Scaffold hopping in drug discovery

Bhupinder Singh Sekhon and Naveen Bimal
PCTE Institute of Pharmacy, Jhanday, near Baddowal Cantt, Ludhiana-142 021, India

ABSTRACT
Molecular scaffolds extracted from bioactive compounds continue to be of high interest in medicinal chemistry. Scaffold hopping is used to identify compounds containing a topologically different scaffold from the parent compound, but with similar or improved activity and other properties from a given database. Different core structures and similar biological activities of the new compounds relative to the parent compounds comprise two key components of scaffold hopping which expands the range of core molecular shapes for lead generation. Various software programs useful to guide scaffold hopping are included.

Keywords: Scaffold hopping, medicinal chemistry, molecular scaffolds, bioactive compounds, molecular hopping.

Due to issues mainly related to intellectual property, physicochemical properties, metabolic stability, toxicity etc., the scaffold of a molecule with an appealing biological activity cannot be developed further in drug discovery. In this context, scaffold hopping offers a good chance to apply computational chemistry methods for solving existing practical problems in drug discovery. The term “scaffold hopping” was coined by former Hoffmann-La Roche researcher Gisbert Schneider for drug discovery to identify isofunctional molecular structure with different molecular backbones having similar or improved properties. A lead compound defined as a compound from a series of related compounds having some of the desired biological activity can be characterized, and modified to produce another molecule with a better profile of required properties to unwanted side effects. Previous studies suggested that leads based on different scaffolds are possible by scaffold hopping technique.1–7

In scaffold hopping approach, one part of a molecule is modified with another one (the chemical scaffold) of a bioactive compound yet binding to the same molecular target. In this process, the bioactivity is retained or improved. Scaffold folding approach has been well illustrated by molecules like diazepam, zolpidem, zaleplon, and zopiclone which exert the same biological response acting as full agonists of GABA-A (γ-aminobutyric acid) receptor at the benzodiazepine site though a structural analogy is barely found8 (Figure 1).

Many scaffolds were obtained in the anti-inflammatory field of the cyclooxygenase ligands.9,10 A common use of pharmacophores (the group of atoms in the molecule of a drug responsible for the drug’s action) is to search 3D databases for molecules containing it. Scaffold hopping using 3D pharmacophores has been successfully carried out.11 The concept of scaffolds and scaffold hopping in the context of molecular topologies has been reported.12 An alteration of the central chemical template of a compound is often desirable for several reasons: i) a replacement of a lipophilic scaffold by a more polar one for increased solubility, ii) a substitution of a metabolically labile scaffold with a more stable or less toxic one for improving the pharmacokinetic properties, iii) a replacement of a very flexible scaffold (such as a peptide backbone) by a rigid central scaffold for
significantly improving the binding affinity and a change in the central scaffold for generating a novel structure that is patentable.\textsuperscript{6,13}

Drug like molecules are built from scaffold (framework) and side chain.\textsuperscript{14} A hierarchical classification of chemical scaffolds (molecular framework obtained by pruning all terminal side chains) has been reported.\textsuperscript{15} The molecular scaffold representations are very active and constitute an important area of current research in chemoinformatics. Representation of full molecular structure (A), scaffold representation (B) and reduced scaffold (C) are shown in Figure 2.

Another example regarding representation of complete structure (D), molecular framework (E), molecular framework (ignoring atom type) (F), molecular framework (ignoring bond type) (G), molecular framework (ignoring atom and bond type) (H), and reduced graph (I) is shown in Figure 3. In context with abstractions of a structure out of the National Cancer Institute comprehensive cancer database (Figure 3) both abstractions rely on the molecular framework obtained by pruning all terminal side chains. The scaffold is further generalized by removing atom- and/or bond-type information and finally by describing it through the SCINS code “framework hierarchy”. The Scaffold tree algorithm allows organizing large molecular data sets by arranging sets of molecules into a unique tree hierarchy based on their scaffolds, with scaffolds forming leaf nodes of such tree and generalizes the scaffold by the iterative removal of rings according to prioritization rules.\textsuperscript{15} The basic principles of the Scaffold tree methodology, its applications including the use of Scaffold trees for visualization of
Figure 2: A: Full molecular representation, B: Scaffold representation, C: Reduced scaffold representation.

