Bioinformatics, 31(6), 2015, 978
doi: 10.1093/bioinformatics/btu821
Advance Access Publication Date: 12 December 2014
Letter to the Editor

OXFORD

Structural bioinformatics

Response to the comment on 'protein-protein binding affinity prediction from amino acid sequence'

K. Yugandhar and M. Michael Gromiha*

Department of Biotechnology, Bhupat and Jyoti Mehta School of BioSciences, Indian Institute of Technology Madras, Chennai-600036, Tamil Nadu, India

*To whom correspondence should be addressed.
Associate Editor: Burkhard Rost

Received on November 3, 2014; revised on November 24, 2014; accepted on December 5, 2014

Contact: gromiha@iitm.ac.in

We would like to thank Moal and Fernandez-Recio (2014) for their insightful comments on our recent publication in Bioinformatics (Yugandhar and Gromiha 2014). The main objective of the work was to develop a method for predicting the binding affinity of protein-protein complexes using sequence-based features. We have utilized both predicted binding sites and amino acid properties for developing the method. The method was formulated with inner feature selection and tested with outer leave-one-out cross-validation. Our method could predict the binding affinity of 135 protein-protein complexes within a mean absolute error (MAE) of 0.759 using leave-one-out cross validation experiment. The statistical significance of Pearson's correlation was assessed with *P*-values. We realized minor changes in the predicted binding sites at the local computer and server due to the usage of different versions of the prediction software, which moderately altered the results. Among the 39 protein-protein complexes used as a test set and six additional complexes mentioned in Moal and Fernández-Recio (2014) the binding affinity of >50% of them were predicted within a deviation of 2 kcal/mol (14 complexes with <1 kcal/mol). Further, three complexes have high sequence identity with our initial set of complexes and binding affinity of one of them differs by >4 kcal/mol due to experimental conditions and/or different binding modes, which are the limitations of our model as discussed in the article.

On the other hand, we have predicted the binding affinity of all the considered 180 protein-protein complexes using average assignment method (Saraboji *et al.*, 2006) for evaluating the performance of the present method. We have assigned the binding affinity of each complex using the average ΔG of all the 180 complexes as well as the average in each group. We obtained a MAE of 2.150 and 1.631 kcal/mol, respectively. The performance of our method is almost twice better than the average assignment method (MAE is 0.972).

We have utilized the single property correlation coefficient to relate the important property with binding affinity and combination of properties for predicting the binding affinity. For enzyme-inhibitor complexes and other enzymes, experimental ΔG varies in the ranges of 9–18 kcal/mol and 6–13 kcal/mol, respectively. The predicted binding affinities also span in these wide ranges. In addition, four typical complexes with high and low affinities in these classes are correctly predicted within an average MAE of 1 kcal/mol. The assignment of random values (or average) showed an average MAE of ∼4 kcal/mol.

Further, we have refined our model, which showed a better performance in both cross-validation and test datasets, and the details will be reported elsewhere.

Conflict of Interest: none declared.

References

Moal,I.H. and Fernández-Recio,J. (2014) Comment on 'protein-protein binding affinity prediction from amino acid sequence'. *Bioinformatics*, **31**, 614–615.

Saraboji,K. *et al.* (2006) Average assignment method for predicting the stability of protein mutants. *Biopolymers*, **82**, 80–92.

Yugandhar,K. and Gromiha,M.M. (2014) Protein-protein binding affinity prediction from amino acid sequence. *Bioinformatics*, **30**, 3583–3589.