Structural insight to mutation effects uncover a common allosteric site in class C GPCRs

Harpsøe, Kasper; Boesgaard, Michael W; Munk, Christian; Bräuner-Osborne, Hans; Gloriam, David E

Published in: Bioinformatics (Oxford, England)

DOI: 10.1093/bioinformatics/btw784

Publication date: 2017

Document license: CC BY-ND

Structural bioinformatics

Structural insight to mutation effects uncover a common allosteric site in class C GPCRs

Kasper Harpsøe, Michael W. Boesgaard, Christian Munk, Hans Bráuner-Osborne and David E. Gloriam*

Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark

*To whom correspondence should be addressed.

Received on October 7, 2016; revised on November 24, 2016; editorial decision on December 3, 2016; accepted on December 7, 2016

Abstract

Motivation: Class C G protein-coupled receptors (GPCRs) regulate important physiological functions and allosteric modulators binding to the transmembrane domain constitute an attractive and, due to a lack of structural insight, a virtually unexplored potential for therapeutics and the food industry. Combining pharmacological site-directed mutagenesis data with the recent class C GPCR experimental structures will provide a foundation for rational design of new therapeutics.

Results: We uncover one common site for both positive and negative modulators with different amino acid layouts that can be utilized to obtain selectivity. Additionally, we show a large potential for structure-based modulator design, especially for four orphan receptors with high similarity to the crystal structures.

Availability and Implementation: All collated mutagenesis data is available in the GPCRdb mutation browser at http://gpcrdb.org/mutations/ and can be analyzed online or downloaded in excel format.

Contact: david.gloriam@sund.ku.dk

Supplementary information: Supplementary data are available at Bioinformatics online.

1 Introduction

Class C of G protein-coupled receptors (GPCRs) consists of five receptor families (Alexander et al., 2015): eight metabotropic glutamate receptors (mGlul-8) sub-divided into three groups (Conn and Pin, 1997), two GABA receptors (GABA_B1-2) (Pinard et al., 2010), three taste 1 receptors, (TAS1R1-3) (Treesukosol et al., 2011), the calcium-sensing (CaS) and GPRC6 receptors, and seven orphan receptors, GPR156, GPR158, GPR179 and GPRC5A-D. The mGlul and GABA_B receptors regulate neuronal transmission throughout the central nervous system and have been implicated in many high-interest diseases as e.g. Parkinson’s. In addition to the well-known functions of the CaS receptor it has been implicated in e.g. Alzheimer’s disease. The taste 1 receptors are responsible for umami and sweet taste and a broad expression pattern may indicate other physiological functions. Physiological function and therapeutic utilization of orphan receptors are unknown and are being debated for the GPRC6 receptor (Clemmensen et al., 2014) (Supplementary Table S1).

Allosteric modulators (AMs) that bind in the transmembrane domain (TMD) comprised of seven helices (TM1-7) have now been discovered for all non-orphan class C GPCR families with potential therapeutic advantages by enhancing or attenuating the normal receptor response without interfering with the N-terminal domain orthosteric site (Lu et al., 2014). Furthermore, AMs potentially offer better receptor subtype selectivity and biased signalling, which in concert can lead to fewer undesirable side-effects (Nickols and Conn, 2014). This allosteric site in class C GPCRs constitute targets for therapeutics as well as for the food industry (Supplementary Table S1), exemplified by the marketed drug cinacalcet, a CaS receptor positive AM (PAM) (Nemeth and Shoback, 2013) and a Parkinson’s disease clinical trial with dipraglurant, a mGlul negative AM (NAM) by Addex Pharma (Tison et al., 2016).
The binding sites of class C GPCRs have been extensively studied using site-directed mutagenesis and receptor structure models (Gregory and Conn, 2015), but the lack of class C structures limited the resolutions of these analyses, hampering the translation of mutagenesis effects and models into pharmacological mechanisms and prospective rational structure-based drug design. Similarly to previously published work on class A GPCRs (Hulme, 2013), the recent mGlu structures (Christopher et al., 2015; Dore et al., 2014; Wu et al., 2014) has herein opened up for the first consistent comparison of all published class C GPCR mutants in the TMD. We pinpoint the location of the allosteric sites, and integrate mutant and structural data towards rational design of new modulators.

