Journal of Medicinal Chemistry

Perspective

pubs.acs.org/jmc

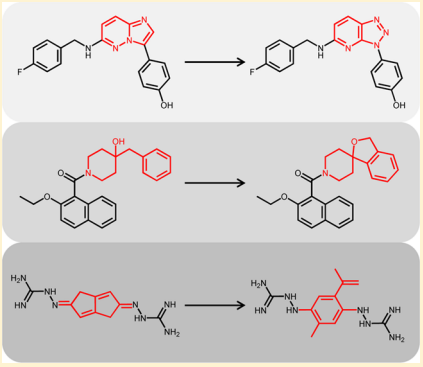
Recent Advances in Scaffold Hopping

Miniperspective

Ye Hu, Dagmar Stumpfe, and Jürgen Bajorath*

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

ABSTRACT: Scaffold hopping refers to the computer-aided search for active compounds containing different core structures, which is a topic of high interest in medicinal chemistry. Herein foundations and caveats of scaffold hopping approaches are discussed and recent methodological developments analyzed. Despite the conceptual prevalence of pharmacophore methods for scaffold hopping, a variety of computational approaches have been successfully applied. In recent years, scaffold hopping calculations are increasingly carried out at the level of scaffolds rather than compounds, and scaffold queries increasingly abstract from chemical structures. In addition, relationships between compounds, scaffolds, and biological activities are beginning to be globally explored, beyond individual applications. Going forward, computational scaffold hopping is thought to benefit from the consideration of new scaffold concepts and the development of methods capable of guiding search calculations toward scaffolds that are likely to represent potent compounds.



INTRODUCTION

The expression *scaffold hopping* was coined by Schneider and colleagues in 1999[1] and has been widely used ever since. It intuitively refers to the quest for compounds with different structures but similar activity, which continues to be a major topic in medicinal chemistry, for several reasons. The term *scaffold* focuses this quest on *core structures* of compounds, which is a key aspect of the exercise and a potential caveat at the same time. Importantly, scaffold hopping calculations have thus far mostly been carried out by comparing compounds, but the results are evaluated by comparing the scaffolds that candidate compounds contain. As discussed in a recent perspective on the scaffold concept in medicinal chemistry,[2] molecular cores aka scaffolds can be defined in different ways and the use of the term scaffold in the literature often remains ambiguous. Thus, for the assessment of scaffold hopping, clear definitions are required and must be consistently applied.

Scaffold hopping is preferentially viewed in the context of computational (virtual) compound screening, for which the original work of Schneider et al. has set the stage. However, scaffold replacements can also be attempted on a case-by-case basis from a chemical viewpoint, for example, by ingenious design or systematic chemical modifications of core structures representing compound series. Such modifications might include, for example, replacements of heterocyclic rings and ring closure or opening reactions.[3]

For systematic scaffold hopping applications, computational approaches are essential. First and foremost, computational methods are employed that extrapolate from structures of known reference molecules and attempt to depart from them to identify compounds containing different scaffolds. Such methods explore and exploit the concept of molecular similarity and dissimilarity in different ways.[4] Alternatively, docking methods can also be employed to search for novel active compounds that structurally depart from known ones,[5] which represents an indirect approach to scaffold hopping.

In the practice of medicinal chemistry, scaffold hopping studies are carried out for different reasons. For example, one might be interested in circumventing an intellectual property position by identifying novel chemical entities having a desired activity, replacing a chemically complex natural product with a synthetically accessible molecule, or improving pharmacological properties of known actives.

The focus of this contribution is on the evaluation of recent methodological developments for scaffold hopping rather than case studies (nonetheless, exemplary case studies are described). Initially, foundations of scaffold hopping are discussed.

CLASSICS

Pharmacophores. A variety of computational methods have been adapted or developed for scaffold hopping applications.[6] Among these, approaches based upon the pharmacophore concept have traditionally been preferred,[1,7] for good reasons. The basic definition of a pharmacophore as the (spatial) arrangement of atoms or groups of atoms in a molecule that determine its bioactivity implies that critical contacts should mostly be formed by R-groups, although scaffold atoms can certainly participate in such interactions. Hence, the underlying premise of scaffold hopping via pharmacophore searching is the following. If pharmacophores

Received: September 27, 2016
Published: December 21, 2016

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

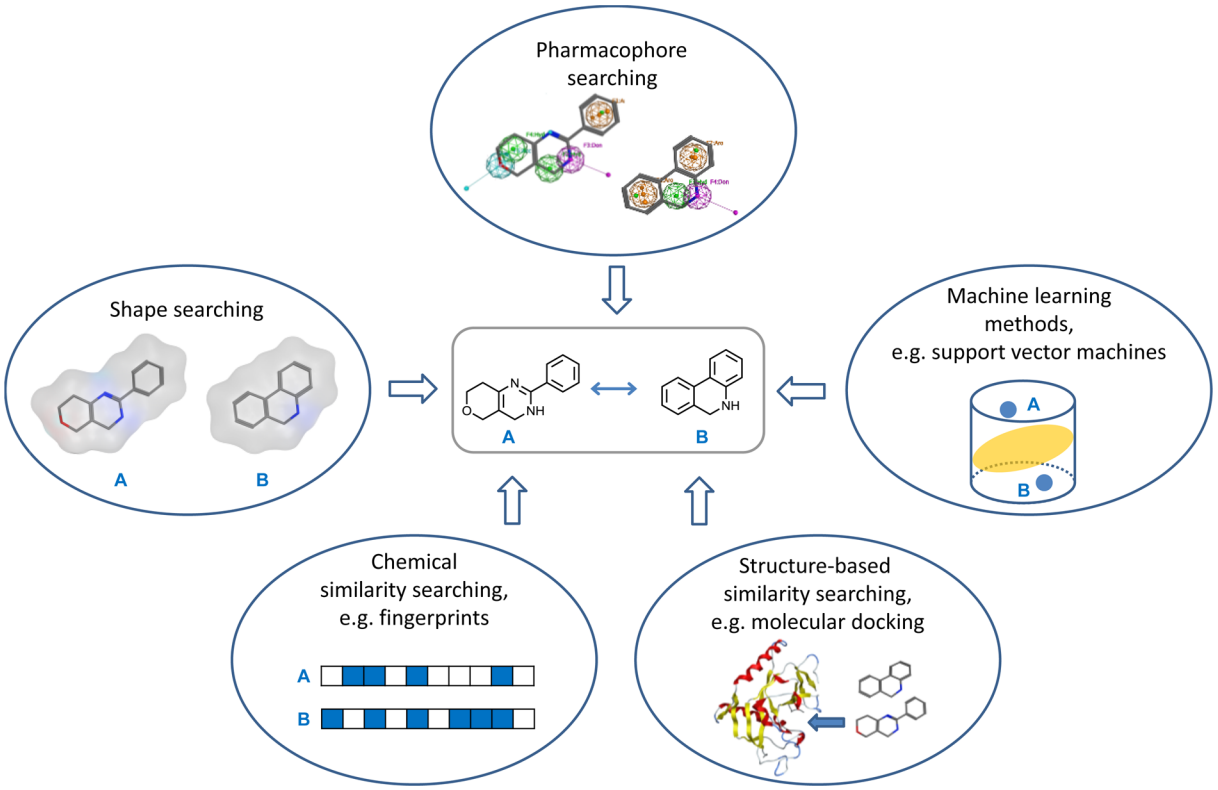


Figure 1. Computational methods. Exemplary computational approaches for scaffold hopping are shown.

