Gut hormone GPCRs: structure, function, drug discovery

Arnau Cordomi¹, Daniel Fourmy², Irina Tikhonova³

¹ Laboratori de Medicina Computacional, Unitat de Bioestadística, Facultat de Medicina, Universitat Autònoma de Barcelona, Barcelona, Spain.
² Laboratory of Chemistry and Physics of Nanoobjects, ERL1226 INSERM, University of Toulouse, CNRS, INSERM, Toulouse, France.
³ Molecular Modelling Laboratory, School of Pharmacy, Queen’s University Belfast, Northern Ireland.

Address for correspondence: Daniel Fourmy, Laboratory of Chemistry and Physics of Nanoobjects, ERL1226 INSERM, University of Toulouse, CNRS, INSERM, Toulouse, France.
Email: Daniel.Fourmy@inserm.fr; phone: 33(0)5 61 32 30 57

Abstract:
Crystallization and determination of the high resolution three-dimensional structure of β2-adrenergic receptor in 2007 was followed by structure elucidation of a number of other receptors, including those for neurotensin and glucagon. These major advances foster the understanding of structure-activity relationship of these receptors and structure-based rational design of new ligands having more predictable activity. At present, structure determination of gut hormone receptors in complex with their ligands (natural, synthetic) and interacting signalling proteins (G-proteins, arrestins...) represents the forthcoming challenge which promise to revolution gut hormone endocrinology.

Introduction

Gut hormones represent an important class of peptides regulating many functions related to (but not limited to) absorption of nutrients and energy homeostasis. Gut hormones exert their biological functions by interacting with specific G-protein coupled receptors within tissues in the gut and outside the gut, including the central nervous system. Well known gut hormones are secretin, gastrin, cholecystokinin, glucose-dependent insulinotropic polypeptide, glucagon, glucagon-like peptide-1, vaso-intestinal polypeptide, motilin, somatostatin, galanin, ghrelin, etc...

Following purification and cloning of the cDNA coding for the β2-adrenergic receptor in 1986 [1] which was preceded by the prediction of the secondary structure of rhodospin three
years earlier [2], a large number of G-protein coupled receptors cDNAs, including those encoding receptors of gut hormones, were cloned in the 90’ leading to access to their structure. From their amino acid sequence, it rapidly appeared that all GPCRs share a similar 7-transmembrane domain (7-TMD) structure. Furthermore, this enabled their classification in several sub-families or sub-groups. Thus, gut hormone receptors belong to either family A or rhodopsin-like receptors (receptors for gastrin, cholecystokinin, motilin, somatostatin, galanin, ghrelin, ...) and to family B (sometimes named B1) or secretin-like receptors (receptors for secretin, glucagon, glucagon-like peptide-1, vaso-intestinal polypeptide ...). Crystallization and determination of the high resolution 3-D structure of β2-adrenergic receptor in 2007 [3] opened the way for structure determination of many other GPCRs and for the understanding of their structure-activity relationship. Moreover, it permitted structure-based rational design of new ligands having more predictable activity on these receptors.

In this short article, we attempt to illustrate how, since the first evidence that signal of hormones was transmitted to cells via GPCRs, a large number of approaches, culminating with determination of 3-D structures, have been employed to understand structure-activity relationship in this family of mediators.

