Drawing the PDB

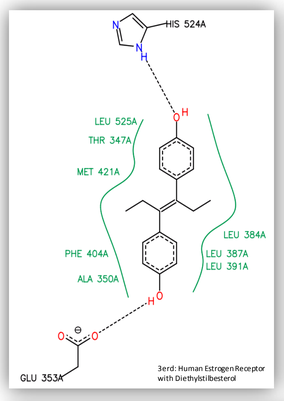
A large-scale application study of the 2D drawing tool PoseView

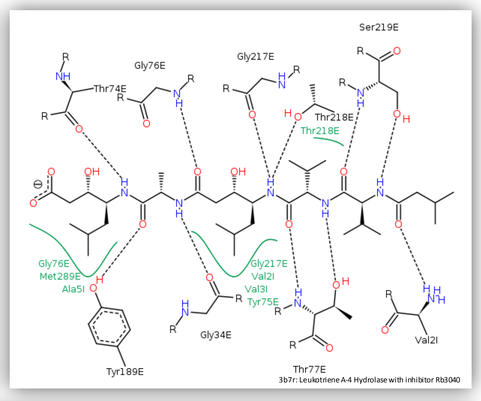
Katrin Stierand | Matthias Rarey

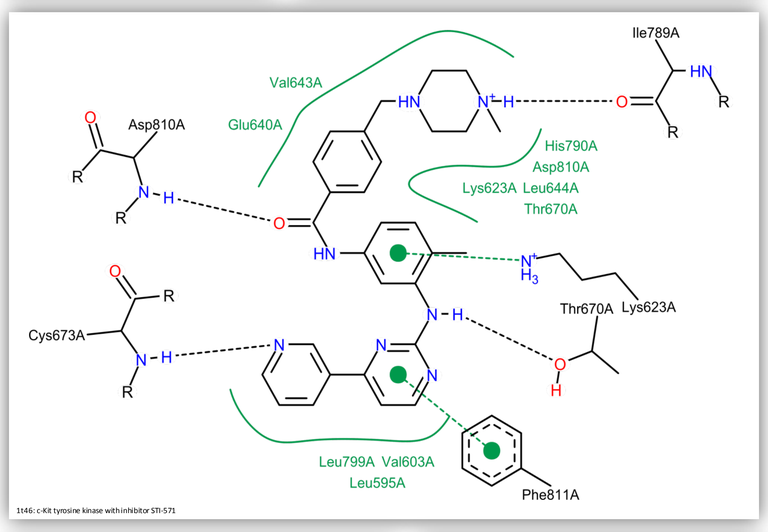
http://poseview.zbh.uni-hamburg.de

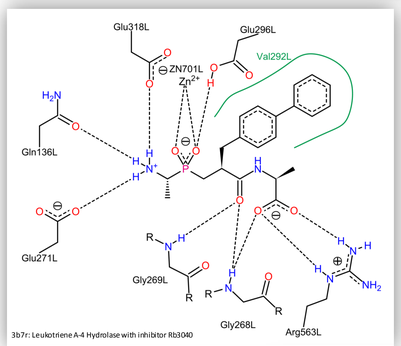
Two-dimensional visualization of molecular structures is a well established procedure to communicate results in drug design processes. We developed PoseView[1,2], a tool that automatically generates diagrams of macromolecular complexes showing the ligand and the interacting receptor compounds as structure diagrams and the interaction pattern in between. An application study on complexes taken from the PDB[3] showed high success rates; 92% of 189,000 input complexes could be drawn. In the following, we will present the interaction model, a method overview and the results of the application study.

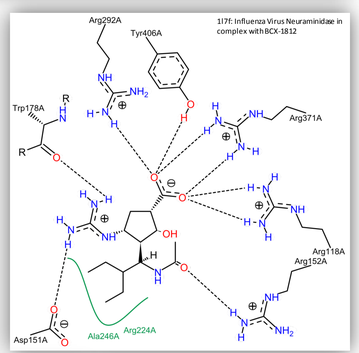
GALLERY











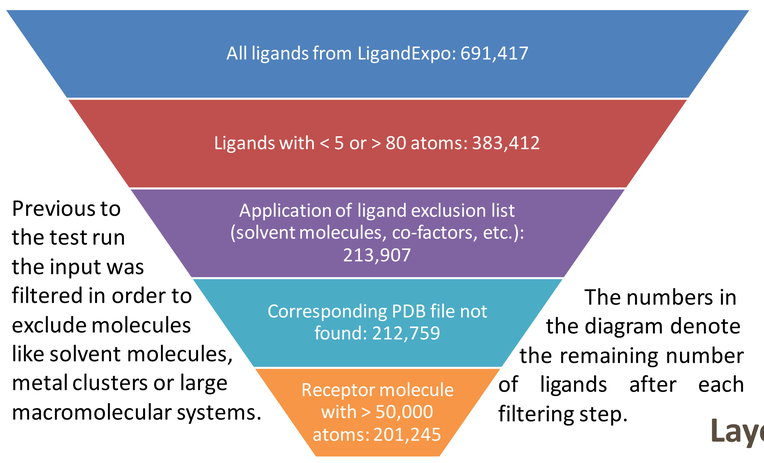
CONCLUSION

PoseView offers the opportunity to facilitate the evaluation of different protein-ligand complexes. It is available as web-service and can help scientists to quickly compare research results with related complexes. The application on PDB data showed that the tool is able to draw large parts of its input complexes in good quality. Due to short computing times it is also possible to generate diagrams online, see *http://poseview.zbh.uni-hamburg.de*.

