Computational and AI Approaches for GPCR Prediction and Hierarchical Classification

Khodeza Begum Mitchell,[1,2] Jonathon Mohl,[2,3,4] and Ming-Ying Leung[2,3,4]

[1]Department of Biological Sciences, [2]Bioinformatics Program, [3]Border Biomedical Research Center, and [4]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX 79968, USA kbegum@utep.edu

Abstract

G protein-coupled receptors (GPCRs) are one of the largest families of membrane proteins and play key roles in cellular signaling. GPCRs consist of seven transmembrane helices with conformational changes induced by ligand binding. They are major targets of prescribed drugs and are widely studied across many diseases including neurological disorders, metabolic conditions, and cancer. GPCR function is determined by how ligands interact with the receptor, yet ligand characterization remains incomplete for many GPCRs. In the past two decades, a range of methods has been developed for GPCR prediction and classification, including homology-based techniques, feature-based machine learning models, and more recent AI-based approaches such as deep neural networks and transformer-based models. These methods use amino acid sequence-derived descriptors and learned representations to capture patterns related to GPCR function, supporting classification at the level of family and subfamily, with limited extension to subtype. Many existing approaches are designed for specific GPCR classes or datasets, and their performance varies across receptor groups due to differences in sequence characteristics and data availability. This talk presents an overview of computational and AI approaches for GPCR prediction and classification, with emphasis on commonly used descriptors, current limitations, and areas for further development.