AIMS Biophysics, 4(4): 543-556.
DOI: 10.3934/biophy.2017.4.543
Received: 04 August 2017
Accepted: 17 September 2017
Published: 26 September 2017

http://www.aimspress.com/journal/biophysics

Research article

Common evolutionary binding mode of rhodopsin-like GPCRs: Insights from structural bioinformatics

Eda Suku [1] and Alejandro Giorgetti [1,2,*]

[1] Department of Biotechnology, University of Verona, Ca' Vignal 1, Strada le Grazie 15, 37134 Verona, Italy
[2] Computational Biomedicine, Institute for Advanced Simulation IAS-5 and Institute of Neuroscience and Medicine INM-9, Forschungszentrum Jülich, 52425 Jülich, Germany

* **Correspondence:** Email: alejandro.giorgetti@univr.it.

Abstract: G-protein Coupled Receptors (GPCRs) form the largest membrane protein superfamily in vertebrates. Advances in crystallization techniques so far resulted in the resolution of 44 unique receptors available for the GPCRs researcher's community, 37 of which belong to rhodopsin-like GPCRs class. We performed here the first systematic analysis of GPCRs binding cavities based on the available pool of rhodopsin-like solved structures. We pinpointed ten positions shared between all the solved receptors, namely 3.32, 3.33, 3,36, 6.48, 6.51, 6.52, 6.55, 7.35, 7.39 and 7.43, as interacting with ligands. We analyzed the conservation of amino acids present in these positions and clustered GPCRs accordingly to the physicochemical properties of binding cavities' residues. Clustering supplied new interesting insights into the common binding mode of these receptors. In particular, the 3.32 position turned out to have an important role in ligand charge detection. Finally, we demonstrated that residues in these ten positions have co-evolved together, sharing a common evolutionary history.

Keywords: GPCRs; bioinformatics; structural analysis; clustering; GPCRs evolution

1. Introduction

The intricate process of cell's signal transduction begins with a transmembrane receptor, which interacts with signaling molecules in the extracellular part and transmits the signal towards the

internal part of the cell. The major contribution to cell signal transmission cascades (80%) is given by a class of transmembrane receptors, i.e. G-protein-coupled receptors (GPCRs) [1]. Indeed, the GPCRs superfamily is the largest group of eukaryotic membrane receptors, with about 850 members in the human proteome [2,3]. Upon extracellular agonists binding, GPCRs are fully-activated by interacting with heterotrimeric G-proteins (Figure 1), regulating in this way downstream second messengers or protein kinase cascades [1].

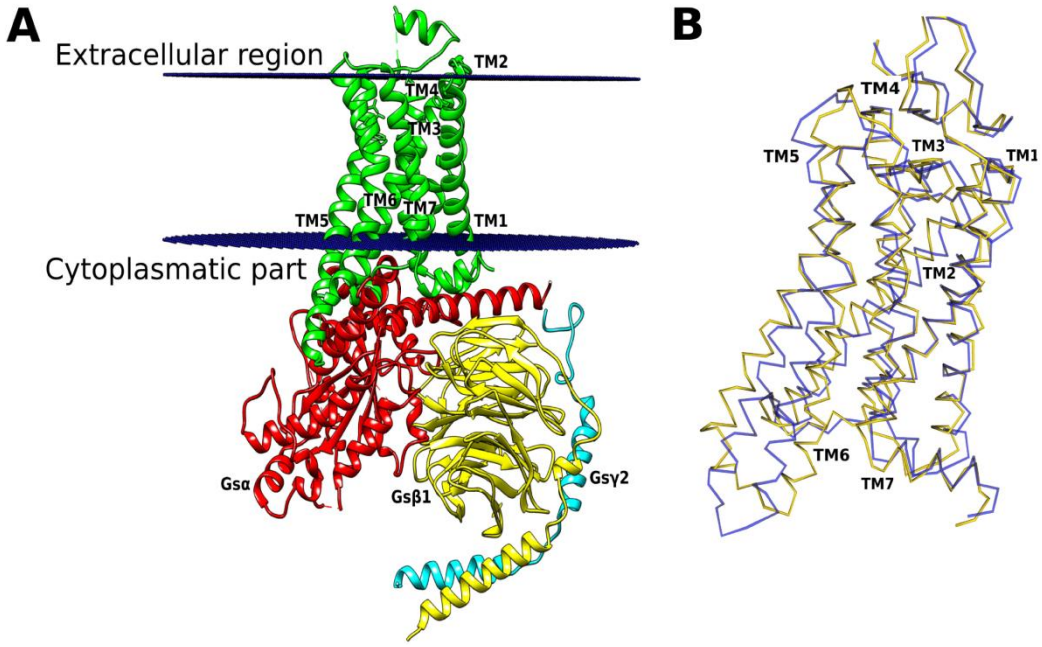

Figure 1. Class A GPCR structure and activation. (A) Schematic representation of a GPCR, the β2-adrenergic receptor in complex with Gs-protein. (B) Activation-induced conformational changes in the backbone of class A GPCRs exemplified here by the β2-adrenergic receptor: its inactive structure (PDB ID: 2RH1) is in blue and its active structure (PDB ID: 3SN6) is in yellow. The receptor has the same orientation in (A) and (B).

The activation of GPCRs is reflected in specific conformational changes of the seven transmembrane helixes (TM1-7). Indeed, upon activation, a huge rearrangement of the so-called highly conserved "intramolecular switches" takes place, that includes helixes 2, 3, 6 and 7 [4,5]. These switches connect the ligand binding site with the cytoplasmic part of GPCRs, transmitting the signal through the transmembrane helixes as pieces of a connected puzzle. More specifically, TM6 moves away from the helical core of ~10Å, while TM7 moves toward the core [6,7]. Together with TM2, 3 and 6, the movement of TM7 disrupts the "hydrophobic barrier" formed between these helixes. This disruption causes the collapse of the allosteric sodium binding site, positioned in the half down part of the receptor. As the "allosteric" sodium ion stabilizes the inactive conformation of the GPCR, the collapse of its binding site thus leads to the activation state of the receptor [8].