PDB APPLICATION STUDY

We tested the performance of PoseView on complexes provided by the PDB[3] . While the protein structures originate from the PDB, the ligands are taken from the Ligand Expo database [4], which provides for many of the PDB structures one or more co-crystallized small molecules. The presented diagrams below summarize statistics for input filtering, success rates and computing times.

Input filtering

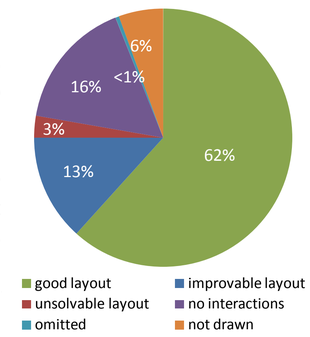


Previous to the test run the input was filtered in order to exclude molecules like solvent molecules, metal clusters or large macromolecular systems.

The numbers in the diagram denote the remaining number of ligands after each filtering step.

Result statistic

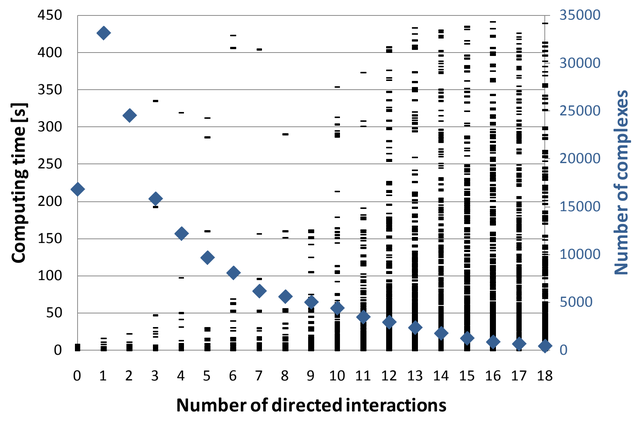
After the filtering step 201,245 complexes remained. For one sixth of them no interactions were found by the interaction model. For ~155,000 complexes a layout could be computed. They are subdivided in three quality categories: Good layouts without collisions , improvable layouts whose collisions could be solved by additional strategies and layouts with unsolvable collisions. Examples for each quality category are given below.



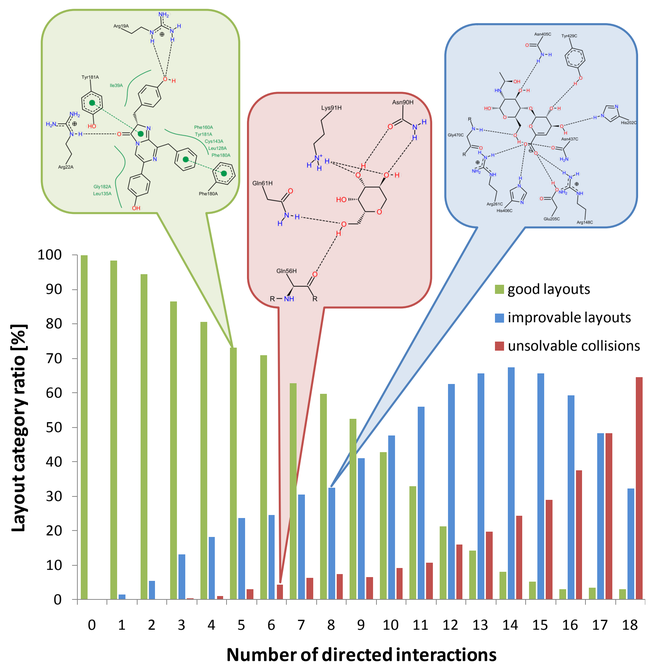
A small part of all complexes could not be drawn due to technical reasons and less than 1000 complexes were omitted because they have more than 18 interactions or more than 14 residues participating in the interactions.

Computing times

This diagram shows the computing time depends on the number of interactions. The blue diamonds denote the number of complexes for each bin. Over 90% of all complexes have less than 11 interactions and their 2D layout can be computed in the range of milliseconds ton seconds.



