Analysis of the interactions in protein binding sites as a tool aiding inhibitory activity prediction

Wiktoria Jedwabny

The main goal of the dissertation was to investigate the physical nature of the protein–inhibitor binding with the first principle-based methods, in order to develop possibly simple and universal nonempirical scoring model of inhibitory activity. For that purpose, reference MP2 interaction energies were partitioned into electrostatic multipole and penetration, exchange, delocalization and correlation components for each studied complex. Performed analyses of five different groups of inhibitors indicate that the optimal model should account for two long-range components of the interaction energy, i.e., electrostatic multipole and dispersion terms, constituting MED (*Electrostatic Multipole and Dispersion*) model. In MED, the former term is represented by Cumulative Atomic Multipole Moment (CAMM) expansion, whereas the latter by dispersion function D_{as} approximating accurate SAPT(DFT) dispersion and exchange-dispersion values and developed by Pernal and coworkers.

MED approach was validated with reference MP2 results, and subsequently compared with both public domain and commercially available empirical scoring methods. MED validation was performed for five different protein–ligand systems, including receptors involved in protein–protein interactions (PPI):

- Fatty acid amide hydrolase (FAAH) with 22 N-cyclohexylcarbamic acid biphenyl-3-yl ester based inhibitors
- Menin with 18 thienopyrimidine class of inhibitors targeting menin–MLL (mixed lineage leukemia fusion protein) interaction
- *Trypanosoma brucei* pteridine reductase 1 (*Tb*PTR1) with 6 benzimidazol derivative inhibitors
- Erythropoietin producing hepatocellular carcinoma A2 receptor (EphA2) with 15 litoholic acid based inhibitors targeting EphA2–ephrin A1 interaction
- Phosphodiesterase 5 (PDE5) with 5 halogen-substituted monocyclic pyrimidinone inhibitors

Obtained results indicate that the best correlation with experiment could be obtained by the inhibitory activity scoring model MED. The proposed method covers the most important long-range interactions, i.e. multipole electrostatic and approximate dispersion and exchange-dispersion terms (D_{as}), which occur between protein and ligand molecule (here, an inhibitor). The reported results are compared to classical, commonly used empirical scoring functions and indicate the promising potential of employing the *ab initio* derived MED model reproducing rankings of relative stabilization energies into the process of inhibitor design. In addition, due to the neglect of short range exchange term, this approach is not sensitive to structural defects often present in models of the protein–inhibitor complexes.

Extension of D_{as} function presented in the dissertation includes halogen atoms and makes MED model applicable to numerous halogen-bearing inhibitors. Since the number of such molecules being exploited in the drug design process constantly rises, new D_{as} parameters considerably extend the applicability range of the MED model. Performance of MED model exploiting different D_{as} revisions available to date was also examined for series of phosphonic inhibitors of urokinase-type plasminogen activator (uPA) enzyme. In addition, MED application limits were tested for systems where solvation effects could be significant.