can be transferred from one compound to another, i.e., if they are conserved in reference and test compounds, scaffolds representing them can be replaced. Essentially the same concept can be applied to structure-based virtual screening as exemplified by the pioneering development of the CAVEAT program.[8] In this case, key interactions between ligands and residues in protein binding sites were represented as vectors and searches for alternative scaffolds carried out that had corresponding attachment points and were capable of presenting interaction vectors in corresponding positions. It is important to note that pharmacophore features can be represented in rather different ways, for example, as 3D or 2D models. In addition, pharmacophore descriptors of different design and greatly varying complexity are available. Regardless of specific representations, pharmacophore-based approaches share a local assessment of similarity that is confined to predefined pharmacophore features.

Despite the conceptual link between pharmacophores and scaffold hopping, other computational approaches have also been successfully applied. For example, shape searching is a promising approach to scaffold hopping.[9] Furthermore, a variety of machine learning methods including self-organizing maps, which also enable visualization of compound distributions, have been adapted for scaffold hopping.[10,11] Moreover, even chemical similarity searching using simple 2D fingerprints, which assesses whole-molecule similarity and disregards pharmacophores, succeeds in scaffold hopping, provided that the characteristics of these search calculations are carefully considered and candidate compounds are selected accordingly.[12] Different computational approaches for scaffold hopping are schematically represented in Figure 1.

Formalized Scaffolds. If scaffold hopping is attempted computationally, a formal scaffold definition is required that can be consistently applied. The definition of scaffolds according to Bemis and Murcko (BM scaffolds) introduced 20 years ago has become a hallmark for computational analysis.[13] BM scaffolds are extracted from compounds by removal of all substituents while retaining ring systems and linker moieties between rings. From a chemical perspective, this definition has shortcomings.[2] However, it has provided the basis for a systematic assessment of the scaffold hopping potential of computational methods, in benchmark comparisons and beyond. In addition, the BM scaffold concept has been further extended in different ways that are also relevant for scaffold hopping. For example, scaffold networks and tree structures have been introduced that generate hierarchies of BM scaffolds and virtual scaffolds derived from them.[14,15] The HierS method decomposes BM scaffolds into all possible ring fragments and organizes fragments and parent scaffolds in network representations.[14] Furthermore, the Scaffold Tree algorithm[15] applies predefined structural rules to systematically decompose BM scaffolds and captures resulting scaffold pathways in tree structures. Scaffold pathways overlap at shared decomposition products. Virtual scaffolds that are neighbors of BM scaffolds from active compounds can be prioritized as scaffold hops and compounds containing these scaffolds evaluated.[16]

In addition to BM scaffolds, other scaffold definitions are also considered in medicinal chemistry,[2] but they are less relevant for computational analysis and scaffold hopping assessment.

HOPPING, LEAPING, OR CRAWLING?

Formally defined scaffold hopping events are often of different magnitude. For example, on the basis of their definition, BM scaffolds might be very similar, e.g., they might only be distinguished by a heteroatom substitution in a ring. Alternatively, BM scaffolds might be completely distinct, e.g.,

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

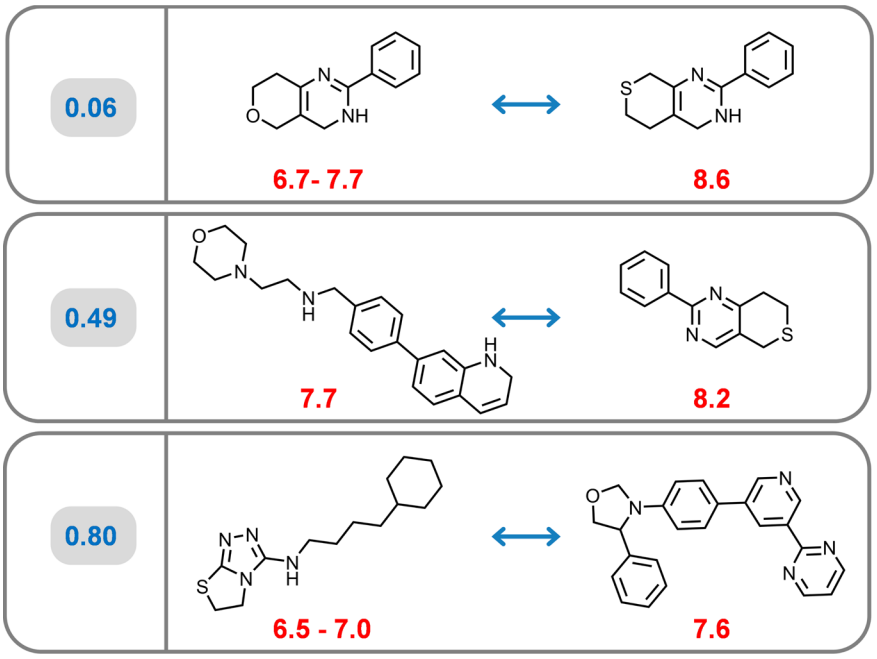


Figure 2. Exemplary scaffold hops. Three scaffold hops with increasing dissimilarity between scaffolds (top to bottom) are shown. The scaffolds were extracted from tankyrase-2 inhibitors. At least one inhibitor represented by each scaffold had nanomolar potency. Blue numbers give scaffold distances for each pair according to ref 17, and red numbers report the range of pIC_{50} values for all compounds represented by a scaffold.