According to the GRAFS system, human GPCRs can be grouped into five main subfamilies [2,9] named: glutamate, rhodopsin, adhesion, frizzled/taste2 and secretin. Glutamate receptors respond to the neurotransmitter glutamate [10]; adhesion receptors are characterized by

extremely large N-terminus, containing various adhesion domains responsible for mediating cell-cell and cell-matrix interactions [11]; frizzled receptors serve as receptors in the Wnt (the acronym Wnt is a fusion of a Drosophila's gene name, *wingless*, and the correspondent vertebrate homolog, *integrated* or *int*) signaling pathway [12] and secretin receptors bind secretin, a hormone that regulates water homeostasis [13]. The rhodopsin-like GPCRs class (rho-GPCRs hereafter), is a widespread protein family that adopts the common structural framework of rhodopsin, the first structurally solved GPCR [14,15]. The rho-GPCRs class is the most heterogeneous one. Indeed, receptors belonging to this class are activated by a broad range of endogenous agonists such as small organic molecules, peptides, proteins, ions, and even light as in the case of rhodopsin.

The first successful GPCR crystallized structure, in 2000 [15], was the bovine rhodopsin receptor and so far, 44 unique receptor structures have been determined by X-ray protein crystallization and nuclear magnetic resonance (NMR) techniques, 37 of which belong to the rho-GPCRs class [16]. Very recently, the activated glucagon (GLP-1) receptor structure, in complex with a G-protein was solved, using, for the first time on a GPCR, the cryo-electron microscopy (cryo-EM) strategy [17].

At a pharmaceutical level, GPCRs are crucial targets since they bind ~40% of all the commercial drugs [18,19]. Thus, solving the structure of GPCRs and characterizing the molecular mechanism of activation could have a great importance for human health. On this purpose, a massive number of studies have been proposed since the first GPCR structure was solved. These studies regard principally the characterization of rho-GPCRs structural features [3, 20,21,22] or analysis of the interactions with their cognate ligands [23,24]. Regarding the first point, the analysis of solved rho-GPCRs has revealed some common structural features, such as the presence of a highly conserved residue in each of the transmembrane helices. Indeed, this is the origin of the Ballesteros-Weinstein numbering (or generic GPCRdb numbering) of GPCRs, in which the first number indicates the helix and the second number indicates the residue position with respect to the most conserved residue in that helix (x.50). For example, position 3.52 refers to a residue in helix 3, two positions after the most conserved residue, the 3.50. Members of the rho-GPCRs class share a set of highly conserved motifs, such as for example, D[E]R$^{3.50}$Y in helix III [25]. This common motif forms the so-called "ionic lock", that is formed only during the inactivation of GPCRs. Another crucial motif, the NP$^{7.50}$xxY in helix VII, is considered to play an important role in GPCRs rearrangement upon activation [26,27]. Finally, a highly conserved disulfide bridge between Cys$^{3.25}$ located in helix III and a cysteine residue in ECL2 was observed in most of the GPCR structures [28].

Concerning the general ligand-binding modes of the GPCRs, size and shape of the binding cavities change dramatically between classes. The binding cavity of glutamate receptors is located much deeper compared to those of the other classes [29]. Contrarily, in the (smoothened) SMO structure, a frizzled class receptor, the ligand lies in a long and narrow binding pocket, interacting mostly with extracellular loops [29]. Similar differences can be observed also within the same classes. In rho-GPCRs class, the ligands that bind to chemokine receptors are positioned in cavities that are mostly located in the extracellular part, showing a wider extracellular opening compared to the ligand binding cavities of the receptors that bind small ligands like aminergic receptors [30].

What is missing so far, to get deeper insights into the function of the GPCRs superfamily and in particular rho-GPCRs class, is a global characterization of the binding cavities of these receptors with the perspective of uncovering common binding characteristics, preserved in this class along millions of years of evolution. Gaining knowledge on common features for ligand recognition shared

among all the five classes may become crucial to deeply understand the mechanism used by these receptors to transmit the signal into the cell.

The aim of this work is to analyze the rho-GPCRs solved structures [16,31] in order to assess if and how shared binding features, can give insights into the evolutionary history of these receptors. We analyzed 85 unique receptor-ligand crystallographic structures and for each complex we extracted all the positions involved in ligand (hereafter in the definition of ligand we include small non-peptide compounds and peptides) interactions. Following the reasoning that ligand binding could be part of the general GPCRs architecture, we expect to find, on one hand, well-conserved positions required for ligand interactions and on the other hand, highly variable residues located in these positions, responsible for the selective differences within each GPCRs subfamily. These differences should result in a recognition of a wide and different range of ligands, as described before. Indeed, by restricting our analysis to the binding sites of GPCRs and clustering the receptors accordingly to their shared features, we identified and validated ten functionally relevant co-evolved positions. As a cross-check, we showed that these positions are enough to cluster all rho-GPCRs accordingly to their evolutionary history.