2 Methods

2.1 Annotation of literature single-point mutations
We manually annotated 1670 data points from 35 publications with single-point mutations tested on 70 AMs and made them available online in the GPCR database, GPCRdb (Isberg et al., 2016). These mutants cover 99 TMD helix and several extracellular loop positions within the CaS, GABAB2 (not shown, as all reported single-point mutations had low effect on PAM function), GPRC6a, mGlu1, mGlu10, mGlu4, mGlu9, and TAS1R3 receptors. Residue positions, numbered according to the GPCRdb generic scheme (Isberg et al., 2015), were considered as potential ligand binding if they have >5-fold effect on AM binding or function (Figs 2A, 3A).

2.2 Sequence alignment, similarities and phylogenetic trees
The sequence alignments and similarities (Supplementary Figs S1, S2 and Table S2) of the 22 human class C GPCRs was calculated and downloaded from GPCRdb (Isberg et al., 2016) The phylogenetic trees (Fig. 1) were calculated with the PHYLIP package, v. 3.695 (Felsenstein, 1989), based on the above alignments using the protdist and neighbor programs with UPGMA clustering and rendered by an in-house javascript using D3.js.

2.3 Crystal structure binding site residues
The mGlu3 (PDB codes, 4OO9, 5CGC and 5CGD) and mGlu1 (PDB code 4OR2) crystal structures were downloaded from the Protein Data Bank (www.rcsb.org) and the PyMOL Molecular Graphics System, version 1.8, Schrödinger, LLC was used to prepare Figures 2B and 3B and determine the residue positions in the allosteric binding site by selecting residues with Cz or side chain atoms ≤5 Å of any co-crystallized NAM atom.

3 Results

3.1 The class C GPCR families group consistently for the transmembrane domain and allosteric modulator site
Class C GPCRs are classified into families by endogenous ligand (Southan et al., 2016) (Supplementary Table S1), which bind in the extracellular N-terminal domain, while the reported AMs bind in the distinct TMD. Thus, it is crucial to establish if these receptor families are valid also for the analysis of AMs. We constructed phylogenetic trees based on the TMD and common allosteric site (below) sequence alignments (Figs 1, S1 and S2). In both trees, the receptor families with a known physiological ligand display a coherent grouping of all members. In addition, the orphan receptors are divided into two groups with the exception of GPR156, which, in agreement with earlier reports (Calver et al., 2003), groups with the GABAB receptors. This confirms that the established receptor families can indeed serve as a basis to map allosteric data, and as target profiles for ligand selectivity and inference.

3.2 Class C GPCR negative and positive allosteric modulators bind in one common transmembrane domain site
An overall TMD allosteric site in the class C GPCRs can be delineated from the mGlu structures and the collated mutagenesis data. In total, 28 residue positions line the TMD pocket and together represent the overall potential AM contacts (Fig. 2). Notably, these exhibit strong experimental evidence: 26 positions are in close vicinity (≤5 Å) of the co-crystallized modulators, and 23 have demonstrated effect (>5-fold) upon mutagenesis (Fig. 2A). Nine AM binding hotspots, 2x56, 3x32, 3x36, 5x43, 6x50, 6x53, 6x57, 7x37 and 7x40, span three receptor families. The latter six positions are supported by mutant effect on 10 or more different ligands while mutations in the central part of the binding site, i.e. positions 6x50 and 6x53 show effect on 37 and 24 different ligands, respectively, in six different receptors (Supplementary Table S3). Furthermore, 68% of the positions re-occur in at least two families demonstrating a
3.1 The mGlu5 receptor features a unique site extension

