Supplementary Material

Figure S1. Overview of *in silico* PPI chemical libraries resulting from this work. 25 chemical libraries from the main providers representing more than 8.3 million compounds were filtered with our SVM-based $2P2I_{HUNTER}$ tool and 'Ro4' rules to lead to the reference PPI library ($2P2I_{REF}$ composed of 143,218 compounds corresponding to 22,845 diverse compounds as estimated with Tanimoto index 0.8). a) We then selected compounds containing at least one privileged scaffold as recently defined by Welsch and colleagues [27]. The resulting 35 sets of compounds containing a given privileged scaffold were pooled together and duplicates were removed. The resulting $2P2I_{PRIV}$ library contained 51,476 compounds. b) A diverse representative subset ($2P2I_{DIV}$) composed of 8,217 compounds was built from $2P2I_{PRIV}$ in three steps. i/ Optimizable k-dissimilarity (OptiSim) selection implemented in TRIPOS [24] (Tanimoto cut-off= 0.8) was applied to each of the 35 privileged structure sets, ii/ compounds were then combined in one library, and duplicates were removed iii/ a diversity selection was performed on the combined set with OptiSim (Tanimoto cut-off= 0.8) . c) Finally, compounds with higher three-dimensionality were selected to escape from flatland, using carbon bond saturation descriptor as defined by fraction sp3 (Fsp3) where Fsp3 is the number of sp3 hybridized carbons divided by the total carbon count [37]. The resulting diverse PPI targeted library ($2P2I_{3D}$) is composed of 1,683 small molecules.

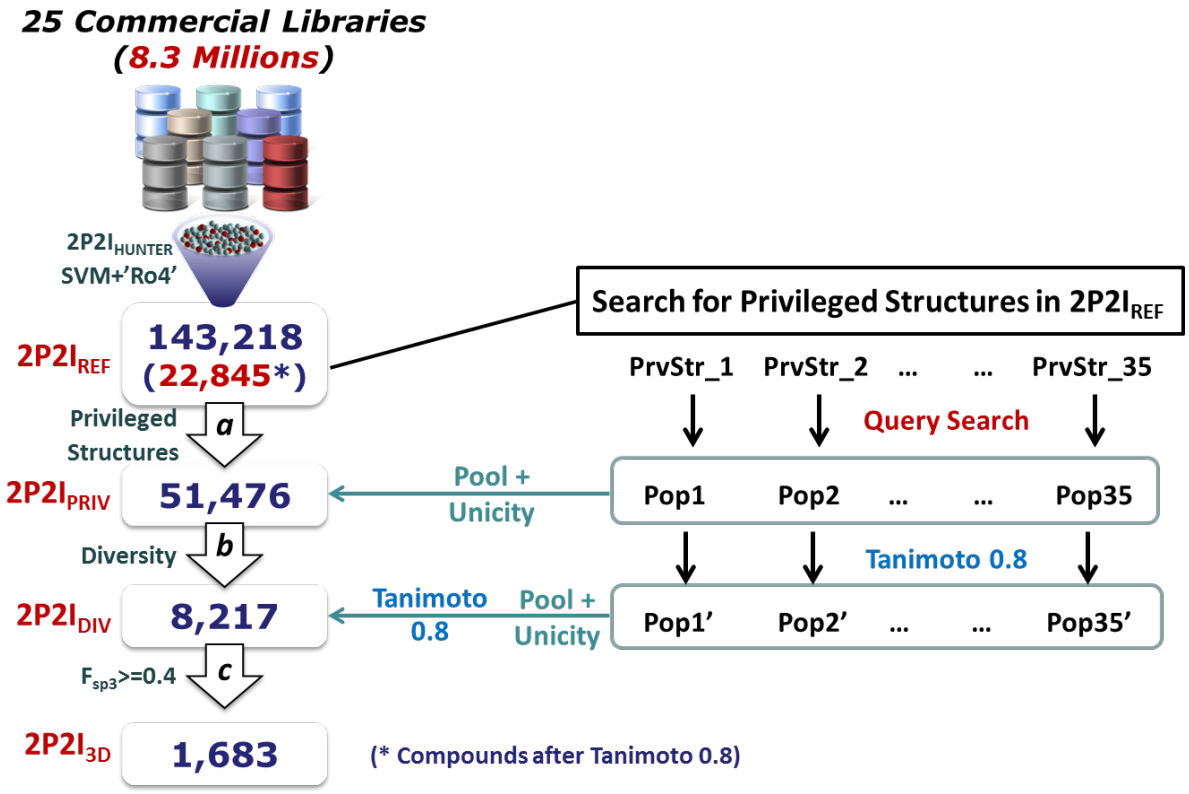

Figure S1

Figure S2: Cumulative percent distribution of 2-ring scaffolds in $2P2I_{REF}$. Another representation of these data expressed in numerical population for each of the 3,973 scaffolds is found in the bubble-shaped Figure 2. Here, the graph shows a rapid growth as the top 1% of the total scaffolds (top 40 most populated scaffolds) stands for the ancestors of as much as 50% of the compounds in $2P2I_{REF}$ (inset).

