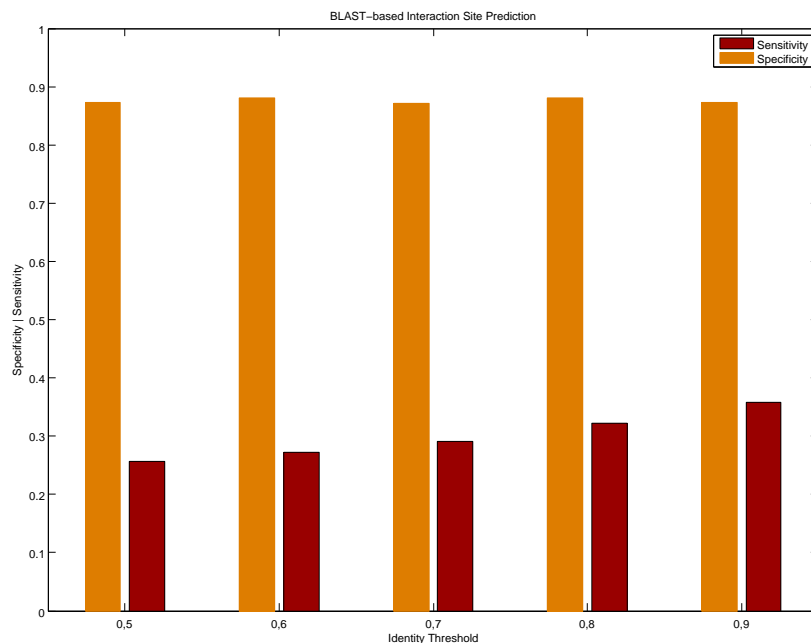# BLAST-based interaction site prediction

A simple method for the transfer of knowledge concerning the location of functional sites relies on a BLAST search. This comparative approach was applied to the training data of the ipHMM method. All homologous sequences were detected for every sequence from this set using BLAST (version 2.2.13). Clustering effects of sequences could be observed in the data set due to sequence redundancies within the PDB. In order to establish a better comparability to the HMM-based method, only hits up to an identity threshold ranging from 90 - 50 % and more than 60 % coverage were considered as a source of prediction to get an overview of the decrease of predictive power from closely to remotely related sequences. The coverage is calculated as the percentage of the query sequence which is covered by the BLAST alignment to the considered hit. The interaction site profile of each first hit sequence that fits the thresholds was subsequently transferred to the query sequence according to the alignment of the BLAST program. Predictions were restricted to the area which was covered by the BLAST alignment. The prediction quality was evaluated in the same way as described for the ipHMM methodology. As expected, the prediction quality drops with lower sequence identity. For an identity threshold of 90% the method predicted 35.88% of all observed interaction sites correctly while 87.23% of all non-binding positions were precisely detected. In case of 50% maximum identity the sensitivity decreases to 25.65% whereas the specificity remains at the level of 87%. Regarding the more important prediction quality of interaction sites the BLAST-predictor performed significantly worse than ipHMMs.