A Novel 2D Depiction Method Using Breadth-First Ordering and an Adapted 2D Force Field

A. Dornhofer[†], M. Biely[†], G. Wolber[†] and T. Langer[‡]

[†] Inte:Ligand GmbH, Mariahilferstraße 74B/11, 1070 Vienna, Austria and
[‡] Computer Aided Molecular Design Group, Institute of Pharmacy,
Dpt. Pharmaceutical Chemistry, University of Innsbruck, Austria

inte:ligand
Your partner for in-silico drug discovery.

INTRODUCTION

The continuously growing number of chemical and pharmaceutical software packages reaching from lead optimization via library design and virtual screening to virtual high-throughput screening deal with huge numbers of molecules. The method of choice for presenting molecules in a fast and convenient way is a two-dimensional representation. Unfortunately, freely available depiction applications do not meet standards for high quality whereas programs that generate high-quality representations are only available commercially which motivates our efforts.
Conceptually three problems arise when implementing a two-dimensional depiction algorithm. The first is to perceive rings from a given molecule, which has been solved by implementing Figueras efficient breadth-first search approach to identify the smallest set of smallest rings (SSSR) [1]. The second problem is the suitable depiction of ring systems which may contain highly connected sub-ring systems and the third problem are dense molecules, for which in many cases the best solution is difficult to find using standard bond angles.

METHODS

The presented approach is based on a graph-theoretical abstraction of non-ring atoms and complete ring systems as nodes and resembles previously presented ideas using reduced graphs [2, 3]. In fact, this step transforms an arbitrary molecular graph into an acyclic graph which is subsequently easier to handle. Since rings and ring systems can be regarded as single graph nodes, the calculation of their representation becomes isolated from the regular coordinate assignment algorithm. In the next step templates are assigned to all inner nodes, the outer nodes stay as leaf nodes (L). This is illustrated in figure 1: (a) contains the acyclic graph where the ring system (RS) as well as the bridged ring (BR) is each contracted to a single template, while T_0, T_1, and T_2 are regular templates. An initial template (in this example RS) has to be chosen and the others are obtained by breadth-first searching. Hence, the final ordering is {RS, T_0, T_1, BR, T_2}. As shown in (b), the initial template has assigned positions to itself and its neighbor templates T_0 and T_1. In (c) T_0 assigns positions to its neighbors L and BR, then T_1 assigns positions to its neighbor T_2 (d), and so on until all templates have assigned positions to their neighbors. (f) shows the final structure.

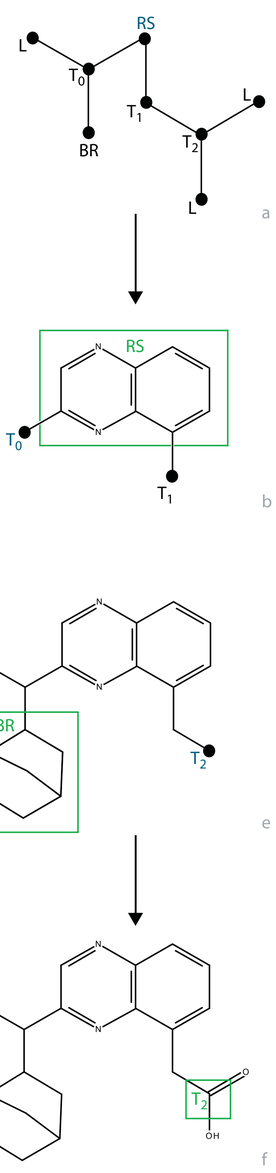

Fig. 1. Basic algorithm for assigning positions to structures using templates in breadth-first ordering