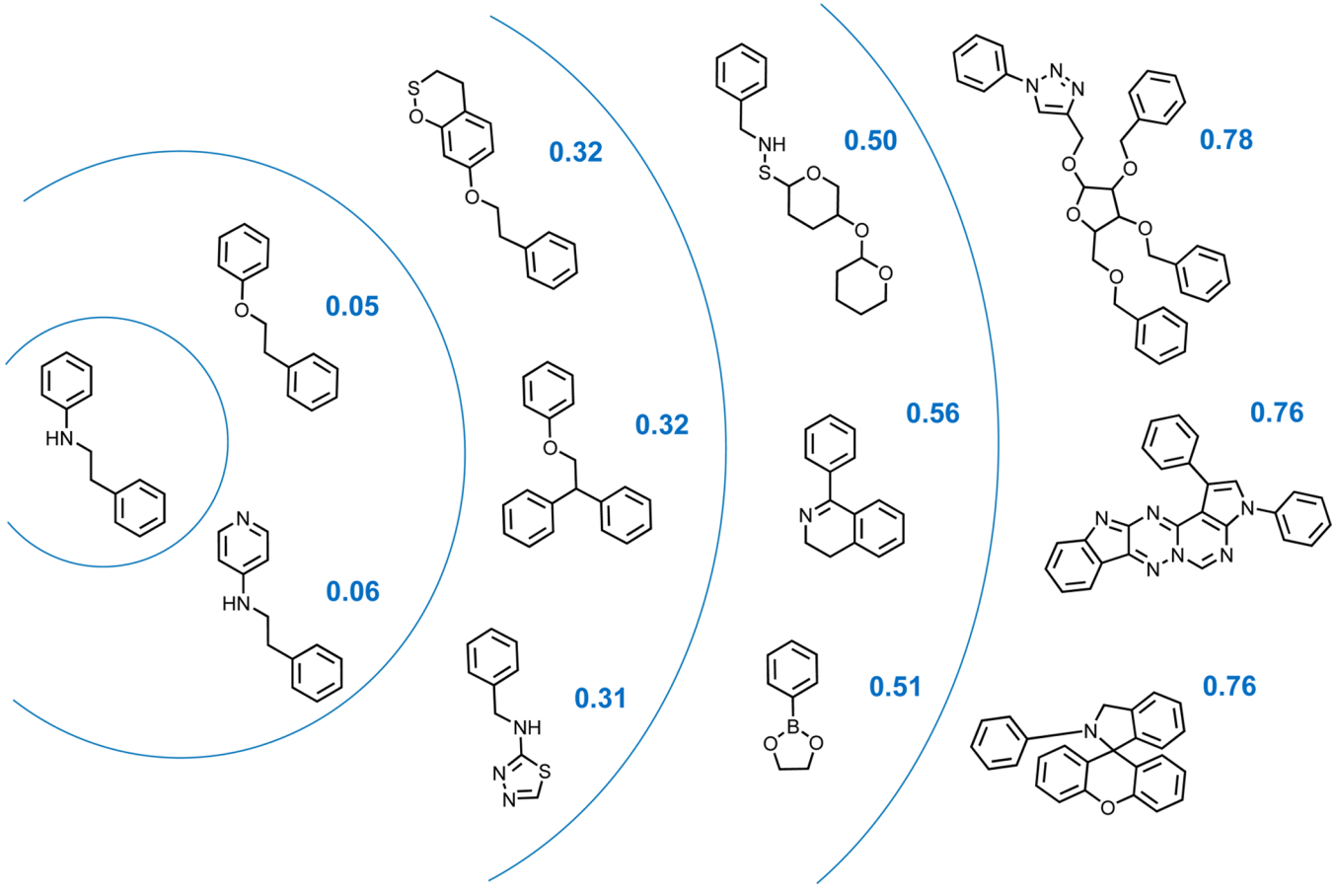


Figure 3. Scaffold spectrum. Shown are BM scaffolds from carbonic anhydrase II inhibitors that are increasingly dissimilar to a reference scaffold (encircled on the left). Inhibitors represented by these scaffolds had comparable mean logarithmic potency (pK_i) values between 7.3 and 8.6. Numbers provide scaffold distances for pairwise comparison to the reference scaffold.

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

consist of different ring systems with different topology. In addition, BM scaffolds might exhibit substructure relationships because (by definition) the addition of a ring to a BM scaffold creates a new one. From a chemical perspective, such scaffold relationships might often be debatable. From a computational viewpoint, they complicate the assessment of scaffold hopping events. Figure 2 shows exemplary scaffold hops that involve very similar (top), remotely similar (middle), or distinct scaffolds. Remotely similar scaffolds might, for example, share some of their structural elements but have different topologies. Detecting compounds that contain such distantly related scaffolds but share similar activity would be considered a meaningful scaffold *hopping* event. In such cases, there is at least some chemical resemblance but the scaffold hop would be difficult to predict. By contrast, detecting very similar scaffolds such as the ones shown at the top of Figure 2 also formally qualifies as a scaffold hop. However, from a computational viewpoint, simple similarity calculations might suffice in such instances for *crawling* from one scaffold to the other. By contrast, for recognizing activity relationships between compounds containing structurally distinct scaffolds (bottom of Figure 2), which are impossible to predict via a chemist's eye, hopping might not be sufficient. Instead, more challenging computational *leaping* might be required.

These semantic connotations are meant to illustrate the critical issue that scaffold hopping covers a wide spectrum of structural relationships, ranging from very close ones, which are trivial to detect, to distant or nonexisting ones, for which activity relationships would be impossible to predict. For a systematic assessment of scaffold hopping beyond subjective judgment this presents a problem. Therefore, a dissimilarity function was introduced to quantify chemical distance relationships between scaffolds.[17] Taking the composition and topology of ring systems comprising BM scaffolds into account, pairs of scaffolds were subjected to iterative editing procedures and transformed into one-dimensional atom sequences that were then compared using sequence alignment methods. Hence, the resulting scaffold distances ranged from 0 to 1 and were a measure of dissimilarity. In Figure 2, pairwise scaffold distances are reported for the exemplary scaffold hops. Figure 3 shows scaffolds of increasing structural diversity compared to a reference scaffold and reports corresponding distance values. The methodology was applied to monitor scaffold distances in 790 compound activity classes, leading to the prioritization of 50 classes for scaffold hopping analysis and benchmark calculations.[18] On the basis of systematic scaffold distance comparisons across many activity classes, distance threshold values were established indicating similar ($\leq$0.34) and dissimilar ($\geq$0.74) scaffolds, respectively.[18] The computational method and the data sets prioritized for scaffold hopping analysis have been made publicly available.

