Using small molecule crystal structure data to improve drug discovery

Paul C. Sanschagrin
The Cambridge Crystallographic Data Centre, Piscataway, NJ 08854, USA

Successful modern drug discovery research makes extensive use of structural data – from target proteins, candidate drug molecules, and complexes of the two. The value of protein-ligand structural information is well accepted. However, knowledge of molecular conformations and interactions derived from small molecule structures alone can have a significant impact in the drug discovery endeavor. The Cambridge Structural Database (CSD) contains structures for over 915,000 organic and metal—organic molecules. Over 80% of the organic compounds with fully resolved atomic structures in the CSD meet the Lipinski rule of 5 for drug-likeness, indicating that the CSD is a vast repository of structural data relevant to drug discovery. Knowledge-based methods, such as Mogul and IsoStar, apply what we already know about small molecules to design even better ones. This presentation will include several examples where such small molecule data provided key insights and direction to improve drug candidates, including the use of Mogul geometric analysis to optimize the binding mycophenolic acid analogs to inosine monophosphate dehydrogenase (IMPDH) and to identify aryl sulfones with improved binding to the glucokinase–glucokinase regulatory protein (GK-GKRP) complex.