2. Materials and Method

We retrieved all the solved rho-GPCRs structures from the PDB database [32]. We eliminated all the apo structures and receptor-ligand complexes that were not unique. We performed our analysis on 85 GPCRs complexes (see Supplementary Table 1). Receptors' residues that were distant not more than 6 Å from the ligand or peptide were extracted using an in-house Python script. We defined a contact between an atom of the ligand and one from the receptor as formed, if their distance is smaller than their van der Waals radii plus a cutoff of 0.8 Å as in [33,34]. The latter takes into account the difference between the different X-Ray resolutions [33,34]. Extracted interactions were then visually checked. We generalized the numbering of the residues involved in ligand or peptide interaction using the GPCRdb numbering [16]. The presence of each position in ligand binding was calculated using in-house Python scripts. For the conservation analysis, we downloaded the rho-GPCRs curated eukaryotic alignment from the GPCRdb database (1618 sequences). We used Python scripts to extract the conservation percentage of residues for each position and build a LOGO of the alignment using the WebLogo server [35]. Coevolution and mutual information (MI) were calculated using the MISTIC server [36]. The mutual information theory is often used to estimate the coevolutionary relationship between two positions in a multiple sequence alignment. For example, crucial mutations that change the protein function can take place only if a compensatory mutation occur elsewhere in the protein to preserve its function [37,38].

Regarding the clustering step, we used an agglomerative hierarchical cluster method from Python Scipy library, generating 10 steps for the bottom up (agglomerative) approach as the number of the shared positions in the binding cavities [39]. Each step results in a defined number of clusters, accordingly to the number of receptors that share similar residues in the same positions of the binding cavity. For example, for step number one of the agglomerative clustering, we have N number of clusters with receptors sharing similar residues in all the ten positions clustered together.

In order to cross-check our results, we downloaded the endogenous ligand structures from the ZINC database [40] and the non-endogenous ligands from the GLASS database [41] (see Figure 2). The latter are ligands previously proven with experimental data to bind a specific GPCR. We then

calculated their charges by using the MarvinSketch tool v.17.2 of the Chemaxon package (http://www.chemaxon.com). All the scripts used in this manuscript are available upon request.

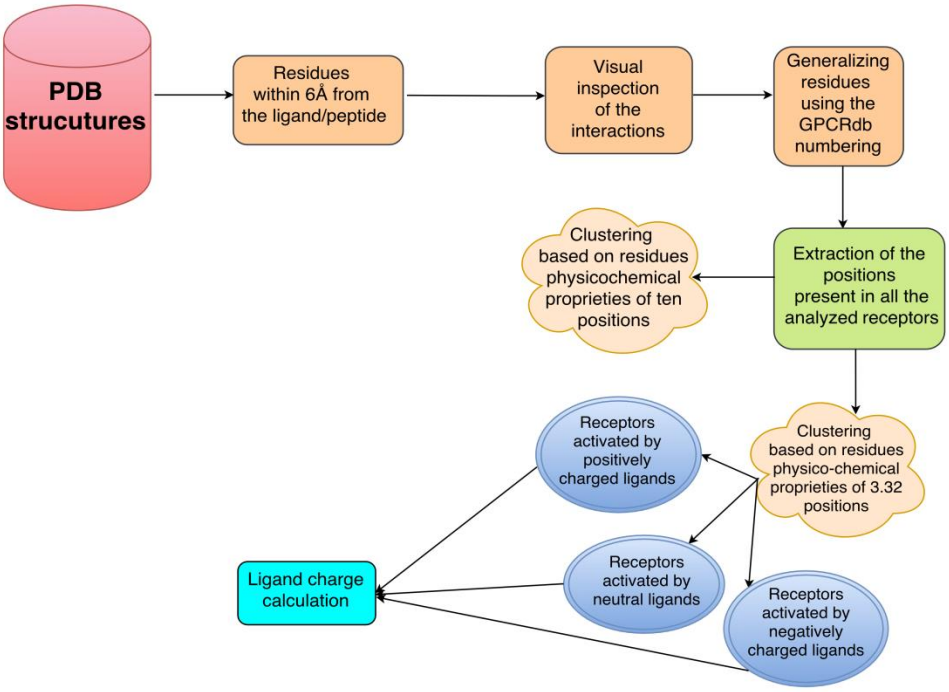

Figure 2. Flowchart of the pipeline used in this work. The pool of the analyzed PDB structures is indicated with a red cylinder, the clustering steps with two orange cloud symbols and the validations of the results steps with a cyan rectangle.

3. Results

We used a pool of 85 complexes with known structure (Supplementary Table 1), belonging to the rho-GPCRs class to perform our analysis. First, we manually checked and deleted all the apo-complexes and redundant receptor-ligand complexes. Then we calculated, for each receptor, all the residues involved in ligand binding. Finally, we manually checked the type of interactions established between the ligand and the receptor, i.e. hydrogen bonds, salt bridges, π-stacking or covalent interactions. In order to refer to a unique system of coordinates, we used the generic numbering from the GPCRdb database [16]. We extracted the binding positions by selecting only those shared in ligand binding between all the solved receptors. By using this approach, we were able to distinguish ten positions, namely 3.32, 3.33, 3.36, 6.48, 6.51, 6.52, 6.55, 7.35, 7.39, 7.43, that turned out to be functionally (positions involved in binding in all the solved GPCRs structures) conserved in 100% of the complexes, meaning that they are key positions for binding and function in all the structurally analyzed GPCRs (see Supplementary Material section "Agonist-antagonist differential bioinformatic analysis" for a detailed analysis regarding agonists and antagonists interactions with the solved receptors). These positions include residues present in only three helixes, i.e. TM3, TM6, and TM7, in agreement also with a previous study [42]. Afterwards, we used the GPCRdb mutant browser tool to check for the existence of experimental data involving residues in

these ten positions. Indeed, significant site directed mutagenesis data, involving all the ten positions, were shown to reduce ligand binding/potency of about >5-fold further confirming the functional importance of these positions (Figure 3) [16].