PREREQUISITES OF SUCCESS

Another fundamental question concerning scaffold hopping is why it frequently works (excluding trivial cases). Why are similarity-based computational methods capable of detecting compounds with different core structures sharing the same activity, although activity is not considered as a search parameter? Insights into such questions were provided by scaffold diversity analysis across bioactive compounds.[19] From nearly 500 compound activity classes, BM scaffolds were extracted and compounds active against nearly 400 targets were found to contain five to 99 different scaffolds per target set. In addition, compounds active against 28 other targets contained 100 or more scaffolds.[19] These findings were corroborated by a recent analysis of target promiscuity that monitored the distribution of scaffolds over pharmaceutical targets.[20] The study revealed that for the majority of current targets, active compounds are available that contain large numbers of scaffolds,[20] as shown in Figure 4. Thus, one of the foundations

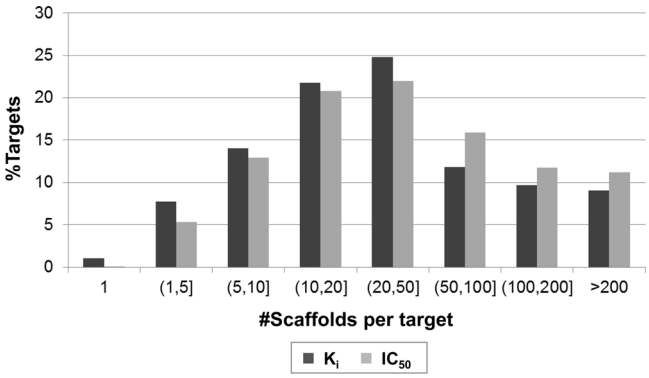


Figure 4. Scaffold–target distribution. Monitored is the percentage of targets for which active compounds contain increasing numbers of unique BM scaffolds. The distribution was generated on the basis of high-confidence activity data extracted from ChEMBL release 21. K_i (dark gray) and IC_{50} (light gray) measurements were separately considered and scaffolds isolated from 464 (K_i) and 980 (IC_{50}) target sets with at least 10 compounds.

of scaffold hopping is the ability of many targets to specifically interact with structurally diverse compounds. Given this ability, there should a priori be a high probability of identifying scaffold hops for many targets, irrespective of the computational methods that are applied. In cases where known active compounds already contain an abundance of different scaffolds, it should not be very difficult to identify additional scaffold hops. Perhaps one frequently overestimates the challenges of scaffold hopping, which might also explain why many different computational methods are successfully applied. On the other hand, a given scaffold hop is not the same as any other and the identification of distantly related or distinct scaffolds representing active compounds remains challenging. It is also noted that computational and chemical viewpoints on scaffold hopping do not necessarily coincide. Chemists might not consider the ability of many targets to recognize structurally diverse compounds when confronted with a scaffold hopping challenge. Rather, they will primarily concentrate on transforming a given active compound into another that sufficiently departs from the original structure, for example, to establish a new patent position.

NEW DEVELOPMENTS

Having discussed some of the foundations and caveats of scaffold hopping, we focus on recently introduced concepts and computational methods.

Hopping from Natural Products. Given the renaissance of natural products in drug discovery, many attempts are being made to replace active natural compounds with synthetic mimics that are easier to access chemically. To these ends, scaffold hopping is an attractive approach. In an exemplary investigation, the peptidic scaffold of the natural product belactosin A, a proteasome inhibitor, was replaced with nonpeptidic scaffolds by so-called *topology-based scaffold*

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238–1246

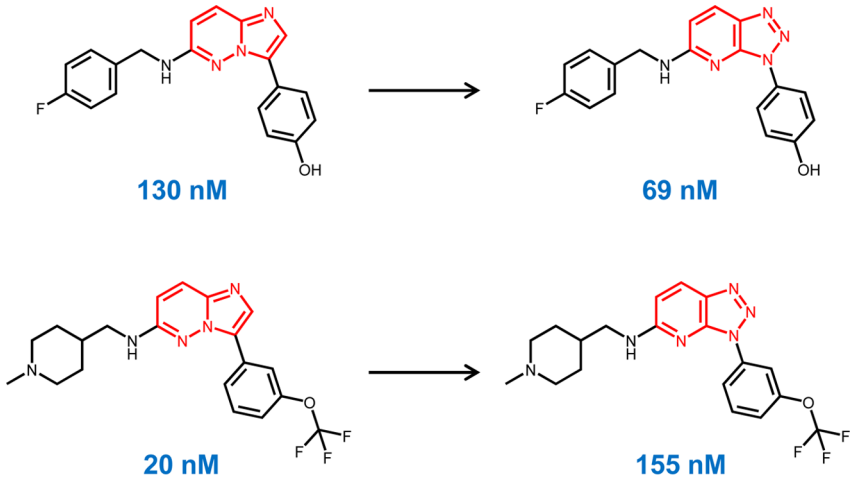


Figure 5. PIM-1 kinase inhibitors. Shown are two new PIM-1 kinase inhibitors (right) in which the imidazopyridazine scaffold contained in known inhibitors (left) was replaced by triazolopyridine.[20] The scaffold replacement is highlighted in red. For each inhibitor, the IC_{50} value is reported.

hopping on the basis of a pharmacophore model derived with the aid of X-ray data.[21] Interestingly, in this case, the scaffold hop was not identified by pharmacophore searching, as is typically attempted, but rather by pharmacophore-guided interactive design. Thus, on the basis of the crystallographic binding mode of a known active compound and the superimposed pharmacophore model, chemical modifications were designed in a stepwise manner to substitute the peptide moiety of the natural product with nonpeptidic groups. The study also nicely illustrates that the BM scaffold definition is not suitable for many compounds from natural sources that contain individual rings and aliphatic or peptidic moieties such as belactosin A.

Fragments, Fingerprints, and Scaffold Similarity. Scaffold hopping techniques that concentrate on compound fragments have also been introduced including a method for scaffold hopping by fragment replacement.[22] Scaffolds were generated by cleaving acyclic bonds of known compounds according to user-defined rules. These graph-based fragment scaffolds were then converted into 3D structures yielding a searchable scaffold conformer database. An indexing scheme was devised accounting for geometrical arrangements of attachment vectors in 3D scaffolds, which rendered search calculations sensitive to the relative orientation of the attachment vectors and enabled fast pruning of the database during fragment scaffold searching. In addition, a scaffold shape descriptor was generated for querying scaffolds with a single attachment vector. The vector-based search, reminiscent of CAVEAT,[8] was used to search for suitable core fragment replacements in active compounds. It was successfully applied to identify a number of bioisosteric fragment scaffolds.

