**Text S1. PS-HomPPI (Partner-Specific Sequence Homology based Protein-Protein Interface Prediction**

Our interface predictor PS-HomPPI consists of two major modules:

- PS-interface conservation analysis: PS-interface conservation analysis (shown as a PCA biplot in Figure S1) is based on a dataset of transient dimers with experimentally determined interface residues. For each dimer $A : B$, sequence homologs $A' : B'$ with known interfaces are retrieved along with sequence alignment measures (e.g., E-value returned by BLASTP). An interface conservation score (IC-score) is calculated based on the similarity of the interfaces of query pair $A : B$ and those of $A' : B'$. The higher the IC-score, the more similar the interfaces of query pair and its homo-interolog. Based on the IC-scores, three interface conservation zones are identified: the Safe Zone, where the interfaces are most conserved among homo-interologs; the Twilight Zone, which denotes the area of medium interface conservation; and the Dark Zone, which denotes the area of lowest interface conservation. A regression model of the relation between IC-score and the sequence alignment measures is constructed. This model is used to estimate the IC score of a homo-interolog during the prediction phase.

- PS-interface prediction: To predict the interface residues between a pair of proteins $A : B$, the homo-interologs of $A : B$ with known interfaces are retrieved. For each homo-interolog $A' : B'$ of $A : B$, sequence alignment measures between $A$ and $A'$, and between $B$ and $B'$ are obtained using BLASTP. The regression model is used to estimate the IC-score for each homo-interolog of $A : B$. If at least one homo-interolog in the Safe Zone is found,, PS-HomPPI uses the Safe Zone homo-interolog(s) to infer the interfaces of the query protein. Otherwise, the search is repeated for homo-interologs in the Twilight and Dark Zones. The top $K$ ($K = 10$ by default. $K < 10$ if there are no enough homo-interologs available) homo-interologs within a single zone are used for interface inference. The zone of homo-interologs used in predictions provides the prediction confidence level.

The default parameters of PS-HomPPI were used in this study. For detailed parameter settings, please refer to Xue et al. 2011 [42].