DETECTION OF CLEAVAGE SITES FOR HIV-1 PROTEASE IN NATIVE PROTEINS

Liwen You
Computational Biology and Biological Physics Group
Department of Theoretical Physics, Lund University
Sölvegatan 14A, SE-22362, Lund, Sweden

Intelligent Systems Lab
School of Information Science, Computer and Electrical Engineering,
Halmstad University
Box 823, SE-30118, Halmstad, Sweden

Predicting novel cleavage sites for HIV-1 protease in non-viral proteins is a difficult task because of the scarcity of previous cleavage data on proteins in a native state. We introduce a three-level hierarchical classifier which combines information from experimentally verified short oligopeptides, secondary structure and solvent accessibility information from prediction servers to predict potential cleavage sites in non-viral proteins. The best classifier using secondary structure information on the second level classification of the hierarchical classifier is the one using logistic regression. By using this level of classification, the false positive ratio was reduced by more than half compared to the first level classifier using only the oligopeptide cleavage information. The method can be applied on other protease specificity problems too, to combine information from oligopeptides and structure from native proteins.