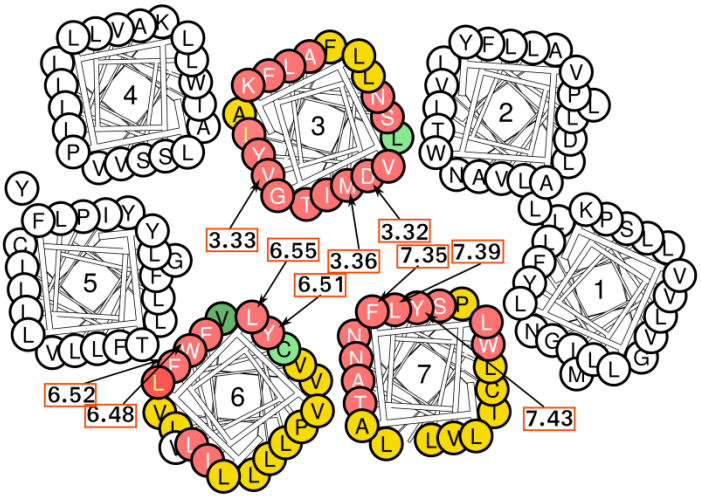

Figure 3. Schematic representation of GPCRs experimental mutagenesis data, obtained by using the GPCRdb server [16]. Red, green, and yellow positions that reduce the ligand binding/potency by >5-fold, increase the ligand binding/potency by >5-fold and have no or low effect in the binding affinity (<5-fold), respectively. Helixes are numbered from one to seven.

The observation that these ten positions are present in all the rho-GPCRs binding cavities, without distinction of the sub-family, prompted us to suggest a role in the evolutionary history of the receptors. We thus performed a coevolution and a mutual information study on a rho-GPCRs curated alignment of 1618 eukaryotic sequences. We observed high values of cumulative mutual information (cMI) for the ten previously calculated positions, meaning that these positions could have played an important evolutionary role (Figure 4 and Supplementary Figure 1) [43].

We then calculated the amino acid conservation of the ten shared positions. From this analysis emerged a high value of conservation, of about 77%, for a tryptophan in position 6.48 (see Table 1 and Supplementary Figure 2). This position is well known in literature since it is a hub involved either in ligand detection and receptor activation [22,44]. The other positions, except for 3.32, have hydrophobic amino acids, i.e. valine, methionine, leucine (Table 1) as the most conserved amino acids. Position 3.32 drawn our attention because it presents an aspartic acid in 22% of the rho-GPCRs (Table 1 and Supplementary Figure 2). This aspartic acid is known in the literature to be responsible for interacting with amines, small positively charged molecules [45]. We thus investigated if, using residues present in the 3.32 position as features for the clustering, could lead to a discrimination between receptors with similar physicochemical properties. Indeed, this first clustering-step showed three principal groups: (i) receptors with a negatively charged residue (amine cluster), (ii) receptors with a hydrophobic residue (hydrophobic cluster) and (iii) receptors with an aromatic residue (aromatic cluster) (see Figure 5). Then, in order to verify if these results could be correlated with the different type of ligands that activate these receptors, we cross-checked the clustering-step by grouping the endogenous ligands accordingly to their charges. We

calculated the charges of 35 endogenous agonists (see Supplementary Table 2). The ligands were clustered into three main groups as well: (i) positively charged ligands that bind amine receptors as acetylcholine, histamine, adrenaline etc., that correspond to the receptors' amine cluster, (ii) neutral molecules as the case of retinol, that binds opsin receptors or adenosine that binds adenosine receptors, as well as melatonin that bind melatonin receptors and anandamide that binds cannabinoid receptors, all grouped in the hydrophobic cluster and (iii) negatively charged molecules as those that activate lipid or nucleotide receptors, present in the aromatic cluster. This cross-check confirmed that physicochemical characteristics of the residues in position 3.32 were enough to discriminate receptors based on the charge of their endogenous agonists. The same behavior was observed also in the case of peptide receptors that have a negatively charged residue in position 3.32 (Supplementary Figure 3). Indeed, melanin-concentrating hormone, opioids, neuropeptide W/B, somatostatin and urotensin receptors came out to be activated mostly by positively non-endogenous ligands (Supplementary Table 3). Taken together these results indicate a huge importance of 3.32 position in ligand charge detection during evolution and an indirect involvement of this position in receptor activation.

