PREDICTING THE BINDING AFFINITY OF MHC CLASS II PEPTIDES

Fatih Altiparmak[1], Altuna Akalin[2], Hakan Ferhatosmanoglu[1]
[1]Computer Science and Engineering, The Ohio State University
[2]Computational Biology Unit, Bergen Center for Computational Science,
University of Bergen

MHC (Major Histocompatibility Complex) proteins are categorized under the heterodimeric integral membrane proteins. The MHC molecules are divided into 2 subclasses, class I and class II. Two classes differ from each other in size of their binding pockets. Predicting the affinity of these peptides is important for vaccine design. It is also vital for understanding the roles of immune system in various diseases. Due to the variability of the locations of the class II peptide binding cores, predicting the affinity of these peptides is difficult. In this paper, we proposed a new method for predicting the affinity of the MHC Class II binding peptides based on their sequences. Our method classifies peptides as binding and non-binding. Our prediction method is based on a 3-step algorithm. In the first step we identify the informative n-grams based on their frequencies for both classes. In the next step, the alphabet size is reduced. At the last step, by utilizing the informative *n*-grams, the class of a given sequence is predicted. We have tested our method on the MHC Bench IV-b data set [13], and compared with various other methods in the literature.