

Scaffold Hunter – Interactive Exploration of Chemical Space

Karsten Klein¹, Nils Kriege¹, Petra Mutzel¹,
Herbert Waldmann², and Stefan Wetzel²

¹ Technische Universität Dortmund, Germany

{karsten.klein,nils.kriege,petra.mutzel}@cs.tu-dortmund.de

² Max-Planck-Institute for Molecular Physiology, Dortmund, Germany

{herbert.waldmann,stefan.wetzel}@mpi-dortmund.mpg.de

Abstract. Scaffold Hunter is a Java-based software tool for the analysis of structure-related biochemical data. It facilitates the interactive exploration of chemical space by enabling generation of and navigation in a scaffold tree hierarchy annotated with various data. The graphical visualization of structural relationships allows to analyze large data sets, e.g., to correlate chemical structure and biochemical activity.

The search for small molecules that are biologically relevant, e.g., to design new drugs and diagnostics, is an important task in chemical biology. Even though high throughput methods are available to test large numbers of chemical compounds, the overwhelming size of chemical space—containing up to 10^{160} different molecules—renders any exhaustive search infeasible. Methods are needed that allow to focus the search on the tiny fraction of chemical space that contains the most promising candidates for biological activity.

Schuffenhauer et al. [1] developed a hierarchical scaffold classification strategy to chart chemical spaces. A rule set is defined to reduce molecules to *scaffolds*, on which a hierarchy is defined based on a substructure relation, cf. Fig. 1. This way, the compounds are classified by structural similarity, reflecting the fact that the biological relevance is closely coupled to the molecular structure.

We developed Scaffold Hunter, a software tool that admits the graphical representation of chemical compound databases based on this classification strategy. The software visualizes the set of scaffold trees resulting from the hierarchy generation step using graph layouts, and allows interactive exploration of the data

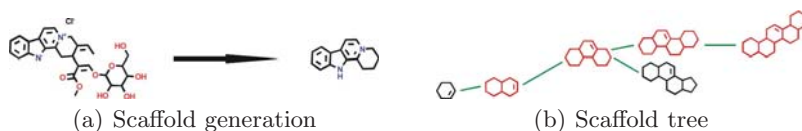


Fig. 1. Example for a reduction step that derives a scaffold structure from a given molecule structure (a) and a partial scaffold tree (b)

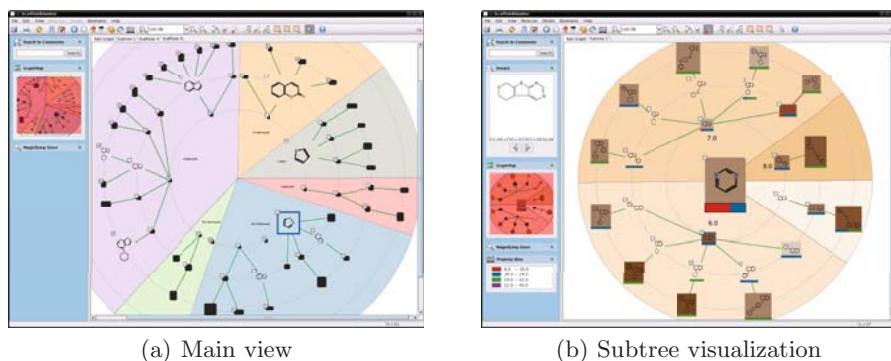


Fig. 2. Screenshots of the main view and the visualization of a subtree generated by a filtering step, both with property-based color shading

based on this representation, cf. Fig. 2(a). Scaffold Hunter provides several layout styles—adaptions of radial, tree, and balloon layout—and iterative scaffold property filtering can be used to reduce the number of scaffolds displayed. Color shading based on selected properties, both of scaffolds and specific segments of the dataset, facilitates orientation and navigation within the dataset. The set of molecules corresponding to scaffolds of interest and also subtrees derived by filtering and selection steps can be viewed in a separate window, cf. Fig. 2(b).

The use of Scaffold Hunter fits into the chemist's workflow for the analysis of large compound libraries and allows to intuitively explore the underlying chemical space. It helps to identify 'holes' in the structure space analyzed that may serve as promising starting points for compound library design. Scaffold Hunter identifies virtual scaffolds that do not represent molecules in the dataset and that should share bioactivity properties with their parents or child scaffolds. These virtual scaffolds may provide new opportunities for the identification of new biologically relevant scaffold classes. An initial proof-of-concept analysis demonstrated the usefulness of our approach [3].

Scaffold Hunter is implemented in Java and freely available under the GPL [2].

References

1. 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.* 47(1), 47–58 (2007)
2. Scaffold Hunter Project Webpage, <http://sourceforge.net/projects/scaffoldhunter/>
3. 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.* 5(8), 581–583 (2009)