



Figure S6: **Classification framework** Classification framework of gradient boosting trees, to deal with a large number of residue-based features as well as allow for intuitive selection of the most predictive residues. A separate gradient boosting tree (GbT-s) is trained on each kind of feature across all aligned residue indices. The weights obtained from all the GbT-s classifiers are pooled (by taking the sum of the normalized weights) to select the top 30 residues. All feature values of these selected residues are used as input to the final classifier GbT. The parameter settings of the GbT-s and GbT classifiers are listed in Methods.