Identification of co-evolving sites in the ligand binding domain of G protein-coupled receptors using mutual information$^1$ SAROSH N. FATAKIA, STEFANO COSTANZI, CARSON C. CHOW, LBM, NIDDK, NIH — G protein-coupled receptors (GPCRs) are the largest superfamily of membrane proteins in humans. They are involved in signal transduction in numerous cellular processes and are the most common target for pharmacological intervention via activation or inhibition. Identification of functionally important sites is relevant for better understanding the ligand-receptor interaction and therefore for drug delivery.

In a superfamily of proteins, functionally important but co-evolving sites are not easily identified in a multiple sequence alignment (MSA). Using a MSA of transmembrane (TM) domains of GPCR superfamily, we identify sites which co-evolve, and may therefore be functionally important. Assigning the TM site as a node and the MI of site pairs as an inverse inter-node distance, a MI graph is established. Co-evolving sites are then identified via this graph. Nodes characterized by high connectivity are located within the commonly accepted ligand binding site of GPCRs, suggesting that concerted co-evolution of a number of neighboring residues gave rise to a multitude of subfamilies each recognizing a specific set of ligands. MI and graph analysis may serve as a tool for the identification of topologically conserved binding pockets in the families of evolutionarily related proteins.

$^1$This work was sponsored by the intramural research program of NIH, NIDDK