Bates, D.; Chambers, J.; *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2010.

(16) Franke, L.; Schwarz, O.; Müller-Kuhrt, L.; Hoernig, C.; Fischer, L.; George, S.; Tanrikulu, Y.; Schneider, P.; Werz, O.; Steinhilber, D.; Schneider, G. Identification of Natural-Product-Derived Inhibitors of 5-Lipoxygenase Activity by Ligand-Based Virtual Screening. *J. Med. Chem.* **2007**, *50*, 2640–2646.

(17) Chaudhaery, S. S.; Roy, K. K.; Shakya, N.; Saxena, G.; Sammi, S. R.; Nazir, A.; Nath, C.; Saxena, A. K. Novel Carbamates as Orally Active Acetylcholinesterase Inhibitors Found to Improve Scopolamine-Induced Cognition Impairment: Pharmacophore-Based Virtual Screening, Synthesis, and Pharmacology. *J. Med. Chem.* **2010**, *53*, 6490–6505.

(18) Neves, M. A. C.; Dinis, T. C. P.; Colombo, G.; Sa e Melo, M. L. An Efficient Steroid Pharmacophore-Based Strategy to Identify New Aromatase Inhibitors. *Eur. J. Med. Chem.* **2009**, *44*, 4121–4127.

(19) Foloppe, N.; Benwell, K.; Brooks, T. D.; Kennett, G.; Knight, A. R.; Misra, A.; Monck, N. J. T. Discovery and Functional Evaluation of Diverse Novel Human CB1 Receptor Ligands. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4183–4190.

(20) Moree, W. J.; Kataoka, K.; Ramirez-Weinhouse, M. M.; Shiota, T.; Imai, M.; Sudo, M.; Tsutumi, T.; Endo, N.; Muroga, Y.; Hada, T.; Tanaka, H.; Morita, T.; Greene, J.; Barnum, D.; Saunders, J.; Kato, Y.; Myers, P. L.; Tarby, C. M. Small Molecule Antagonists Of the CCR2b Receptor. Part 2: Discovery Process and Initial Structure-Activity Relationships of Diamine Derivatives. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5413–5416.

(21) Byvatov, E.; Sasse, B. C.; Stark, H.; Schneider, G. From Virtual to Real Screening for D3 Dopamine Receptor Ligands. *ChemBioChem* **2005**, *6*, 997–999.

(22) Enyedy, I. J.; Sakamuri, S.; Zaman, W. A.; Johnson, K. M.; Wang, S. Pharmacophore-Based Discovery of Substituted Pyridines as Novel Dopamine Transporter Inhibitors. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 513–517.

(23) Funk, O. F.; Kettmann, V.; Drimal, J.; Langer, T. Chemical Function Based Pharmacophore Generation of Endothelin-A Selective Receptor Antagonists. *J. Med. Chem.* **2004**, *47*, 2750–2760.

(24) Narum., L.; Norskov-Lauritsen, L.; Olesen, P. Scaffold Hopping and Optimization towards Libraries of Glycogen Synthase Kinase-3 Inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1525–1528.

(25) Shoda, M.; Harada, T.; Kogami, Y.; Tsujita, R.; Akashi, H.; Kouji, H.; Stahura, F. L.; Xue, L.; Bajorath, J. Identification of Structurally Diverse Growth Hormone Secretagogue Agonists by Virtual Screening and Structure-Activity Relationship Analysis of 2-Formylaminoacetamide Derivates. *J. Med. Chem.* **2004**, *47*, 4286–4290.

(26) Al-Sha'er, M. A.; Taha, M. O. Elaborate Ligand-Based Modeling Reveals New Nanomolar Heat Shock Protein 90α Inhibitors. *J. Chem. Inf. Model.* **2010**, *50*, 1706–1723.

(27) Tang, H.; Wang, X. S.; Huang, X.-P.; Roth, B. L.; Butler, K. V.; Kozikowski, A. P.; Jung, M.; Tropsha, A. Novel Inhibitors of Human Histone Deacetylase (HDAC) Identified by QSAR Modeling of Known Inhibitors, Virtual Screening, and Experimental Validation. *J. Chem. Inf. Model.* **2009**, *49*, 461–476.

(28) Hajjo, R.; Grulke, C. M.; Golbraikh, A.; Setola, V.; Huang, X.-P.; Roth, B. L.; Tropsha, A. Development, Validation, and Use of Quantitative Structure-Activity Relationship Models of 5-Hydroxytryptamine (2B) Receptor Ligands to Identify Novel Receptor Binders and Putative Valvulopathic Compounds among Common Drugs. *J. Med. Chem.* **2010**, *53*, 7573–7586.

(29) Koide, Y.; Uemoto, K.; Hasegawa, T.; Sada, T.; Murakami, A.; Takasugi, H.; Sakurai, A.; Mochizuki, N.; Takahashi, A.; Nishida, A. Pharmacophore-Based Design of Sphingosine-3 Receptor Antagonists That Include a 3,4-Dialkoxybenzophenone Scaffold. *J. Med. Chem.* **2007**, *50*, 442–454.

(30) Lescot, E.; Sopkova-de Oliveira Santos, J.; Dubessy, C.; Oulyadi, H.; Lesnard, A.; Vaudry, H.; Bureau, R.; Rault, S. Definition of New Pharmacophores for Nonpeptide Antagonists of Human Urotensin-II. Comparison with the 3D-structure of Human Urotensin-II and URP. *J. Chem. Inf. Model.* **2007**, *47*, 602–612.

(31) Peukert, S.; Brendel, J.; Pirard, B.; Strübing, C.; Kleemann, H.-W.; Böhme, T.; Hemmerle, H. Pharmacophore-Based Search, Synthesis, and Biological Evaluation of Anthranilic Amides as Novel Blockers of the Kv1.5 Channel. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2823–2827.