Conformer generation is the computational prediction of three-dimensional molecular conformations. Conformer generation is a key component of virtual screening methods since these methods of computationally identifying potential drug-protein interactions require realistic three-dimensional structures. Current methods of conformer generation are energy-based, where conformers are generated to minimize an empirically determined energy function. An alternative approach is enabled by the recent explosion of structural data, as exemplified by the Protein Data Bank (102,158 structures) and the Crystallography Open Database (293,056 structures). In this knowledge-based approach, conformers are generated by stitching together pieces of experimentally determined structures.

In this project we analyze the PDB and COD and show that the majority of bioactive conformers in the PDB can be reconstituted using experimental data from the COD. Conformers now need to be generated and tested to measure the efficacy of these results.

References


Acknowledgements

- David Koes
- Joseph Ayoob, Tim Lezon, Leah Russell, Mark Schurdak, and the entire DISCoBio Program
- Megan Houlihan, Michael Lotze, and the entire UPCI Academy staff
- Sean Geiger and my family

Conclusions

These results suggest that knowledge-based conformer generation methods can be successful in reconstituting PDB ligands from the experimental data in the COD. Conformers now need to be generated and tested to measure the efficacy of these results.

Future Directions

This figure shows the percentage match of the average molecule in the PDB Database based on the most popular clusters found from the COD. The X-Axis follows the increase in K, the highest cluster analyzed, while the Y-Axis follows the percentage match of a molecule to the K highest clusters.