

DeepCys: structure-based multiple cysteine function prediction method trained on deep neural network: case study on Domains of Unknown Functions (DUFs) belonging to COX2 family

Vamsi Mohan Nallapareddy¹, Shubham Rajendra Bogam¹, Himaja Devarakonda¹, Shubham Paliwal¹, and Debashree Bandyopadhyay¹

¹Birla Institute of Technology and Science, Pilani, Hyderabad Campus

October 16, 2020

Abstract

Cysteine (Cys) is the most reactive amino acid participating in a wide range of biological functions. In-silico predictions complement the experiments to meet the need of functional characterization. Multiple Cys function prediction algorithm is scarce, in contrast to specific function prediction algorithms. Here we present a deep neural network-based multiple Cys function prediction, available on web-server (DeepCys) (<https://deepcys.herokuapp.com/>). DeepCys model was trained and tested on two independent datasets curated from protein crystal structures. This prediction method requires three inputs, namely, PDB identifier (ID), chain ID and residue ID for a given Cys and outputs the probabilities of four cysteine functions, namely, disulphide, metal-binding, thioether and sulphenylation and predicts the most probable Cys function. The algorithm exploits the local and global protein properties, like, sequence and secondary structure motifs, buried fractions, microenvironments and protein/enzyme class. DeepCys outperformed most of the multiple and specific Cys function algorithms. This method can predict maximum number of cysteine functions. Moreover, for the first time, explicitly predicts thioether function. This tool was used to elucidate the cysteine functions on domains of unknown functions (DUFs) belonging to cytochrome C oxidase subunit-II (COX2) like transmembrane domains. Apart from the web-server, a standalone program is also available on GitHub (<https://github.com/vam-sin/deepcys>)

Hosted file

DeepCys_Structure_12102020.pdf available at <https://authorea.com/users/367699/articles/487033-deepcys-structure-based-multiple-cysteine-function-prediction-method-trained-on-deep-neural-network-case-study-on-domains-of-unknown-functions-dufs-belonging-to-cox2-family>