In conceptually related studies, fragment databases were generated as sources for *fragment hopping*.[23,24] Starting from X-ray conformations of reference compounds, 3D similarity (shape) searching was carried out to prioritize candidate fragment scaffolds. In addition, scaffolds matching pharmacophore models were identified and re-evaluated via docking on the basis of X-ray structures. The fragment hopping approach was applied to identify new melanin concentrating hormone antagonists[23] and, more recently, to search for novel inhibitors of the PIM-1 kinase.[24] In the latter case, triazolopyridine was detected as a potential replacement of the imidazopyridazine scaffold present in known inhibitors. Two compounds were synthesized that contained this replacement scaffold but retained the same R-groups as known PIM-1 inhibitors, as shown in Figure 5. These compounds also inhibited PIM-1 kinase.

In another prospective scaffold-centric virtual screening application aiming at identifying new protein kinase TTK inhibitors, level 1 decomposition fragments of BM scaffolds of known inhibitors and database compounds were taken from scaffold trees[15] and used for 2D fingerprint and 3D shape similarity searching.[25] Scaffold-based search calculations led to the identification of several inhibitors containing new scaffolds that were confirmed experimentally. In addition, for scaffold searching, a *quantitative 2D scaffold fingerprint* was introduced that consisted of 1033 bit positions, encoded descriptors of scaffold topology, shape, and pharmacophore features as well as sp^3 carbons, chirality, and attachment points for substituents.[26] It also provided the option of feature weighting. The fingerprint was used to calculate the similarity of query and database scaffolds and identify scaffold hops and bioisosters. In benchmark calculations, scaffold searching achieved higher accuracy than conventional compound-based 2D fingerprint or 3D shape similarity searching.

In a systematic analysis of scaffolds from active compounds, it was found that the majority of scaffolds only contained one or two attachment points and that only less than one-third contained three or four.[27] From the latter subset, a scaffold replacement database with more than 7000 entities was generated. Similarity between query and database scaffolds was calculated on the basis of connectivity descriptors and others to identify bioisosteric scaffold replacements.[27]

Scaffold similarity was also assessed on the basis of a set of 32 *scaffold keys* representing categorized structural patterns derived from scaffolds using topological descriptors.[28] Scaffold keys were prioritized based on medicinal chemistry relevance and used to organize scaffold populations according to the keys they contained. To quantify scaffold similarity, the sum of weighted absolute differences in keys was normalized to a mean of 0 and standard deviation of 1 (with identical scaffolds yielding a value of 0). Scaffold keys were also encoded in a fingerprint format for scaffold similarity searching to identify bioisosteric replacements.[28] In addition, for automated design of bioisosteric replacements, the IADE program was developed and

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

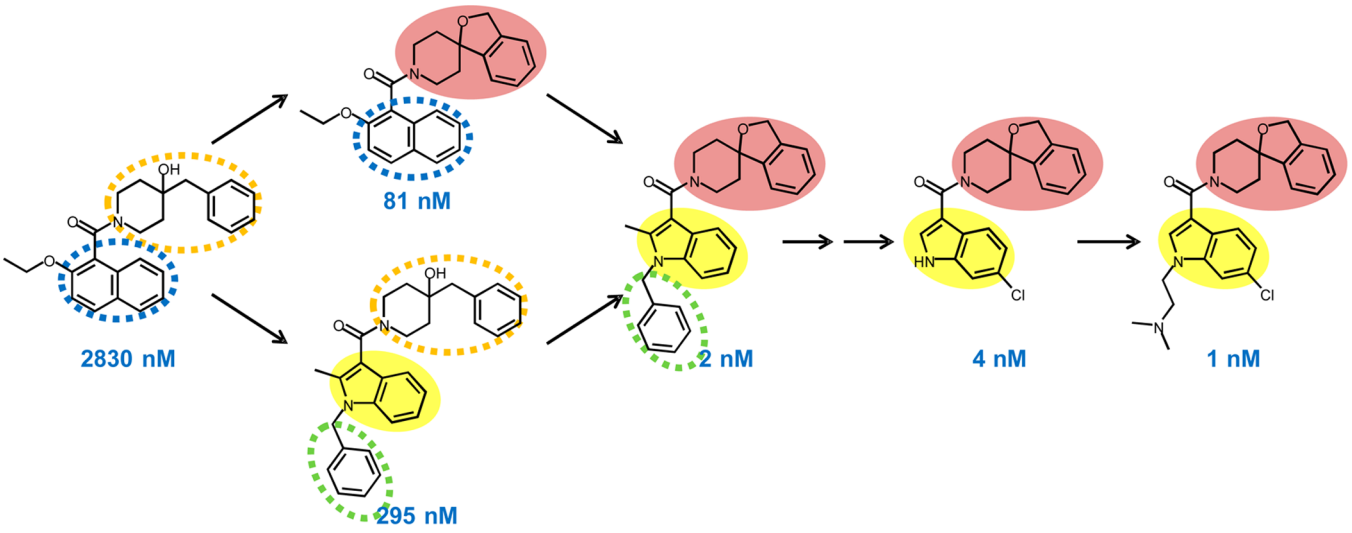


Figure 6. Vasopressin 1a receptor antagonists. Starting from a high-throughput screening hit with micromolar potency against human vasopressin 1a receptor, a series of highly potent antagonists was generated.[28] The optimization involved scaffold replacements. For each antagonist, its K_i value is given. Dashed circles indicate substructures that were replaced during optimization, and newly introduced fragment scaffolds are highlighted in pink and yellow, respectively.

successfully applied to generate a new farnesyltransferase inhibitor.[29]

As a recurrent theme, several of the studies discussed above indicated that similarity searching at the level of scaffolds was often more accurate than calculations using conventional fingerprints encoding entire compounds.

Abstracting from Scaffold Structure. Scaffold keys represent a form of *meta-scaffolds* and provide another layer of structural classification. Other approaches have also been introduced to render search calculations increasingly independent of structural representations. For this purpose, molecular shape is an attractive property. A 3D shape-based similarity method was applied to encode consensus shape patterns of an ensemble of active reference compounds compared to a set of decoys as well as pharmacophore features as a 4D fingerprint.[30] For test compounds, similarity to the ensemble fingerprint was calculated. Furthermore, the approach was combined with ligand docking to assess pharmacophore feature and shape complementarity to a binding site using a specialized scoring function. For the $MycP_1$ protease, a target for tuberculosis treatment, several new, albeit weakly active inhibitors were identified using this approach.[30] Shape searching was also successfully applied to identify novel inhibitors of dual leucine zipper kinase via scaffold hopping.[31]

Shape representations were also an integral component of the so-called *chemical similarity network analysis pull-down 3D* approach that combined 3D similarity measures with network algorithms for structure-based target profiling and automated scaffold hopping.[32] The methodology included a *ShapeAlign* protocol for scaffold hopping that combined 3D shape and pharmacophores with 2D similarity scoring. When applied to six targets and 206 known active compounds, the approach detected scaffold hops that had low 2D chemical similarity but displayed similar binding modes.[32]