Figure 3: Molecular scaffold representations – D: Complete structure; E: Molecular framework; F: Molecular framework (ignoring atom type); G: Molecular framework (ignoring bond type); H: Molecular framework (ignoring atom and bond type); I: Reduced graph.
large chemical data sets, compound clustering, and the identification of novel bioactive molecules have been reported.\textsuperscript{16}

The selection of a template structure is the starting point for scaffold hopping. This is followed by hopping iso-functional, but structurally dissimilar, scaffolds into different parts of the template structure. Various novel approaches have recently been developed and applied in this direction.\textsuperscript{17} In drug discovery, it has a number of potential applications such as replacement of a patented core scaffold with a different scaffold free from IP restrictions, development of a back-up series of compounds and early property exploration of different chemical series for drug development. Scaffold hopping is one of the strengths of the feature tree software which can be realized to full effect as the vastness of the search space covers an innumerably large number of products based on quite sizeable numbers of scaffolds. The searches can be performed quickly and have been shown to produce active hits from novel chemical series in Big Pharma.\textsuperscript{18}

The identification of novel active compounds against a preselected biological target with acceptable pharmacological properties according to marketed drugs specification is a general goal of drug discovery. Researchers classified scaffold hopping into four major categories, namely heterocyclic replacements, ring opening or closure, peptidomimetics and topology-based hopping incorporating structural diversity of original and final scaffolds with respect to each category. The advantages and limitations of small, medium and large-step scaffold hopping modifications were also reported.\textsuperscript{2} Swapping carbon and nitrogen atoms in an aromatic ring or replacing carbon with other hetero-atoms in a ring representing small step hops results in a low degree of structural novelty. The transformation from morphin to tramadol obtained by cleaving six ring bonds and opening up three fused rings (ring opening) was found more flexible, resulted in reduced potency and side effects. Ring closure design was obtained by converting – an alkyl chain to cyclohexane, piperizine or piperidine, -o-hydroxylbenzoyl group to quinazoline and an arylamine or arylamide to a fused ring system. Peptidomimetics compounds mimics a natural peptide or protein in 3D space while retaining the ability to interact with the biological target and produced the same biological effect. Further, topology-based hops always lead to a high degree of novelty.\textsuperscript{3} Computational scaffold extraction methods have been frequently used in approximate benchmarks for scaffold hopping.\textsuperscript{19} Researchers summarized software that found frequent use to facilitate different kinds of scaffold-hopping methods.\textsuperscript{2} A series of derivatives with α-L-amino acid moieties in place of ureido group in (I) resulted in (II) through scaffold hopping method (Figure 4).\textsuperscript{20} Other examples are available in literature.\textsuperscript{21, 22} Key methods in scaffold hopping to replace only a core motif of a known ligand while keeping the key substituents conserved include shape matching,\textsuperscript{23} pharmacophore searching,\textsuperscript{24} fragment replacement,\textsuperscript{25} and similarity searching.\textsuperscript{26} The programs for the different categories of scaffold hopping are shown in Table 1.\textsuperscript{19, 27} In pharmacophore-driven scaffold hopping, various pharmacophore descriptors determine the correlation between newly generated scaffold and the original scaffold. In addition, the 2D and 3D pharmacophore fingerprints, designed specific pharmacophore descriptors for scaffold hopping are shown in Table 1. Reduced graphs offer the potential to retrieve compounds that were similar in terms of their gross features rather than at the atom-bond level. In this context, researchers exploit the low cardinality of the reduced graph in graph-based similarity searching. They showed graph matching as an effective retrieval method using fully connected reduced graphs. Moreover, the actives retrieved were topologically different from those retrieved using conventional 2D methods.\textsuperscript{28}

