



Figure S6. The ROC curves of MEDscore and two simple conservation scoring methods [i.e. the Shannon entropy (SE) and Shannon entropy with residue properties (RPSE)]. SE of a residue is defined as: $SE_i = - \sum_{\chi \in AA} p_i(\chi) \times \ln p_i(\chi)$ where χ belongs to one of the 20 amino acids (AA) and $p_i(\chi)$ is the probability of χ in the i^{th} column of the multiple sequence alignment (MSA). To obtain the RPSE value of a given residue, we first classified 20 amino acids into the following groups: aliphatic [AVLMC], aromatic [FWYH], polar [STNQ], positive [KR], negative [DE] and [GP]. Then, the RPSE value could be obtained by computing the entropy according to this classification. The same MSAs used in the calculation of CONscore were used and the program developed by Capra and Singh (2007) was employed to determine SE and RPSE.