Biological Similarity. Scaffold hopping can also be accomplished by completely disregarding structural information of reference compounds, although this might be perceived as being counterintuitive at a first glance. However, the key is exploring biological similarities instead, which can be facilitated at the level of ligands or targets. For example, biological assay profiles of compounds can be compared to identify structurally distinct chemical entities with similar activity. This was demonstrated by the development of a *high-throughput screening fingerprint* capturing nearly 200 target- and cell-based assays for a screening collection of ~1.5 million compounds.[33] Biological fingerprints are compared like structural fingerprints and their overlap is quantified as a measure of biological relatedness. Since there is no structural information included in the assessment of biological activity profiles, compounds with completely distinct structures that share a specific activity can be identified. Another attractive opportunity of this approach is that activity predictions are not limited to those activities comprising the investigated assay profile. However, biological descriptors might also lead to the identification of active compounds with different mechanisms of action (e.g., orthosteric vs allosteric inhibitors). It is also conceivable that compounds with similar biological activity profiles in cell-based assays might be active against different targets.

Going beyond biological similarities of ligands, targets within families can be compared to establish relationships between them. This approach is based on the idea that *similar targets should bind similar ligands*, a concept from chemogenomics. If active small molecules are available for at least some of the targets under comparison, they can be tested against closely related ones to identify new hits or possible scaffold replacements. Such a chemogenomics-type approach was successfully applied in the search for vasopressin 1a antagonists.[34] In this case, scaffold replacements for a weakly active screening hit were identified on the basis of target relationships that ultimately yielded nanomolar leads.[34] Sequence segments comprising the ligand binding sites in G protein-coupled receptors related to vasopressin 1a receptor were encoded using physicochemical descriptors and nearest neighbors were identified by principal component analysis. Candidate compounds were then selected from available antagonists of a closely related receptor, several of which were found to be active against the vasopressin 1a receptor. The compounds provided guidance for scaffold modifications of the screening hit during optimization, as illustrated in Figure 6.[34]

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

The use of ligand or target information to establish biological relationships is not mutually exclusive. For example, proteochemometric methods use descriptors that combine ligand and target information[35] and can also be applied for scaffold hopping. Taken together, approaches exploring biological similarities at the ligand and/or target level can be applied in a variety of ways to identify structurally diverse active compounds.

PERSPECTIVE

The scaffold hopping concept intuitively refers to the quest for compounds with different core structures sharing a specific activity, which is of high interest in medicinal chemistry. Major motivations for scaffold hopping include finding alternative starting points for lead optimization, generating backup candidates for advanced compounds, or establishing intellectual property positions in a crowded therapeutic area. An important aspect of scaffold hopping is that it focuses the search for novel active compounds on molecular core structures. The concept of scaffold hopping has consistently been linked to computational methods and virtual screening, although meaningful scaffold replacements can also be accomplished on the basis of chemical knowledge and intuition. Despite the intrinsic advantage of pharmacophore and shape methods to facilitate scaffold hopping, many different computational approaches have been successfully applied including simple 2D similarity searching. The latter approach is conceptually based on the *similarity-property principle*,[36] stating that *similar structures should have similar activity*. Thus, the identification of structurally distinct scaffolds essentially falls outside its applicability domain.

It is often questionable if complicated and/or multilayered computational approaches are truly required for identifying structurally diverse active compounds, which is a recurrent issue in the assessment of computational studies.[37] The fact that many different computational approaches have been (and continue to be) successfully applied for scaffold hopping can perhaps be more attributed to small molecule binding promiscuity of many pharmaceutical targets than special algorithmic features. Moreover, for the assessment of scaffold hopping potential, no generally accepted standards exist in the computational community. Nonetheless, new computational approaches and protocols continue to be reported specifically for scaffold hopping applications, which further advance the field. In recent years, three trends can be observed. First, search calculations are increasingly performed at the level of scaffolds, rather than compounds, and scaffold searching tends to be more successful than conventional whole-molecule search calculations. Second, it is attempted to increasingly abstract from chemical structure in the search for new scaffolds and bioisosteric replacements. To these ends, relevant approaches make use of, for example, fuzzy pharmacophores[38] and biological descriptors that do not take any structural information into account or combine ligand and target information. Clearly, the fewer the constraints that are put on original structural features, the easier computational hopping or leaping becomes. On the other hand, new compounds with unknown modes of actions might be identified. Third, scaffold relationships and hopping potential are often more globally viewed, beyond individual case studies. Exemplary approaches include large-scale analyses of scaffold−activity relationships,[2] the use of network methods to associate compounds and scaffolds with targets, or chemogenomics-type methods. Such global analysis schemes also provide opportunities to further assess scaffold hopping potential across different protein families and emerging therapeutic targets. Taken together, these trends mirror the evolution of computational scaffold hopping in recent years.

So, where might be room for further improvements and new developments? An "old" issue that still plagues computational predictions including scaffold hopping is that many benchmark studies are not reproducible and/or yield misleading conclusions.[39] In addition, reported performance in benchmark calculations rarely, if ever, scales with success rates in prospective applications. Here, the lack of generally accepted standards for the assessment and publication of test calculations is a major issue and the computational community is challenged to, finally, establish and implement such standards in a concerted manner. This would clearly help to better judge about scaffold hopping potential of different methodologies and reduce the gap between theory and practical applications. More specifically, for computational scaffold hopping, there are at least two opportunities for future developments that are expected to have a major impact on the field, if successfully addressed. One relates to the formal assessment of scaffolds, the other to the computational treatment of relative compound potency. First, although the definition of BM scaffolds has paved the way for systematic scaffold analysis and the assessment of scaffold hopping, it would be beneficial considering alternative scaffold concepts for compound comparison and scaffold hopping that might be more relevant synthetically and more consistent with lead optimization schemes.[40] Second, computationally identified compounds containing new scaffolds are typically weakly potent. This can be explained by the fact that such compounds are hits that provide new starting points for chemical optimization, regardless of whether reference compounds were also screening hits or optimized leads. Importantly, current scaffold hopping approaches do not take compound potency as a parameter into account. Thus, it is neither possible to computationally estimate the potency of compounds resulting from scaffold hopping analysis nor possible to direct a search toward scaffolds having a high probability of representing potent compounds. To these ends, new computational methods are required, which provide substantial opportunities for future research. Both the consideration of alternative scaffold concepts and methods incorporating potency criteria into the search for scaffold replacements would be expected to further increase the impact of scaffold hopping investigations on medicinal chemistry programs.