The mGlu5 receptor structures contain a unique site extension located between TM2, 3 and 7 (Fig. 3A, B) that can be exploited to achieve selective NAM binding (Harpsøe et al., 2015). Its uniqueness is supported by the site extension position 7x44 in four different receptors with a selective effect in mGlu5 (Supplementary Table S3 and online in GPCRdb), and a gateway position 3x40 holding a mGlu5-specific Pro, which significantly decreases or abolishes AM binding or activity when mutated (Gregory et al., 2013; Molck et al., 2012). Most other class C GPCRs contain bulkier sidechains in position 3x40 and in another crucial position, Gly2x49 (Harpsøe et al., 2015) (Supplementary Fig. S1). In spite of this, a 6.7-fold PAM affinity decrease by a CaS receptor mutation in position 7x44 (Fig. 3A) in one of two assays (Leach et al., 2016) challenges the uniqueness of the site extension. However, as a drastic F3x40A mutation in the site extension gateway position does not support this finding, it is tempting to consider it as an experimental outlier. In all, the data suggest that the extension of the modulator site in mGlu5 is unique not only within the mGlu receptors, but within all class C GPCRs.

3.2 Mapping of other class C GPCR modulators structures. Structural mapping pinpoints an allosteric site that, as for most class A GPCR ligands, is located mainly between TM3 and TM5-7, with additional contacts to TM2 and the second extracellular loop (Fig. 2B). Notably, 87% of the mutated binding site positions are involved in both NAM and PAM binding and/or function (Fig. 2A, C) and the concordant structural and mutagenesis data strongly suggests the existence of a common allosteric binding site for both types of modulators with substantial overlap across the class C receptor families.

3.3 The second extracellular loop is involved in binding

The second extracellular loop of GPCRs typically forms a disulphide bridge to the top of TM3, placing the loop as a lid on the TMD pocket (Wu et al., 2014), often in direct ligand contact (de Graaf et al., 2008). It is constituted by a pair of cysteine residues (3x29 and 45x50) only absent in GPR156, GPRCSA and GPRCSD (Supplementary Fig. S1), indicating that the second extracellular loop is close to the AM site in nearly all class C GPCRs. Mutagenesis data reports effect on the position before (45x49) and four consecutive positions after (45x51-54) the conserved cysteine (Figs 2A, B, 3A, B and online in GPCRdb), while the mGlu structures only show one position (45x52) in direct contact with a NAM and an additional position (45x49) with the side chain pointing towards the AM site. This is expected as the mGlu3 receptor structures (Wheatley et al., 2012) show the flexibility and varying length of the second extracellular loop to bring alternative positions into vicinity of the AM site.

3.4 Glycine mutations perturb TM4 and 7 helical structure

Relative to other residues, glycines allow for additional backbone torsional angles and introduce more flexibility (Bywater and Veryazov, 2015). Mutation of two Gly positions, 4x42 and 7x43, that face the cell membrane and TM6, respectively, have been shown to affect modulator binding. Specifically, a Gly-to-Val
mutation of position 4x42 yielded >55-fold effect on the mGlu3 PAM, [JN]-41482012 (Farinha et al., 2015), and a Gly-to-Ala/Val mutation in 7x43 reduced abolished activation of TAS1R3 by the allosteric agonist NHDC (Winning et al., 2007). Importantly, mutation in 4x42 has no effect when the native residue is not a Gly: Val-to-Ile and Leu-to-Val in mGlu3 and mGlu1, respectively, have no effect on several AMs (online in GPCRdb). Together, this indicates that these subtype specific Gly residues are associated with distinct receptor structures, such as a helical kink or rotation implicating additional class C GPCRs with a Gly at 4x42 (GPR179) or 7x43 (CaS and GPRC6 receptors) (Supplementary Fig. S1). It would be of interest to mutate Gly4x42 in the CaS and GPRC6 receptors, whereas mutations of Gly7x43 can currently not be tested on the orphan GPR179 due to lack of ligands.

3.8 Mutated positions with weak and rare effects

The remaining mutated positions in Figure 3, 1x42, 3x41, 4x50, 4x51, 5x37, 7x38 and 7x47, are distant to the common allosteric site and have relatively subtle effects reported for only one AM on one receptor (Supplementary Table S3). Of note, the AMs affected by mutations in these positions are all more pronouncedly affected by mutation in other positions within the common allosteric site. Taken together, it is fair to conclude that these mutants probably give their effect through other processes, such as e.g. binding site entry/exit or stabilization of an (in)active receptor state, or represent experimental outliers.

