



**Figure S4. Pair-wise comparisons of the mean of I-RMSDs of the highest ranked conformation selected by different docking scoring methods for each case on the ZDock3-BM3 decoy set using the Nemenyi test.** Methods that are not significantly different (at significance level  $\alpha = 0.05$ ) are grouped together (via connecting lines). For each of the 67 docking cases considered, these scoring functions are ranked based on the I-RMSD of their respective top ranked conformation. Thus, the scoring function whose top ranked conformation has the lowest I-RMSD appears at the top of the ranking of scoring functions. The number on the line of each method is the average rank of each method over all docking cases. The critical difference determined by the Nemenyi test at significance level 0.05 is 1.10. Hence, the difference between any pair of scoring schemes is statistically significant if the difference between the average ranks of the corresponding scoring functions is greater than 1.10. Thus the mean of I-RMSDs of the highest ranked conformations selected by DockRank is significantly smaller than those selected by IRAD, ZRank, and the variants of DockRank's scoring using the predicted residues from NPS-HomPPI, PrISE, and meta-PPISP. The mean of I-RMSDs of the highest ranked conformations selected by DockRank (with partner-specific predicted interfaces by PS-HomPPI) is not significantly different from DockRank's scoring function supplied with actual partner-specific interface-based method (labeled as "PS-Act Int").