![Figure 4: Derivatives with α-L-amino acid moieties in place of ureido group.](image-url)
The application of the shape-comparison program ROCS (Rapid Overlay of Chemical Structures) has provided new scaffolds for small molecule inhibitors of the ZipA–FtsZ protein–protein interaction. In this direction, a set of novel, weakly binding inhibitors with scaffolds presenting synthetic opportunities were identified. Moreover, such ROCS-identified scaffolds could not be achieved using other structural similarity approaches such as ISIS 2D fingerprints. Based on the X-ray crystallographic analysis of one of the inhibitors bound to ZipA reveals that the shape comparison approach very accurately predicts the binding mode, thereby, validating the use of ROCS for scaffold hopping.

By means of shape- and pharmacoaphore-based virtual screening, the identification of a potent PPAR-selective activator from a large compound collection represented a scaffold-hop from known PPAR agonists and provided proof-of-concept for a novel ligand-based virtual screening approach.

Inductive logic programming (ILP) uses the observed spatial relationships between pharmacoaphore types in pretested active and inactive compounds and learns human-readable rules describing the diverse structures of active compounds. The comparison of ILP-based scaffold hopping method to two previous algorithms (chemically advanced template search, CATS, and CATS3D) on 10 data sets with diverse scaffolds showed the ILP-based method significantly better than random selection while the other two algorithms were not. In addition, the ILP-based method retrieved new active scaffolds which were not found by CATS and CATS3D. Normally, one looks for identification of molecules that are structurally dissimilar to a query molecule but has similar biological properties in scaffold hopping. Researchers showed that empirical evaluation highlighted interesting patterns that deviate from the well-established Lipinski’s rule in the domain of drug-likeness. ReCore replaces a given core: Given a pre-defined central unit of a molecule (the core), fragments are searched in a 3D database for the best possible replacement – whilst keeping all connected residues, i.e., the rest of the query compound in place. Additionally, user-defined “pharmacophore” constraints can be employed to restrict solutions.

Shop is useful to guide the scaffold hopping procedure during the drug discovery process. L.Ox 2.0 identify and optimize drug lead candidates through preserving and linking critical features with new scaffolds, altering the core scaffold of a lead series to maximize activity, optimize physical properties and to create novel intellectual property. Chemically intuitive scaffold distances have been obtained for pairs of scaffolds with varying composition and topology. Further, distance threshold values for close and remote structural relationships between scaffolds were also determined. An approach (LigCSRe) to the 3D ligand similarity search of drug candidates combines a 3D maximum common substructure search algorithm independent on atom order with a tunable description of atomic compatibilities to prune the search and increase its physico-chemical relevance. Researchers demonstrated that for one or multiple templates of a given chemotype, other chemotypes were reached during de novo compound generation, thus, indicating successful scaffold-hops. Researchers have made an assessment of global scaffold hopping potential by systematically analyzing topologically distinct scaffolds in currently available bioactive compounds with defined targets and activity annotations. The analysis revealed that scaffold hops occurred with rather high frequency among active compounds. Self-organizing map (SOM, Kohonen network) and variations thereof have found widespread application. SOMs cover such diverse fields of drug discovery as screening library design, scaffold-hopping, and repurposing.

The application of a GQSAR method to assist in lead optimization of multikinase (PDGFR-beta, FGFR-1 and SRC) and scaffold hopping of multiserootonin target (serotonin receptor 1A and serotonin transporter) inhibitors has been reported. The developed GQSAR models were found to be useful for scaffold hopping and lead optimization of multitarget inhibitors. Selected literature data on biological activity of the known conformationally restricted bicyclic secondary diamines derivatives demonstrated utility of the conformationally restricted bicyclic secondary diamines scaffold hopping in drug design.

