



**Figure S1.** (A) Heat plot of log-odd scores for observed versus expected nucleotide frequencies in negative examples for each amino acid in each of the 16 contacts; (B) the corresponding  $\chi^2$  values.