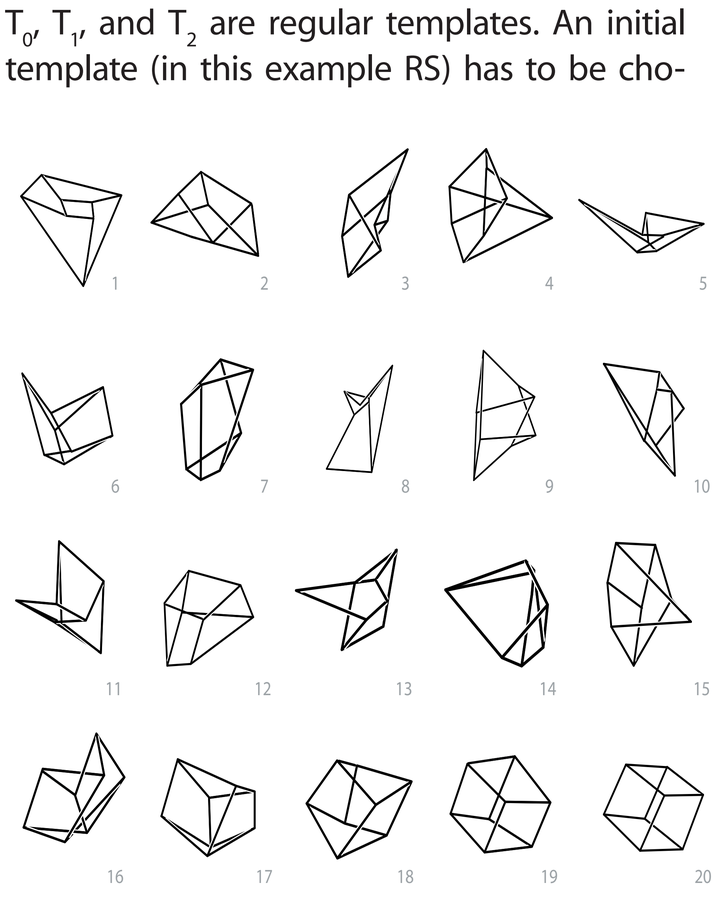

Fig. 2. 2D force field applied to Cubane yields good results after just a few iterations

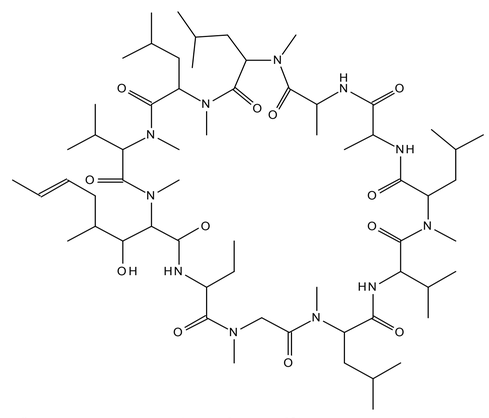

Fig. 3. Large ring handling (example Cyclosporine)

Single rings are assigned symmetrically distributed positions on a circle, whereas all simple annelated and spiro-ring systems are composed of single rings. However, standard geometry is inappropriate for dense ring systems and bridged rings. This problem has been solved by applying an adapted two-dimensional force field [4] to these systems which tends to generate symmetric layouts as illustrated in figure 2.
However, there is still an aesthetical issue to be solved for rings comprised of 10 or more atoms. Symmetric distribution of positions on a circle makes it difficult to perceive the correct number of ring atoms. Therefore large rings are assigned positions on an automatically generated structure for the specified number of atoms with angles of 120 degrees. Figure 3 shows the result for Cyclosporine.
The last problem to be addressed is the problem of dense molecules. Here the power of breadth-first ordering of templates enfolds completely: In conjunction with a backtracking algorithm it ensures that all possible transformations on templates have been performed to find the best solution. If this straight-forward approach does not yield a suitable solution due to crossed edges or nodes being too close, an additional mechanism takes the best solution at this point and bends sub-graph parts to finally improve the existing layout as shown in figure 4.

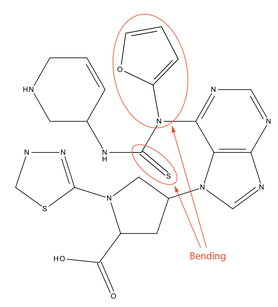

Fig. 4. Conflict resolution through bending

RESULTS

The presented method was manually evaluated for speed and accuracy with virtual molecular libraries generated by the software application CombiGen [5] as well as with ligands from the Protein Data Bank where the software application LigandScout [6] has been used to extract the ligands. The shown algorithm proved to be a robust, usable and fast tool for the fully automated two-dimensional depiction of molecules.

References

[1] Figueras J. Ring Perception Using Breadth-First Search. J Chem Inf Comp Sci. 29(1):43-48; 1979.

[2] Gillet V, Willett P, Bradshow J. Similarity Searching Using Reduced Graphs. J Chem Inf Comp Sci. 2003; 43(2):338-345.

[3] Rarey M, Dixon J. Feature Trees: A New Molecular Similarity Measure Based on Tree Matching. J Comput-Aided Mol Design. 1998; 12:471-49

[4] Frick A, Ludwig A, Mehldau H. A Fast Adaptive Layout Algorithm For Undirected Graphs. In: Tamassia R, Tollis IG, editors. Lecture Notes in Computer Science. Proc. DIMACS Int. Work. Graph Drawing, GD: Berlin: Springer; 1994. p. 338-403.

[5] Wolber G, Langer T. CombiGen: A novel software package for the rapid generation of virtual combinatorial libraries. In: Höltje HD, Sippl W. editors. Rational Approaches to drug design. Prous Science; 2000. p. 390-399; available as “ilib diverse” from Inte:Ligand GmbH; www.inteligand.com

[6] Wolber G, Langer T. LigandScout: 3D Pharmacophores derived from protein-bound ligands and their use as virtual screening filters, J Chem Inf Comp Sci, accepted for publication