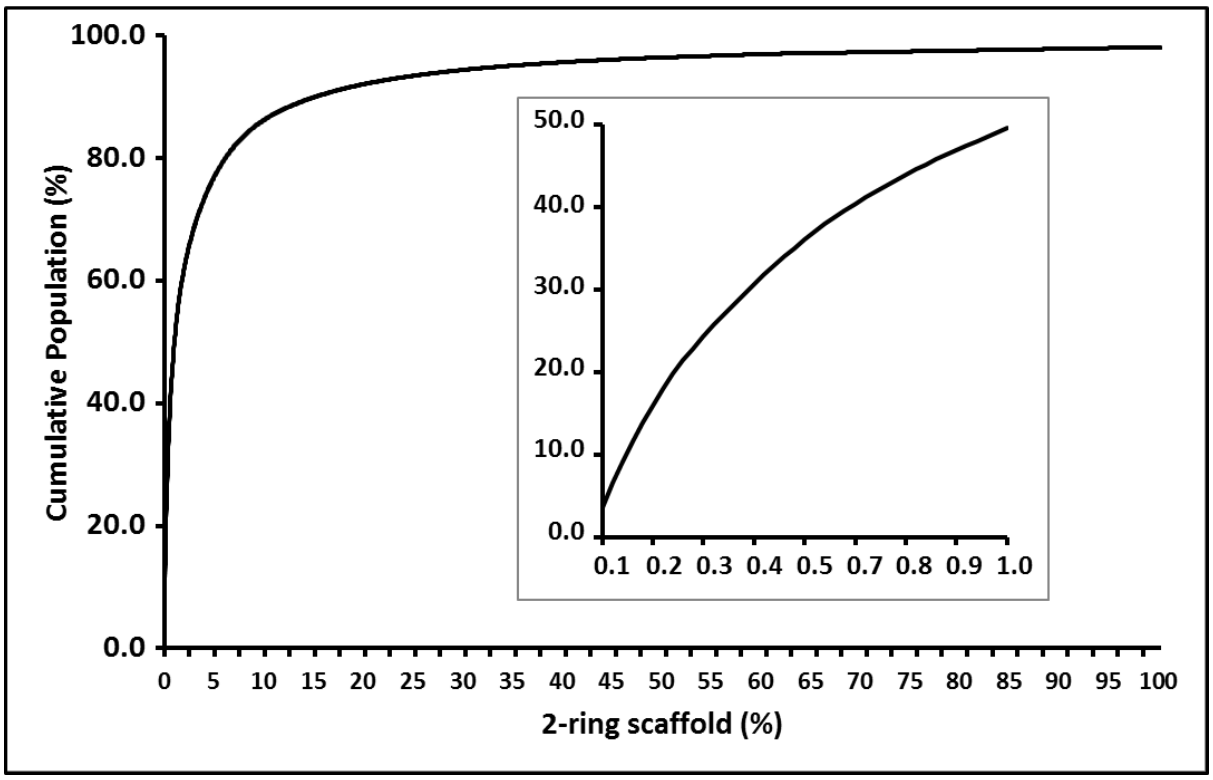

Figure S2

Figure S3: 2D PCA plot of training dataset used in $2P2I_{HUNTER}$ and $2P2I_{3D}$ chemical library. The 1,683 compounds (light blue spheres) from the final *in silico* PPI-targeted library $2P2I_{3D}$ were projected in the chemical space of compounds used as the training set during the development of $2P2I_{HUNTER}$.[22] Orthosteric PPI modulators used as a positive set in the SVM model are shown as red squares, whereas decoy compounds are shown as grey spheres.

