Probabilistic Pharmacophore Matching for Structure-Based Virtual Screening

Boris Klebansky and Richard M. Fine BioPredict, Inc. Oradell, NJ, 07649

Many problems in computational chemistry can be reduced to finding a match of chemical features in three dimensions. Problems that can be tackled in this way include docking ligands into the active site of a protein. Here we apply a Markov Random Fields (MRF) engine to this problem. Ligand docking is performed as a graph match of an abstracted description of the ligand to an abstracted description of the active site. The abstracted descriptions are graphs, whose nodes are chemical entities and whose edges are associated distance constraints. The weighted graph-matching problem is expressed as an MRF model, whose solution minimizes its associated free energy function. The resulting solution is a maximal a posteriori probability (MAP) distribution that statistically describes optimal (in MAP) placements of the fragment into the active site. Individual low-energy placements of the fragment are obtained by marginalizing the MAP distribution. MRFs can be combined, allowing simultaneous probabilistic docking into multiple models (ensembles) of protein active sites to represent protein families and/or to account for protein flexibility. This results in newly defined consensus protein models that mix features of constituent models. MRFs can incorporate sets of probabilistic beliefs, expressed as probabilistic prior distributions. These can be used to bias matches according to known actives (focused docking). We acknowledge grant support from NIH-GM071055.