CONCLUSIONS

Herein scaffold hopping has been discussed from different viewpoints. Emphasis has been put on evaluating foundations of computational scaffold hopping, including the assessment of molecular similarity as well as small molecule binding characteristics of pharmaceutical targets, and on discussing intrinsic limitations that might often not be sufficiently considered. In addition, recent methodological developments have been analyzed that reveal current trends in this field. These include increasing focus on scaffold-based rather than compound-based search calculations, abstraction from chemical structure in formulating scaffold queries, and global scaffold analysis. Furthermore, a perspective on the field has been provided including opportunities for future research such as new scaffold definitions and scaffold hopping methods incorporating compound potency information. These and

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

other extensions of computational scaffold hopping approaches should be attractive goals for computational research, since there are all reasons to believe that scaffold hopping will continue to be of high interest in medicinal chemistry.

AUTHOR INFORMATION

Corresponding Author

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

ORCID

Jürgen Bajorath: 0000-0002-0557-5714

Notes

The authors declare no competing financial interest.

Biographies

Ye Hu studied clinical medicine at the Southeast University, China, from 1999 to 2004. In 2006, she joined the Life Science Informatics Master program at the University of Bonn and obtained her Master degree in 2008. In October 2008, she began her Ph.D. studies in the group of Prof. Jürgen Bajorath focusing on systematic computational analysis of molecular scaffolds of bioactive compounds and associated characteristics. Since July 2011, she is working as a postdoctoral fellow in the department. Her current research interests include large-scale mining of ligand−target interaction data and structure−activity relationship analysis.

Dagmar Stumpfe studied biology at the University of Bonn, Germany. In 2006, she joined the Department of Life Science Informatics at the University of Bonn headed by Prof. Jürgen Bajorath for her Ph.D. thesis, where she worked on methods for computer-aided chemical biology with a focus on the exploration of compound selectivity. Since 2009, Dagmar is working as a postdoctoral fellow in the department, and her current research interests include computational chemical biology and large-scale structure−activity relationship analysis.

Jürgen Bajorath is Professor and Chair of Life Science Informatics at the University of Bonn and also an Affiliate Professor in the Department of Biological Structure at the University of Washington, Seattle. His research interests include computational medicinal chemistry, chemoinformatics, chemical biology, and drug discovery. For further details, see: http://www.limes-institut-bonn.de/forschung/arbeitsgruppen/unit-4/abteilung-bajorath/abt-bajorath-startseite/.

ACKNOWLEDGMENTS

D.S. is supported by Sonderforschungsbereich 704 of the Deutsche Forschungsgemeinschaft.

ABBREVIATIONS USED

BM scaffold, Bemis and Murcko scaffold; 2-, 3-, 4D, two-, three, four-dimensional, respectively

REFERENCES

(1) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem., Int. Ed.* **1999**, *38* (19), 2894−2896.

(2) Hu, Y.; Stumpfe, D.; Bajorath, J. Computational Exploration of Molecular Scaffolds in Medicinal Chemistry. *J. Med. Chem.* **2016**, *59*, 4062−4076.

(3) Sun, H.; Tawa, G.; Wallqvist, A. Classification of Scaffold-Hopping Approaches. *Drug Discovery Today* **2012**, *17*, 310−324.

(4) Maggiora, G.; Vogt, M.; Stumpfe, D.; Bajorath, J. Molecular Similarity in Medicinal Chemistry. *J. Med. Chem.* **2014**, *57*, 3186−3204.

(5) Chen, Y. C.; Totrov, M.; Abagyan, R. Docking to Multiple Pockets or Ligand Fields for Screening, Activity Prediction and Scaffold Hopping. *Future Med. Chem.* **2014**, *6*, 1741−1755.

(6) Schuffenhauer, A. Computational Methods for Scaffold Hopping. *Wires Comput. Mol. Sci.* **2012**, *2*, 842−867.

(7) Hessler, G.; Baringhaus, K.-H. The Scaffold Hopping Potential of Pharmacophores. *Drug Discovery Today: Technol.* **2010**, *7*, e263−e269.

(8) Lauri, B.; Bartlett, P. A. CAVEAT: A Program to Facilitate the Design of Organic Molecules. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51−66.

(9) Rush, T. S., III; Grant, J. A.; Mosyak, L.; Nicholls, A. A Shape-Based 3-D Scaffold Hopping Methods and its Application to a Bacterial Protein-Protein Interaction. *J. Med. Chem.* **2005**, *48*, 1489−1495.

(10) 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.

(11) Schneider, P.; Tanrikulu, J.; Schneider, G. Self-Organizing Maps in Drug Discovery: Compound Library Design, Scaffold-Hopping, Repurposing. *Curr. Med. Chem.* **2009**, *16*, 258−266.

(12) Vogt, M.; Stumpfe, D.; Geppert, H.; Bajorath, J. Scaffold Hopping Using Two-Dimensional Fingerprints: True Potential, Black Magic, or a Hopeless Endeavor? Guidelines for Virtual Screening. *J. Med. Chem.* **2010**, *53*, 5707−5715.

(13) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887−2893.

(14) Wilkens, S. J.; Janes, J.; Su, A. I. HierS: Hierarchical Scaffold Clustering Using Topological Chemical Graphs. *J. Med. Chem.* **2005**, *48*, 3182−3193.

(15) Schuffenhauer, A.; Ertl, P.; Roggo, S.; Wetzel, S.; Koch, M. A.; Waldmann, H. The Scaffold Tree−Visualization of the Scaffold Universe by Hierarchical Scaffold Classification. *J. Chem. Inf. Model.* **2007**, *47*, 47−58.

(16) Wetzel, S.; Klein, K.; Renner, S.; Rauh, D.; Oprea, T. I.; Mutzel, P.; Waldmann, H. Interactive Exploration of Chemical Space with Scaffold Hunter. *Nat. Chem. Biol.* **2009**, *5*, 581−583.

(17) Li, R.; Stumpfe, D.; Vogt, M.; Geppert, H.; Bajorath, J. Development of a Method to Consistently Quantify the Structural Distance Between Scaffolds and to Assess Scaffold Hopping Potential. *J. Chem. Inf. Model.* **2011**, *51*, 2507−2514.

(18) Li, R.; Bajorath, J. Systematic Assessment of Scaffold Distances in ChEMBL: Prioritization of Compound Data Sets for Scaffold Hopping Analysis in Virtual Screening. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1101−1109.

(19) Hu, Y.; Bajorath, J. Global Assessment of Scaffold Hopping Potential for Current Pharmaceutical Targets. *MedChemComm* **2010**, *1*, 339−344.

