Text S2. Estimating the weighting factor β in DockRank's scoring function

In this study, we used a set of transient complexes with experimentally determined interface residues (3.25 Å or better resolution, determined using X-ray crystallography), Trans208, to estimate the weighting factor β used by DockRank's scoring function (see "DockRank's Scoring function" in Methods). Trans208 is derived from Trans212, a dataset of 212 transient complexes originally generated by Mintseris and Weng [71]. Trans208 is obtained by replacing 2 complexes with their more recent versions in PDB and discarding 4 others that have since become obsolete or otherwise unusable. The complexes replaced by new structures in PDB are: 1DN1 (superseded by 3C98) and 1IIS (superseded by 1T83). Four complexes discarded from Trans212 are: 1F83 (1F83 was superseded by 3G94, which was obsoleted on 2009-07-14 and superseded by none), 1DF9 (obsoleted on 2010-02-09 and superseded by none), 4CPA (the latest structure of 4CPA in PDB does not have chain I and 0 as recorded in the original Trans212 dataset), and 1JCH (1JCH cannot be processed by NACCESS [72], a software that we used in the process of calculating interface residues.).

We calculated the partner-specific (PS) transient interface residue ratio using Trans208, which was used as the weighting factor β of DockRank's scoring function. PS transient interface residue ratio was calculated by dividing the total number of PS transient interface residues by the total number of residues computed from the set of partner-specific interfaces in the complexes in Trans208. The PS-interface residues of a chain are defined with respect to each interaction partner. Suppose we have a protein complex AB:C, where chain A and chain B form a transient interaction with chain C. Then chain C has two partner-specific interfaces (one with A and the other with B) and hence two sets of partner-specific interface residues. For each partner-specific interface, we calculate the number of interface residues as simply the sum of interface residues in each chain with its partner. Thus, the number of partner-specific interface residues in the interface formed by chain C with chain A is the number of residues of chain C that interact with chain A plus the number of residues of chain A that interact with chain C. The total number of residues in the interacting pair of chains A and C is simply the number of residues of A plus the number of residues of C. The total number of partner-specific interface residues in Trans208 is obtained by adding up the numbers of interface residues for each partner-specific interface in each of the complexes in Trans208. The total number of residues in chains that are involved in partner-specific interactions is obtained by simply adding up the total numbers of residues in the two chains that participate in each partner-specific interaction in each of the complexes in Trans208. Following this procedure, we find that Trans208 has 11,396 PS transient interface residues and 146,198 total residues (in the sense defined above). Hence, we estimate β, the ratio of the number of PS transient interface residues to the total number of residues to be $\frac{11396}{146198}$=0.08.