Journal of Cheminformatics

POSTER PRESENTATION Open Access

DrugScorePPI for scoring protein-protein interactions: improving a knowledge-based scoring function by atomtype-based QSAR

Dennis M Krüger*, Holger Gohlke

From 5th German Conference on Cheminformatics: 23. CIC-Workshop
Goslar, Germany. 8-10 November 2009

Protein-protein complexes are known to play key roles in many cellular processes. Therefore, knowledge of the three-dimensional structure of protein-complexes is of fundamental importance. A key goal in protein-protein docking is to identify near-native protein-complex structures. In this work, we address this problem by deriving a knowledge-based scoring function from protein-protein complex structures and further fine-tuning of the statistical potentials against experimentally determined alanine-scanning results.

Based on the formalism of the DrugScore approach1, distance-dependent pair potentials are derived from 850 crystallographically determined protein-protein complexes 2. These DrugScorePPI potentials display quantitative differences compared to those of DrugScore, which was derived from protein-ligand complexes. When used as an objective function to score a non-redundant dataset of 54 targets with "unbound perturbation" solutions, DrugscorePPI was able to rank a near-native solution in the top ten in 89% and in the top five in 65% of the cases. Applied to a dataset of "unbound docking" solutions, DrugscorePPI was able to rank a near-native solution in the top ten in 100% and in the top five in 67% of the cases. Furthermore, DrugscorePPI was used for computational alanine-scanning of a dataset of 18 targets with a total of 309 mutations to predict changes in the binding free energy upon mutations in the interface. Computed and experimental values showed a correlation of R2 = 0.34. To improve the predictive power, a QSAR-model was built based on 24 residue-specific atom types that improves the correlation coefficient to a value of 0.53, with a root mean square deviation of 0.89 kcal/mol. A Leave-One-Out analysis yields a correlation coefficient of 0.41. This clearly demonstrates the robustness of the model. The application to an independent validation dataset of alanine-mutations was used to show the predictive power of the method and yields a correlation coefficient of 0.51. Based on these findings, Drugscore-PPI was used to successful identify hotspots in multiple protein-interfaces. These results suggest that DrugscorePPI is an adequate method to score protein-protein interactions.

Heinrich-Heine-University, Computational Pharmaceutical Chemistry, 40225 Düsseldorf, Germany

Published: 4 May 2010

References

1. Gohlke H, Hendlich M, Klebe G: **Knowledge-based scoring function to predict protein-ligand interactions.** *J Mol Biol* 2000, **295(2)**:337-356.
2. Huang SY, Zou X: **An iterative knowledge-based scoring function for protein-protein recognition.** *Proteins* 2008, **72(2)**:557-579.

doi:10.1186/1758-2946-2-S1-P20
Cite this article as: Krüger and Gohlke: DrugScorePPI for scoring protein-protein interactions: improving a knowledge-based scoring function by atomtype-based QSAR. *Journal of Cheminformatics* 2010 **2**(Suppl 1):P20.