Evaluation of Empirical Scoring Functions for Virtual Screening Docking Experiments

Pereira, F.S.S.[1]; Guedes, I.A.[1]; Dardenne, L.E.[1]

[1]Grupo de Modelagem Molecular de Sistemas Biológicos, Laboratório Nacional de Computação Científica, LNCC/MCTI, Petrópolis, Brazil

Introduction: DockThor program (http://www.dockthor.lncc.br) obtained promising results for pose prediction in comparative studies with other molecular docking programs. We recently developed new empirical scoring functions (ESFs), based on the original DockThor scoring function, to predict protein-ligand binding affinities, which demonstrated to be competitive with the best evaluated scoring functions in the benchmarking PDBbind *core set* 2013. Moreover, for virtual screening studies, it is important to access the capacity of our ESFs to distinguish active from decoy compounds. **Objectives:** We aim to evaluate the performance of our ESFs in virtual screening experiments, using a dataset composed of actives and decoys for three relevant targets. **Material and Methods:** We used DUD-E (Directory of Usefull Decoys: Enhanced) as dataset for benchmarking virtual screening studies. We focused on three protein targets available in the DUD-E dataset: Trypsin, CDK2 and Beta-lactamase. The docking experiments for pose prediction of all active and decoy compounds are performed with the DockThor program and the binding affinities are predicted using our ESFs. Accordingly with the classification strategy adopted in the DUD-E dataset, actives and decoys are defined as the compounds with binding affinities less than 1□M and greather than 30□M, respectively. The enrichment factors are calculated to access the performance of the ESFs. **Results and Discussion:** The ESFs trained with experimental structures and also with docking results, demonstrated great performance when evaluating the benchmark PDBbind *core set* 2013 (N = 195). These results motivated the evaluation and validation of these ESFs in virtual screening experiments. **Conclusions:** The accurate classification between actives and decoys is a mandatory step to validate ESFs for virtual screening experiments. The validation of our recently developed empirical scoring functions in the DUD-E datasets is essential to guide and validate the development of the DockThor version for virtual screening.

Keywords: virtual screening, empirical scoring functions, binding affinity prediction
Acknowledgment: CAPES