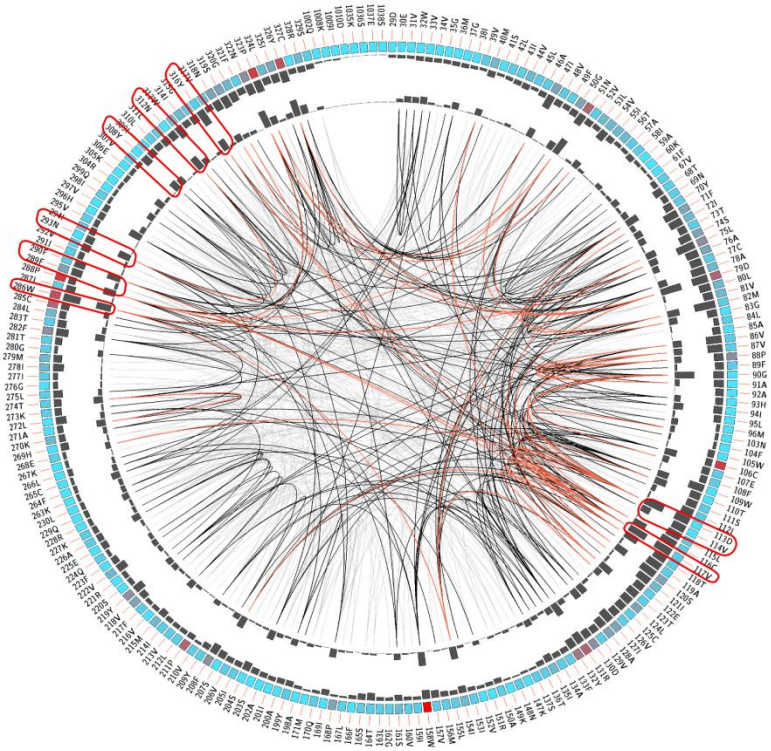

Figure 4. Mutual information circle calculated using the MISTIC server [36], mapped on the β2 adrenoceptor. Labels in the first (outer) circle indicate the alignment position and the amino acid code of the reference sequence, the β2-adrenergic receptor. The colored-square boxes of the second circle indicate the MSA position conservation (highly conserved positions are in red, while less conserved ones are in blue). The third and fourth circles show the Proximity MI (pMI) and Cumulative MI (cMI) as histograms, facing inwards and outwards respectively. In the center of the circle, red lines connect pairs of positions with high values of MI [43].

Table 1. Conservation of the residues for each shared position in the GPCRs binding cavities. Positions (green) correspond to each column and amino acids (blue) correspond to each row. The most conserved residue in shown in red.

	3.32	3.33	3.36	6.48	6.51	6.52	6.55	7.35	7.39	7.43
G	0,0364	0,0444	0,1242	0,0142	0,0067	0,0265	0,0352	0,0191	0,0154	0,0030
A	0,0822	0,0309	0,0160	0,0049	0,0037	0,0859	0,1365	0,0315	0,0822	0,0191
V	0,0618	**0,1699**	0,0976	0,0006	0,0234	0,0253	0,0624	0,0302	0,1093	0,0166
L	0,0556	0,0605	0,1285	0,0018	0,1019	0,0426	**0,1903**	0,1186	**0,1779**	0,0296
I	0,0723	0,0568	0,0265	0,0024	0,0519	0,0185	0,0475	0,0469	0,0920	0,0636
M	0,0550	0,0111	**0,1959**	0,0191	0,0049	0,0074	0,0296	**0,1205**	0,0160	0,0939
F	0,0927	0,0815	0,0896	0,0945	**0,3207**	**0,2379**	0,0358	0,2002	0,0519	0,1279
W	0,0012	0,0024	0	**0,7750**	0,0012	0,0080	0,0179	0,0296	0,0037	0,0222
P	0,0154	0,0037	0,0055	0	0	0	0,0006	0,0006	0,0024	0,0333
S	0,0160	0,0655	0,0920	0,0185	0,0111	0,0432	0,0488	0,0494	0,0525	0,0957
T	0,0642	0,1322	0,0438	0,0037	0,0055	0,0296	0,0352	0,0370	0,0747	0,0358
C	0,0049	0,0710	0,1093	0,0037	0,0055	0,0043	0,0135	0,0012	0,0117	0,0191
Y	0,0852	0,1310	0,0228	0,0222	0,3850	0,0488	0,0228	0,1149	0,0407	**0,2904**
N	0	0,0154	0,0030	0,0006	0,0142	0,1477	0,1174	0,0247	0,0574	0,0117
Q	0,0704	0,0135	0,0129	0,0129	0,0012	0,0407	0,0438	0,0049	0,0154	0
D	**0,2113**	0,0203	0	0,0006	0	0	0,0049	0,0296	0,0098	0,0018
E	0,0074	0,0241	0	0,0006	0,0098	0,0030	0,0080	0,0309	0,0957	0,0055
K	0,0247	0,0241	0,0030	0	0	0	0,0259	0,0525	0,0092	0,1013
R	0,0142	0	0,0098	0,0030	0,0043	0,0055	0,0692	0,0210	0,0358	0,0018
H	0,0086	0,0216	0	0,0006	0,0278	0,2033	0,0234	0,0309	0,0401	0,0216

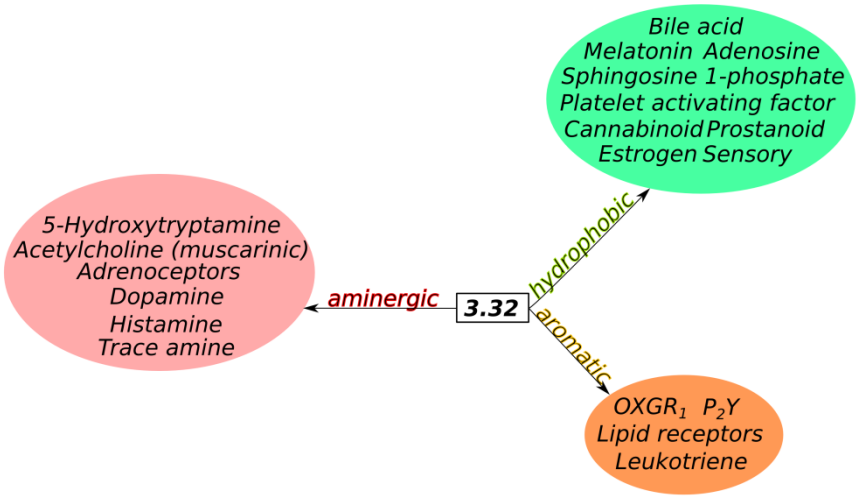

Figure 5. Clustering results using the physicochemical properties of residues corresponding to position 3.32. Amine cluster is shown in red, hydrophobic cluster is shown in green and aromatic cluster in shown in orange.