(20) Hu, Y.; Bajorath, J. Quantifying the Tendency of Therapeutic Target Proteins to Bind Promiscuous or Selective Compounds. *PLoS One* **2015**, *10*, e0126838.

(21) Kawamura, S.; Unno, Y.; Hirokawa, T.; Asai, A.; Arisawa, M.; Shuto, S. Rational Hopping of a Peptidic Scaffold into Non-peptidic Scaffolds: Structurally Novel Potent Proteasome Inhibitors Derived from a Natural Product, Belactosin A. *Chem. Commun.* **2014**, *50*, 2445−2447.

(22) Vainio, M. J.; Kogej, T.; Raubacher, F.; Sadowski, J. Scaffold Hopping by Fragment Replacement. *J. Chem. Inf. Model.* **2013**, *53*, 1825−1835.

(23) Oyarzabal, J.; Howe, T.; Alcazar, J.; Andrés, J. I.; Alvarez, R. M.; Dautzenberg, F.; Iturrino, L.; Martínez, S.; Van der Linden, I. Novel Approach for Chemotype Hopping Based on Annotated Databases of Chemically Feasible Fragments and a Prospective Case Study: New Melanin Concentrating Hormone Antagonists. *J. Med. Chem.* **2009**, *52*, 2076−2089.

(24) Saluste, G.; Albarran, M. I.; Alvarez, R. M.; Rabal, O.; Ortega, M. A.; Blanco, C.; Kurz, G.; Salgado, A.; Pevarello, P.; Bischoff, J. R.; Pastor, J.; Oyarzabal, J. Fragment-hopping-based Discovery of a Novel Chemical Series of Proto-oncogene PIM-1 Kinase Inhibitors. *PLoS One* **2012**, *7*, e45964.

(25) Langdon, S. R.; Westwood, I. M.; van Montfort, R. L.; Brown, N.; Blagg, J. Scaffold-focused Virtual Screening: Prospective

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246

Application to the Discovery of TTK Inhibitors. *J. Chem. Inf. Model.* **2013**, *53*, 1100−1112.
(26) Rabal, O.; Amr, F. I.; Oyarzabal, J. Novel Scaffold FingerPrint (SFP): Applications in Scaffold Hopping and Scaffold-based Selection of Diverse Compounds. *J. Chem. Inf. Model.* **2015**, *55*, 1−18.
(27) Ertl, P. Database of Bioactive Ring Systems with Calculated Properties and its Use in Bioisosteric Design and Scaffold Hopping. *Bioorg. Med. Chem.* **2012**, *20*, 5436−5442.
(28) Ertl, P. Intuitive Ordering of Scaffolds and Scaffold Similarity Searching Using Scaffold Keys. *J. Chem. Inf. Model.* **2014**, *54*, 1617−1622.
(29) Ertl, P.; Lewis, R. IADE: A System for Intelligent Automatic Design of Bioisosteric Analogs. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1207−1215.
(30) Hamza, A.; Wagner, J. M.; Wei, N. N.; Kwiatkowski, S.; Zhan, C. G.; Watt, D. S.; Korotkov, K. V. Application of the 4D Fingerprint Method with a Robust Scoring Function for Scaffold-Hopping and Drug Repurposing Strategies. *J. Chem. Inf. Model.* **2014**, *54*, 2834−2845.
(31) Patel, S.; Harris, S. F.; Gibbons, P.; Deshmukh, G.; Gustafson, A.; Kellar, T.; Lin, H.; Liu, X.; Liu, Y.; Liu, Y.; Ma, C.; Scearce-Levie, K.; Ghosh, A. S.; Shin, Y. G.; Solanoy, H.; Wang, J.; Wang, B.; Yin, J.; Siu, M.; Lewcock, J. W. Scaffold-Hopping and Structure-Based Discovery of Potent, Selective, and Brain Penetrant N-(1H-Pyrazol-3-yl)pyridine-2-amine Inhibitors of Dual Leucine Zipper Kinase (DLK, MAP3K12). *J. Med. Chem.* **2015**, *58*, 8182−8199.
(32) Lo, Y. C.; Senese, S.; Damoiseaux, R.; Torres, J. Z. 3D Chemical Similarity Networks for Structure-based Target Prediction and Scaffold Hopping. *ACS Chem. Biol.* **2016**, *11*, 2244−2253.
(33) Petrone, P. M.; Simms, B.; Nigsch, F.; Lounkine, E.; Kutchukian, P.; Cornett, A.; Deng, Z.; Davies, J. W.; Jenkins, J. L.; Glick, M. Rethinking Molecular Similarity: Comparing Compounds on the Basis of Biological Activity. *ACS Chem. Biol.* **2012**, *7*, 1399−1409.
(34) Ratni, H.; Rogers-Evans, M.; Bissantz, C.; Grundschober, C.; Moreau, J. L.; Schuler, F.; Fischer, H.; Alvarez Sanchez, R.; Schnider, P. Discovery of Highly Selective Brain-Penetrant Vasopressin 1a Antagonists for the Potential Treatment of Autism via a Chemogenomic and Scaffold Hopping Approach. *J. Med. Chem.* **2015**, *58*, 2275−2289.
(35) van Westen, G. J. P.; van den Hoven, O. O.; van der Pijl, R.; Mulder-Krieger, T.; de Vries, H.; Wegner, J. K.; IJzerman, A. P.; van Vlijmen, H. W. T.; Bender, A. Identifying Novel Adenosine Receptor Ligands by Simultaneous Protochemometric Modeling of Rat and Human Bioactivity Data. *J. Med. Chem.* **2012**, *55*, 7010−7020.
(36) Johnson, M., Maggiora, G. M., Eds. *Concepts and Applications of Molecular Similarity*; John Wiley & Sons: New York, 1990.
(37) Stahl, M.; Bajorath, J. Computational Medicinal Chemistry. *J. Med. Chem.* **2011**, *54*, 1−2.
(38) Klenner, A.; Hartenfeller, M.; Schneider, P.; Schneider, G. "Fuzziness" in Pharmacophore-based Virtual Screening and De Novo Design. *Drug Discovery Today: Technol.* **2010**, *7*, e237−e244.
(39) Bajorath, J. Computational Chemistry in Pharmaceutical Research: At the Crossroads. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 11−12.
(40) Dimova, D.; Stumpfe, D.; Hu, Y.; Bajorath, J. ASB Scaffolds: Computational Design and Exploration of a New Type of Molecular Scaffolds for Medicinal Chemistry. *Future Sci. OA* **2016**, *2*, FSO149.

DOI: 10.1021/acs.jmedchem.6b01437
J. Med. Chem. 2017, 60, 1238−1246