**Table 1. Scaffold hopping programs**

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacophore-driven</td>
<td>Inductive logic programming (ILP); CATS(CATS, CATS3D, SURFCATS); Charge 3D, Triple Charge 3D; FEPOPS; the Similog pharmacoaphore keys; SQUID; LIQUID</td>
</tr>
<tr>
<td>Reduced graph-based</td>
<td>Clique detection, ErG, centroid connecting path; structural unit analysis, Recore</td>
</tr>
<tr>
<td>Shape-based</td>
<td>Grid molecular interaction field based approaches (such as MOLPRINT3D, FLAP, and SHOP, ROCS, FieldScreen, extended electrical distribution, field-based similarity search, Surflex-Sim, Topomer, KIN, and ParaFrag).</td>
</tr>
</tbody>
</table>

Bhupinder Singh Sekhon and Naveen Bimal: Scaffold hopping in drug discovery
identified a backup drug candidate for BPR0L075 ([6-methoxy-3-(39,49,59-trimethoxy-benzoyl)-1H-indole]), an indole-based anticancer agent (Figure 5). In this context, 5,6-fused bicyclic heteroaromatic scaffolds were designed and synthesized through shuffling of the nitrogen from the N-1 position or by insertion of one or two nitrogen atoms into the indole core of BPR0L075. Among these, 7-azaindole core showed potent in vitro anticancer activity and improved oral bioavailability (F = 35%) compared with BPR0L075 (F < 10%).

Monoamine oxidase B (MAO-B) inhibitors with potencies in the low nanomolar and low micromolar range have been identified using a scaffold hopping approach with the thiazolinedione class of compounds. Derivatives of the natural product sulfuretin were found potent MAO-A and MAO-B inhibitors. 2D fingerprints provided a simple and computationally efficient way of identifying novel chemotypes in lead-discovery programs.

Compounds with novel scaffolds represent promising starting points of an optimization program against E. histolytica. Compounds with the benzotriazole and indazole scaffolds showed low micromolar activity (IC₅₀ = 0.304 and 0.339 μM) and were found more active than metronidazole, which is used for the treatment of amebiasis. Moreover, the novel compounds have similar properties to approved drugs. Molecule Cloud - a method for compact visualization of the typical substructures present in large collections of molecules was reported. The Molecule Cloud graphs allowed recognition of scaffolds and other substructure features that are typical for particular data set by a single look. Evolvus offers advanced computational approaches for hopping from one scaffold to another for modifying drug selectivity and affinity profile, as well as for optimising physicochemical and ADMET properties. By its virtue to jump into a new chemical space, scaffold-hopping inherits a strong potential for generation of intellectual property.

In view of an urgent need for a database system to provide valuable data in the drug design field, ScafBank (Web-based database system) incorporating scaffolds derived from the bioactive compounds could help researchers in the fields of medicinal and organic chemistry to design novel molecules with properties similar to the original compounds, but built on novel scaffolds. Moreover, each entry in the database was associated with a molecular occurrence and includes its distribution of molecular properties, such as molecular weight, logP, hydrogen bond acceptor number, hydrogen bond donor number, rotatable bond number and ring number.

**CONCLUSION AND PERSPECTIVES**

Scaffold hopping two key components (different core structures and similar biological activities of the new compounds relative to the parent compounds) is a strategy for discovering structurally novel compounds. Basically it refers to the identification of different compound classes having similar biological activity and particularly useful to obtain alternative compound structures when the initial compound under development has unexpected side-effects or when the initial compound is patented by competitors. In fact, scaffold hopping, a central task of modern medicinal chemistry, has demonstrated the capability to discover equipotent compounds with novel backbone having improved properties. In order to avoid prior patents for targets of interest, pharmaceutical companies use scaffold hopping.

To facilitate the chemical syntheses and testing of drug development, the identification of few peptide scaffolds having diverse potential with chemical modifications has been reported. Synthetic peptidomimetics like aptamers, dendrimers, arylamide foldamers, β-peptides, may be exploited for the therapeutic use of scaffold structures.

**REFERENCES**


**Figure 5:**