Prompted by the encouraging result of the clustering method, we then considered all the ten previously identified positions, namely 3.32, 3.33, 3.36, 6.48, 6.51, 6.52, 6.55, 7.35, 7.39, 7.43, for a second clustering-step. As the shared positions are ten, we used ten different thresholds in our approach (see Methods section for more details). At threshold equal to ten (first level of the agglomerative clustering) shared positions, receptors that have residues with similar physicochemical properties in all the positions are clustered together. On the other hand, at threshold equal to one (last level of the agglomerative clustering), receptors sharing only one similar residue are clustered together. We obtained 394 clusters for the first level and 10 clusters for the tenth level (Supplementary Figure 4). Concerning the first level, we noticed that clustering was strongly species-dependent with mammalians sequences mostly clustered together, and other more distant species, like fishes, forming separate clusters. We focused instead our attention on the seventh level of the agglomerative cluster, with a threshold of at least three shared positions. We chose this level because it showed an optimal ratio between number of members and biological relevance of the clusters (Figure 6). Thus, at this point, we were interested in investigating if, sharing only three residues in the binding cavity, could be enough to explain the evolutionary history of rho-GPCRs. We compared our clusters (Figure 6) against the GPCR network phylogenetic tree's sub-branches [47] and against data in literature containing experimental information on the GPCRs.

First, all the amine receptors (Figure 6, cluster α) were clustered within the same cluster, similarly as in the classical GPCR phylogenetic tree [47]. Indeed, all the amine receptors shared five positions out of ten in the binding site (see Figure 6). In other four clusters (clusters λ, κ, μ and ν, Figure 6), we can distinguish, Adenosine, Rhodopsin, P2Y and Chemokine receptors. While these receptors are very similar within their local branches (subfamilies), they differ throughout the rho-GPCRs class. In our clusters analysis, we observed the same trend as in the GPCRs phylogenetic tree [47]. In Cluster β we find Motilin, Neuromedin U and Ghrelin receptors that share six out of ten positions in the binding cavity (Figure 6). From a biological point of view, Motilin and Ghrelin receptors are both used as drug targets for gastrointestinal disorders [48]. Considering only their

binding cavities, our clustering method was able to give informative results regarding the common pathway of these receptors.

Cannabinoid receptors and Melanocortin receptors are included in the same cluster (cluster θ, Figure 6). These two receptors were discovered to be expressed as a chimeric protein in an isolated fragment of a leech CNS, an invertebrate species [49]. Surprisingly, in the eukaryotic species, they share only three out of ten positions in the binding site. These positions, that could have played a key role during the evolution of these two receptors, were captured with our method.

Cluster η grouped together Lipid, Prostanoid, Cholecystokinin and Relaxin receptors (Figure 6). Cholecystokinin is expressed in the gastrointestinal system and is responsible for stimulating the digestion of lipids, and thus belongs to the same biological pathway of lipid receptors. In fact, Harikumar KG and collaborators [50] have shown a high correlation between a microenvironment rich of lipids and the inactive, uncoupled state of the Cholecystokinin receptor.

On the last cluster (cluster δ, Figure 6) we can find Opioid, Endothelin and Oxytocin receptors together. Opioid and Endothelin receptors share a common antagonist, that inhibits both receptors [51]. Regarding the Oxytocin receptor, it seems that both Opioid and Oxytocin receptors play important roles in pain modulation. Indeed, the opioid system is involved in the oxytocin-induced antinociception in the brain of rats [52]. Thus, our clustering method was able to capture important biological features, that are difficult to be captured using classical phylogenetic approaches.

