Text S5. Statistical Analysis of the Highest Ranked Models Selected by DockRank and Other Methods on ClusPro2-BM3_31 Decoys

Our analysis, using the Friedman test, shows that the null hypothesis that the mean L-RMSDs of the highest ranked models selected by DockRank (using PS-HomPPI predicted interfaces), ClusPro (cluster-sized based), ClusPro Center Energy-based and ClusPro Lowest Energy-based scoring functions are the same can be rejected with high confidence (p-value < 0.0001). We applied the Nemenyi test to determine whether the differences of L-RMSDs of the highest ranked conformations identified by any given pair of scoring functions are statistically significant. The results of Nemenyi test (Figure S6) suggest that, at a significance level of 0.05, the mean L-RMSDs of the highest ranked conformations identified by DockRank are significantly lower than those selected by ClusPro Center Energy. However, although the top conformations ranked by DockRank have lower average L-RMSD values than those ranked by ClusPro and ClusPro Lowest Energy scoring functions, the difference is not statistically significant. This lack of statistically significant difference could be due to the limited power of the non-parametric test, which is undermined by the small number of cases (31) and the small number of docked conformations (~30) that DockRank used in the comparisons.