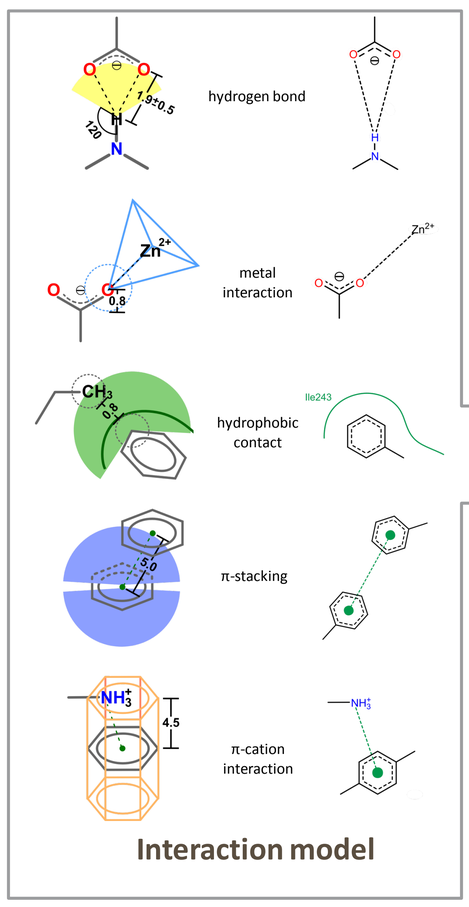
Layout quality ratio

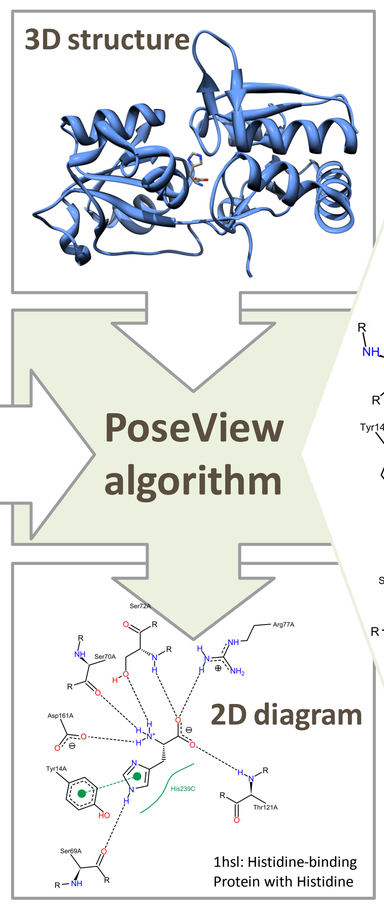


The ratio of diagram layout quality changes with the number of interactions. While the number of good layouts decreases with the growing number of interactions, the layouts containing unsolvable collisions show the opposite behavior. For each of the layout types an example diagram is shown in the corresponding box.

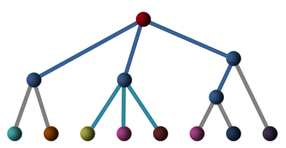
COMPLEX DIAGRAM GENERATION

If interactions between ligand and receptor are not part of the input, they are estimated by a built-in interaction model. It is based on simple distance and angle criteria and considers five different interaction types. In combination with the complex 3D structure it is the input of the layout algorithm.

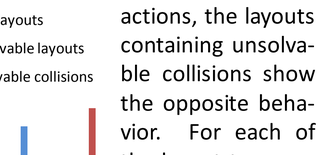




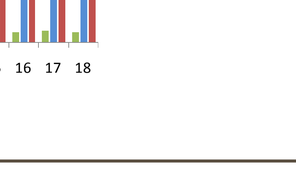
Layout algorithm



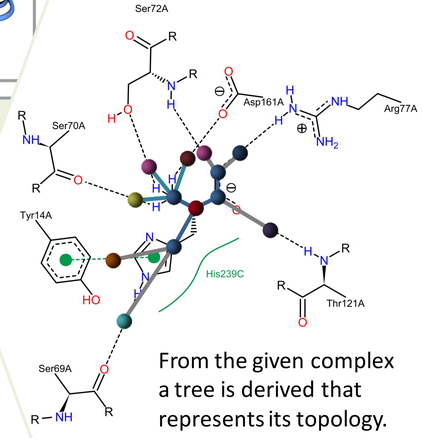
The layout can be modified by exchanging or rotating edges. The edge types are assigned according to the corresponding bond types.



To find a good layout all possible combinations of modifications are enumerated. In this example, the first applied operation is a rotation.



The next step is an exchange of the two red bonds. These operations are applied to the ligand structure diagram.



From the given complex a tree is derived that represents its topology.

[1] K. Stierand, M. Rarey, ChemMedChem, **2007**, 2: 853.
[2] K. Stierand, P. Maaß, M. Rarey, Bioinformatics **2006**, 22, 1710.
[3]H.Berman et al., Acta Cristallographica Section D, **2002**, 58: 899.
[4] Z.Feng et al., Bioinformatics **2004**, 20: 2153.

Katrin Stierand – stierand@zbh.uni-hamburg.de
Matthias Rarey – rarey@zbh.uni-hamburg.de
ZBH Center for Bioinformatics, Bundesstr. 43, 20146 Hamburg
This project is funded by Klaus Tschira Stiftung Gemeinnützige GmbH