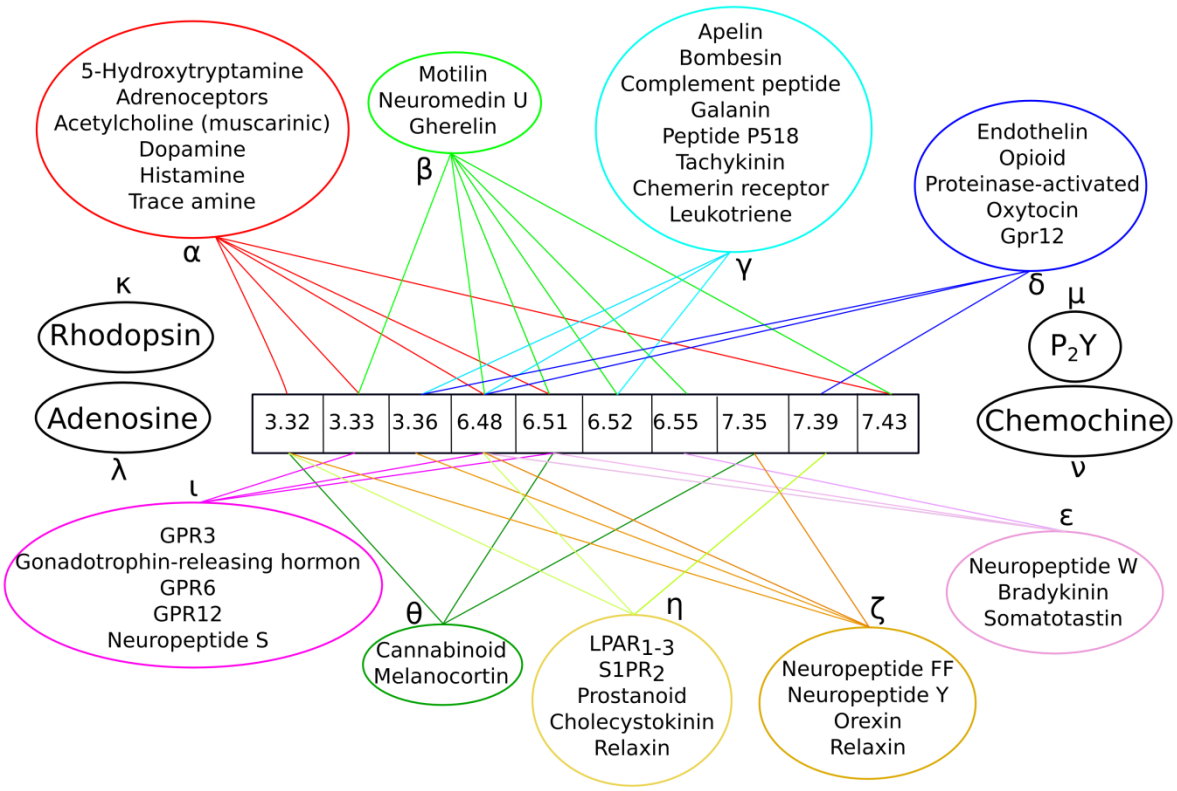

Figure 6. Clustering-step using, as features, residues corresponding to the ten shared binding cavity positions. Each cluster is shown with a differently colored circle and labeled using Greek letters. The shared positions within a cluster are also illustrated with the same colored lines.

4. Conclusions

Here we present the first global analysis on the rho-GPCRs binding cavities, based on all the solved structures of these receptors. In total, we analyzed 85 complexes and from our structural-based GPCRs analysis, we found previously uncovered properties of the binding cavities. First, we found that ten positions of the GPCRs binding cavities, namely 3.32, 3.33, 3.36, 6.48, 6.51, 6.52, 6.55, 7.35, 7.39, 7.43, are shared between all the rho-GPCRs solved structures. They are located in three helices, i.e. TM3, TM6, and TM7. These helices together with TM2 have been previously shown to play a fundamental role in GPCRs activation [5,26,53]. This leads us to believe that our findings could be strictly connected with the activation of these receptors. In fact, the transmission of the signal in GPCRs starts with agonist binding and continue, through hinge residues interactions, towards the G-protein binding cavity. Our findings could be used as the starting point of further studies aiming at a deep learning of GPCRs activation.

Using similar physicochemical properties of residues in these ten positions as features, we then were able to cluster together receptors that are very distant between each other at a sequence level, but very close in ligand recognition and binding cavities similarities. We showed that on one hand, some receptors were clustered together in a very similar way to branches of the GPCR network phylogenetic tree. On the other hand, we found clustered together receptors completely different at a sequence level but belonging to the same biological pathway. An example is the case of Opioid and Bradykinin receptors that interact with the same/very similar ligands. The method of the agglomerative clustering used here was able to capture important features of receptors binding cavities that are very difficult to be recognized using classical phylogenetic approaches. Moreover, position 3.32 that seems to have played an important role in agonists' charge detection during GPCRs evolution.

To the best of our knowledge, this is the first time that a similar analysis is performed on all the rho-GPCRs solved structures. We believe that our results can help in the deorphanization of GPCRs whose ligands are still unknown, as well as in suggesting novel specific drug targets.

Acknowledgments

E.S. was financed by a PhD fellowship from the Italian Ministry of Education, University, and Research (MIUR).

Conflict of Interest

All authors declare no conflicts of interest in this paper.

References

1. Pierce KL, Premont RT, Lefkowitz RJ (2002) Seven-transmembrane receptors. *Nat Rev Mol Cell Biol* 3: 639–650.
2. Fredriksson R, Lagerström MC, Lundin L, et al. (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* 63: 1256–1272.

3. Lagerström MC, Schiöth HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov* 7: 339–357.
4. Tehan BG, Bortolato A, Blaney, et al. (2014) Unifying family A GPCR theories of activation. *Pharmacol Ther* 143: 51–60.
5. Lans I, Dalton JA, Giraldo J (2015) Helix 3 acts as a conformational hinge in Class A GPCR activation: An analysis of interhelical interaction energies in crystal structures. *J Struct Biol* 192: 545–553.
6. Trzaskowski B, Latek D, Yuan S, et al. (2012) Action of molecular switches in GPCRs—theoretical and experimental studies. *Curr Med Chem* 19: 1090–1109.
7. Dalton JA, Lans I, Giraldo J (2015) Quantifying conformational changes in GPCRs: glimpse of a common functional mechanism. *BMC Bioinformatics* 16: 124.
8. Katritch V, Fenalti G, Abola EE, et al. (2014) Allosteric sodium in class A GPCR signaling. *Trends Biochem Sci* 39: 233–244.
9. Krishnan A, Almén MS, Fredriksson R, et al. (2012) The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. *PLoS One* 7: e29817.
10. Willard SS, Koochekpour S (2013) Glutamate, glutamate receptors, and downstream signaling pathways. *Int J Biol Sci* 9: 948–959.
11. Springer TA (1990) Adhesion receptors of the immune system. *Nature* 346: 425–434.
12. Nichols AS, Floyd DH, Bruinsma SP, et al. (2013) Frizzled receptors signal through G proteins. *Cell Signal* 25: 1468–1475.
13. Afroze S, Meng F, Jensen K, et al. (2013) The physiological roles of secretin and its receptor. *Ann Transl Med* 1: 29.
14. Costanzi S, Siegel J, Tikhonova IG, et al. (2009) Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr Pharm Des* 15: 3994–4002.
15. Palczewski K, Kumasaka T, Hori T, et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289: 739–745.
16. Isberg V, Mordalski S, Munk C, et al. (2016) GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res* 44: D356–D364.
17. Zhang H, Qiao A, Yang D, et al. (2017) Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature* 546: 259–264.
18. Lundstrom K (2006) Latest development in drug discovery on G protein-coupled receptors. *Curr Protein Pept Sci* 7: 465–470.
19. Wacker D, Stevens RC, Roth BL (2017) How ligands illuminate GPCR molecular pharmacology. *Cell* 170: 414–427.
20. Moreira IS (2014) Structural features of the G-protein/GPCR interactions. *Biochim Biophys Acta* 1840: 16–33.
21. Venkatakrishnan AJ, Deupi X, Lebon G, et al. (2013) Molecular signatures of G-protein-coupled receptors. *Nature* 494: 185–194.
22. Katritch V, Cherezov V, Stevens RC (2013) Structure-function of the G protein-coupled receptor superfamily. *Annu Rev Pharmacol Toxicol* 53: 531–556.
23. Wolf S, Grünewald S (2015) Sequence, structure and ligand binding evolution of rhodopsin-like G protein-coupled receptors: a crystal structure-based phylogenetic analysis. *PLoS One* 10: e0123533.

