Identifying GPCR/ligand binding sites and binding modes using concensus regression

Martyn Ford,[1] David Salt, [1] Philip Evans, [1] David Whitley[1] and Steve Garland[2]

[1]Centre for Molecular Design, Institute of Biomedical and Biomolecular Sciences, King Henry Building, King Henry I Street, University of Portsmouth PO1 2DY; [2]GlaxoSmithKline, Harlow

G-Protein-Coupled Receptors (GPCRs) are a class of membrane-bound receptors of considerable interest to the pharmaceutical industry [1]. Around 40 - 50 % of marketed drugs exert their effects at this class of receptor [2, 3] and much of current pharmaceutical research effort is directed towards the discovery of novel agonists or antagonists for these receptors. GPCRs are therefore the most important class of drug targets known. The aims of this research are to identify (i) which positions on the trans-membrane helices (TMHs) and (ii) which amino acids at these positions are important for ligand binding. The study also aims to test the hypothesis that the properties of the residues at these positions explain the observed ligand selectivity to GPCRs and are therefore important variables determining the strength of ligand binding.

Current research aimed at identifying ligand binding sites within GPCRs is based on homology modelling and trawling the literature for evidence of binding mode. This can be a lengthy and time-consuming process. Developments in HTS technology have established a need for a more rapid and robust procedure. An *in silico* procedure, based on exhaustive all subsets regression, correspondence analysis and consensus modelling has now been developed at Portsmouth. The approach leads to robust models that can be used for reliable prediction of novel compounds acting at mutated receptors or those from different subfamilies of GPCRs. The methodology is largely automatic and enables drug design scientists to identify the molecular interactions between a ligand and a range of GPCRs with reliability and speed. The identity and location of the side chains involved in binding are determined. Moreover, the key property (size, electrostatic potential or hydrophobicity) of each side chain involved in the interaction is also determined. This information provides a strong experimental basis for interpretation of the selectivity profiles of a candidate drug.

This presentation will describe the algorithm and apply the approach to several examples. These include the familiar aminergic GPCRs, but will also extend to other GPCR families. The predictions have been validated against reports in the literature. The software provides a tool that should help industrial scientists to improve the drug discovery process, leading to more rapid development of ligands with high selectivity and affinity for a named receptor target.

1. Rebois, R.V., B.G. Allen, and T.E. Hebert, *The targetable G protein proteome: where is the next generation of drug targets?* Drug Discovery Today: TARGETS, 2004. **3**(3): p. 104-111.
2. Flower, D.R., *Modelling G-protien-coupled receptors for drug design.* Biochimica et Biophysica Acta, 1999. **1422**: p. 207-234.
3. Gudermann, T., B. Nurnberg, and G. Schultz, *Receptors and G-proteins as primary components of transmembrane signal transduction.* Journal of Molecular Medicine, 1995. **73**: p. 51-63.