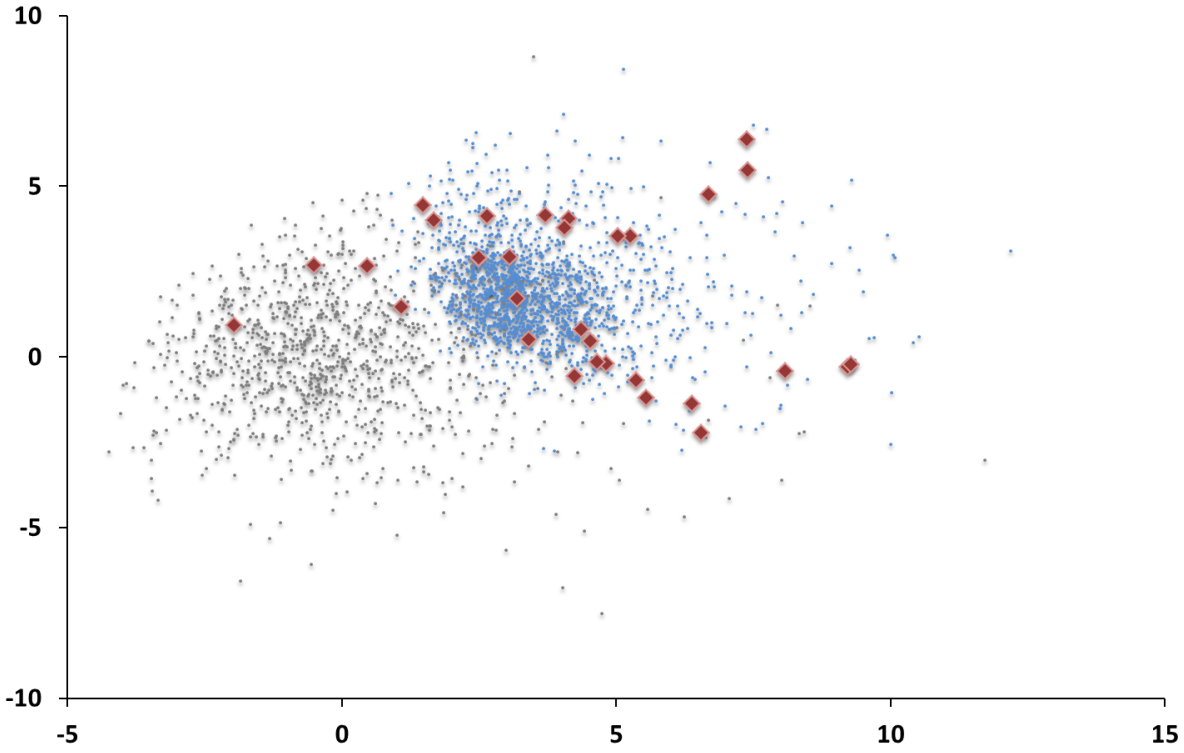

Figure S3

Figure S4: Percentage of compounds in $2P2I_{3D}$ that can be purchased from each provider. Only providers from which at least 5% of the compounds can be acquired are labeled in the picture. (The total amount is greater than 100% because the same compound can be obtained from different providers).

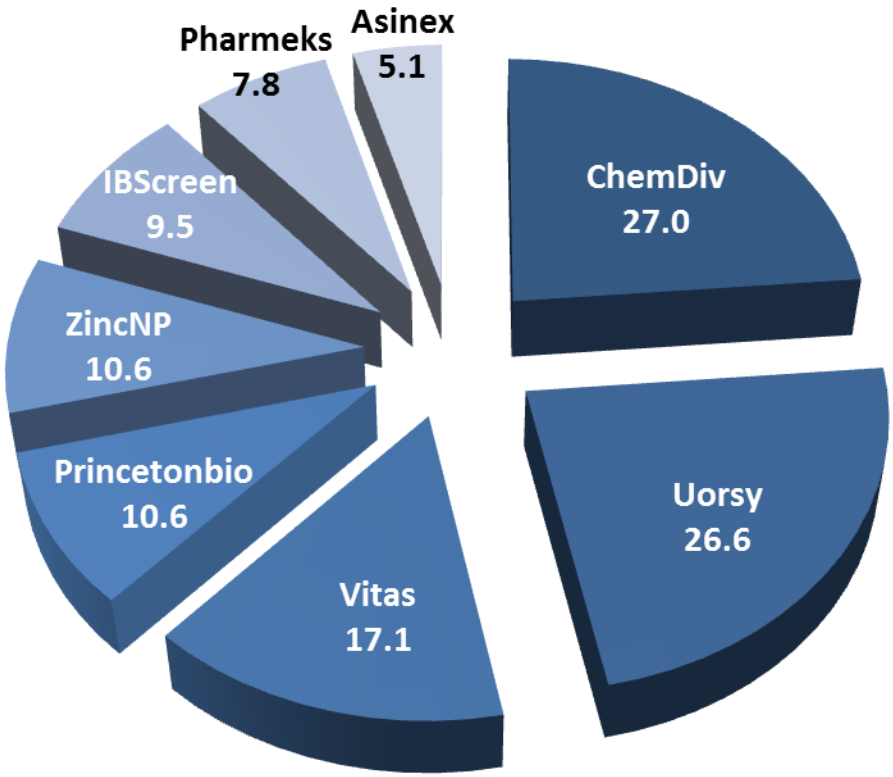

Figure S4