Classification of Transient and Permanent Protein-Protein Interactions (PPI) using machine learning approaches: Implications for biologics design

Kiran Kumar. A

Department of Bioinformatics, Central University of South Bihar, Gaya, Bihar 824236 India

ABSTRACT

Protein−protein interactions (PPIs) have been classified either as permanent or transient type of interactions based on their dissociation constant and life time. Enzyme-inhibitor complex, antigen-antibody, and broad oligomeric enzyme complex structures are composed of proteins that bind tightly and permanently whereas cellular processes that are involved in disease-related pathways, such as the cell cycle, are mostly regulated by transient interactions. Understanding the intricate details of protein-protein interactions in these two different categories should guide the development and design of inhibitors that can serve as therapeutics for a great many diseases. Computational approaches are more promising as the fluorescence resonance energy transfer (FRET), bimolecular fluorescence complementation (BiFC) and other experimental methods are costly and time consuming to understand transient and permanent PPIs, Furthermore, the prediction of the type of interactions has significant implications for understanding the nature and function of PPIs.

In this study, we have classified transient and permanent category of interactions with Waikato Environment for Knowledge Analysis (WEKA) machine learning tool. We have identified and calculated 51structural and physicochemical features, responsible for such interactions for prediction of transient and permanent interactions. Using 10-fold cross-validation, 7-different algorithm were tested. C4.5decision tree model achieved the highest accuracy of 97.24%%, and yielded Matthews correlation coefficient (MCC) of 0.942 and Receiver Operating Characteristic (ROC) Curve as 0.99. The study provided new insights into the mechanism of protein association and predicting the transient and permanent interaction for effective design of biologics.

Keywords: Transient and Permanent Protein-Protein Interactions (PPI), machine learning, biologics