24. Kratochwil NA, Malherbe P, Lindemann L, et al. (2005) An automated system for the analysis of G protein-coupled receptor transmembrane binding pockets: alignment, receptor-based pharmacophores, and their application. *J Chem Inf Model* 45: 1324–1336.
25. Rovati GE, Capra V, Neubig RR (2007) The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. *Mol Pharmacol* 71: 959–964.
26. Urizar E, Claeysen S, Deup íX, et al. (2005) An activation switch in the rhodopsin family of G protein-coupled receptors: the thyrotropin receptor. *J Biol Chem* 280: 17135–17141.
27. Caltabiano G, Gonzalez A, Cordom íA, et al. (2013) The role of hydrophobic amino acids in the structure and function of the rhodopsin family of G protein-coupled receptors. *Methods Enzymol* 520: 99–115.
28. Wheatley M, Wootten D, Conner MT, et al. (2012) Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol* 165: 1688–1703.
29. Zhang D, Zhao Q, Wu B (2015) Structural studies of G protein-coupled receptors. *Mol Cells* 38: 836–842.
30. Ziarek JJ, Kleist AB, London N, et al. (2017) Structural basis for chemokine recognition by a G protein-coupled receptor and implications for receptor activation. *Sci Signal* 10: 471.
31. Stansfeld PJ, Goose JE, Caffrey M, et al. (2015) MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes. *Structure* 23: 1350–1361.
32. Rose PW, Prlić A, Altunkaya A, et al. (2017) The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res* 45: D271–D281.
33. Vangone A, Bonvin AM (2015) Contacts-based prediction of binding affinity in protein-protein complexes. *Elife* 4: e07454.
34. Venkatakrishnan AJ, Deupi X, Lebon G, et al. (2016) Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* 536: 484–487.
35. Crooks GE, Hon G, Chandonia JM, et al. (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188–1190.
36. Simonetti FL, Teppa E, Chernomoretz A, et al. (2013) MISTIC: Mutual information server to infer coevolution. *Nucleic Acids Res* 41: W8–W14.
37. Gloor GB, Martin LC, Wahl LM, et al. (2005) Mutual information in protein multiple sequence alignments reveals two classes of coevolving positions. *Biochemistry* 44: 7156–7165.
38. Martin LC, Gloor GB, Dunn SD, et al. (2005) Using information theory to search for co-evolving residues in proteins. *Bioinformatics* 21: 4116–4124.
39. Ward Jr, Joe H (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 301: 236–244.
40. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45: 177–182.
41. Chan WK, Zhang H, Yang J, et al. (2015) GLASS: a comprehensive database for experimentally validated GPCR-ligand associations. *Bioinformatics* 31: 3035–3042.
42. Lee SM, Booe JM, Pioszak AA (2015) Structural insights into ligand recognition and selectivity for classes A, B, and C GPCRs. *Eur J Pharmacol* 763: 196–205.
43. Buslje CM, Santos J, Delfino JM, et al. (2009) Correction for phylogeny, small number of observations and data redundancy improves the identification of coevolving amino acid pairs using mutual information. *Bioinformatics* 25: 1125–1131.

44. Marco E, Foucaud M, Langer I, et al. (2007) Mechanism of activation of a G protein-coupled receptor, the human cholecystokinin-2 receptor. *J Biol Chem* 282: 28779–28790.
45. Kooistra AJ, Kuhne S, Esch IJP, et al. (2013) A structural chemogenomics analysis of aminergic GPCRs: lessons for histamine receptor ligand design. *Br J Pharmacol* 170: 101–126.
46. Regard JB, Sato IT, Coughlin SR (2008) Anatomical profiling of G protein-coupled receptor expression. *Cell* 135: 561–571.
47. Stevens RC, Cherezov V, Katritch V, et al. (2013) The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. *Nat Rev Drug Discov* 12: 25–34.
48. Sanger GJ, Furness JB (2016) Ghrelin and motilin receptors as drug targets for gastrointestinal disorders. *Nat Rev Gastroenterol Hepatol* 13: 38–48.
49. Elphick MR (1998) An invertebrate G-protein coupled receptor is a chimeric cannabinoid/melanocortin receptor. *Brain Res* 780: 170–173.
50. Harikumar KG, Puri V, Singh RD, et al. (2005) Differential effects of modification of membrane cholesterol and sphingolipids on the conformation, function, and trafficking of the G protein-coupled cholecystokinin receptor. *J Biol Chem* 280: 2176–2185.
51. Bhalla S, Zhang Z, Patterson N, et al. (2010) Effect of endothelin-A receptor antagonist on mu, delta and kappa opioid receptor-mediated antinociception in mice. *Eur J Pharmacol* 635: 62–71.
52. Gao L, Yu LC (2004) Involvement of opioid receptors in the oxytocin-induced antinociception in the central nervous system of rats. *Regul Pept* 120: 53–58.
53. Deupi X, Standfuss J (2011) Structural insights into agonist-induced activation of G-protein-coupled receptors. *Curr Opin Struct Biol* 21: 541–551.

AIMS Press