4 Discussion

The available structural templates, which are all NAM-complexes, map as well to the PAM as the NAM mutagenesis data (Fig. 2). This is in line with the many class A GPCR-agonist and -antagonist structure complexes which exhibit moderate differences in the TMD ligand binding site (Kruze et al., 2013). This suggests that the mGlu4 and mGlu5 structures could be valid templates also for the design of PAMs. However, to avoid bias toward NAM activity, it is advisable to first optimize the structural templates around a high-affinity PAM, which has previously been successful in GPCR virtual screening (Vilar et al., 2011). Furthermore, AMs have increasingly been associated with biased agonism, including CaS (Cook et al., 2015; Leach et al., 2016) and mGlu5 (Sengmany and Gregory, 2016) receptors, and it could be expected that several already existing and future AMs share this phenomenon.

An advantage of allosteric ligands is their ability to achieve higher receptor target selectivity. As shown in Figure 2, the allosteric site is largely the same across the class C receptor families, which could imply less selective binding. However, sequence alignment of the 28 positions that constitute the common allosteric site shows that, in spite of many positions displaying more than 60% conservation of side chain properties (e.g. 86% of the receptors contain an aromatic residue in position 6x53—Supplementary Fig. S2), there are a number of family- and subtype-unique residues. Thus, as shown for the mGlu receptors (Harpsoe et al., 2015), selective modulators may be achieved by exploiting these as selectivity hotspots. This is also supported by the known chemistry, as prototypical AMs (Supplementary Fig. S3) display highly diverse structures.

Modelling of class C GPCR-AM complexes has become more accessible with the possibility to utilize the mGlu structures in combination with our collation of literature mutations, which are available for visualization and download in GPCRdb. We previously demonstrated that such data was sufficient to build conclusive models explaining mGlu NAM selectivity (Harpsoe et al., 2015). Interestingly, the receptors that displays the closest homology within the allosteric site (50–64%) are the orphan GPRC5A-D receptors (Supplementary Table S2). This is encouraging for the prospects of identifying ligands through receptor structure-based virtual screening. Notably, since the orphan receptors lack the extracellular domain (Kniazeff et al., 2011), it is plausible that the TMD site constitutes their orthosteric site in line with the fact that other class C GPCRs can be activated via the allosteric site, even without the N-terminal domain (Binet et al., 2004; Goudet et al., 2004). Identification of the endogenous ligands or even the first tool compounds would be very valuable for the characterization of physiological functions, cellular localizations and therapeutic potential.

In conclusion, the accumulated mutagenesis data for AMs on class C GPCRs almost exclusively maps to one common inter-helical allosteric site for both PAMs and NAMs, which corresponds very well with the allosteric site in the mGlu crystal structures. The combined mGlu-NAM structure complexes, mutagenesis data, sequence comparisons and known NAMs/PAMs lays a stronger foundation for rational design of novel modulators through the construction of high quality homology models. Specifically, structures, mutagenesis data and sequence comparisons have been combined in GPCRdb to create an online tool to design new mutagenesis experiments (Munk et al., 2016). This design aids researchers by selecting the mutations with the highest observed involvement in modulator binding, while avoiding indirect effects through adjacent residues or due to perturbation of the backbone structure. Given the disease relevance of several class C GPCRs, it will be intriguing to explore the integration of data for the identification of potent target- and maybe pathway-specific novel modulators towards more efficient drugs with less adverse effects.

Funding

This work was supported by the Lundbeck Foundation [R163-2013-16327 to D.E.G., C.M. and K.H.]; the European Research Council [DE-ORPHAN 639125 to D.E.G., C.M. and K.H.]; the Augustinus Foundation [15-2818 to M.W.B and H.B.-O.]; and the COST Action CM1207 ‘LISTEN’ [to D.E.G.].

Conflict of Interest: none declared.

References


Binet, V. et al. (2004) The heptahelical domain of GABA(B2) is activated directly by CGP7930, a positive allosteric modulator of the GABA(B) receptor. J. Biol. Chem., 279, 29085–29091.


Gregory, K.J. et al. (2013) Probing the metabotropic glutamate receptor 5 (mGlu5) positive allosteric modulator (PAM) binding pocket: discovery of point mutations that engender a “molecular switch” in PAM pharmacology. Mol. Pharmacol., 83, 991–1006.