The pre-crystallisation age of gut hormone GPCRs

Structure-activity relationship studies of gut hormones

Historically, identification and structure determination of gut hormones was followed by elucidation of their physiological functions and mechanisms of action, as well as by an explosion of structure-activity relationship studies conducted in common by medicinal chemists, pharmacologists and biologists. At that time, it was expected discovery of molecules intended to treat human diseases. Thus, individual synthetic substitution of all amino acids in the original hormone sequences led to identify the pharmacophore of the hormones, namely the region(s) which contain(s) binding and bioactive domain(s). It was found in many cases that shortened peptides reproduced the full spectrum of biological activity of the natural hormone. A second important achievement was the identification of peptidic antagonists actually representing the first generation of antagonists of gut hormone GPCRs. These advances greatly facilitated subsequent works of gut endocrinologists. However, it was rapidly appreciated that peptidic agonists and antagonists presented important disadvantages (low bioavailability and half-life) in a therapeutic perspective; this motivated medicinal chemists to launch important programmes to design non-peptidic ligands. Different strategies were used, including rational design based on structure-activity relationships of the endogenous hormones and active/inactive analogs or random screening of libraries of compounds followed by hit optimization. These works highlighted the variable degree of difficulty (and as a consequence
of success) to obtain small-molecule ligands. For instance, obtaining non-peptidic ligands for GPCRs of rhodopsin-like family which present a compact binding domain site, was much more easy than for secretin-like family which have their binding and activation sites scattered in different regions of the receptors (see below). In some cases, such as for cholecystokinin and gastrin, a huge number of non-peptidic ligands having a spectrum of activity ranging from full antagonism to full agonism were generated [4]. Clinical trials have been conducted https://clinicaltrials.gov/, but it is however disappointing to notice that none of these numerous molecules has been marketed so far.

Structure-activity relationship of gut hormone GPCRs using site-directed mutagenesis and computational 3-D modelling

Pioneer structure-activity relationship studies using site-directed mutagenesis of GPCRs, most often conducted in synergy with computational approaches of homology modelling and ligand docking, provided important data regarding the location of the binding site(s) of natural and synthetic ligands in the receptors, as well as activation mechanism and G-protein coupling determinants. For example, functional regions of the neurotensin receptor-1 (NTR1) were identified [5]. As another example, binding modes for endogenous and non-peptidic ligands of the cholecystokinin receptors CCK1R and CCK2R were studied in great detail [6]. Based on the findings, a mechanism explaining how two very similar non-peptide ligands of CCK2R differing only by the absence or presence of a methyl group display inverse agonism or partial agonism activity, respectively [7]. On the other hand, a biased antagonist of the CCK2R was identified and the underlying mechanism delineated [8]. Obviously, these data require confirmation by 3-D structure determination. Concerning class B GPCRs, a huge amount of data was generated by site-directed mutagenesis highlighting the conservation of ligand binding and activation mode in this group.

Crystal Structures of the Neurotensin and the Glucagon receptors

Seven similar crystal structures of the rat NTR1 in complex with the carboxy-terminal portion of neurotensin are available [9,10]. These are particularly interesting because they are some of the few GPCR structures solved with bound peptides. Binding site of neurotensin is located at the extracellular half of the receptor at a crevice formed mainly by residues at TM helices II, VI and VII and at the extracellular loop 3 (Figure 1 left). This binding site seen in the crystal structure of the liganded NTR1 is in excellent agreement with that deduced in studies using site-directed mutagenesis and molecular modeling. Most of the residues that constitute the binding pocket are not conserved in other peptide receptors and therefore, as evidenced
from the other known GPCR structures, the binding mode is not conserved. The 3-D structure of the neurotensins/NTR1 complex revealed an opened and solvent accessible binding pocket which is common to all the structures of class A peptide receptors and is facilitated by the formation of a beta-sheet in ECL2 [11]. The proximal N-terminal region of NTR1 (residues 50-60) partially occludes the binding pocket as has also been observed in the structures of the CXCR4 [12,13]. Such occlusion does not occur for instance in the opioid receptors [14]. NTS occupies a less deep location relative to amine receptor ligands (such as the β2-adrenergic receptor ligand carazolol, see Figure 1 left).

The structures of both the ECD and the TMD of the glucagon receptor (GLR) have been released during the last years [15-17]. The ECD is a highly conserved globular domain (100-160 residues) that consists of two antiparallel β-sheets defined by three conserved disulphide bonds and an N-terminal α-helix. The structure of the TMD is similar to that of the CRF1R [18], the other crystallized class B receptor, which reveal remarkable similarity with rhodopsin-like GPCRs (Figure 1 right) though with a larger chalice-like pocket between the extracellular ends of the helices [19]. Unfortunately, neither the ECD nor the TMD of the GLR have been solved in complex with glucagon or any peptidic analogue. However, the GLR(ECD)/glucagon complex will surely resemble the related complexes of the closely related GLP1R [20] and GiPR [21]. No structure has been obtained containing both the ECD and the TMD, thus, determining their relative orientation still represents a challenge [16,22,23]. However, Siu et al constructed a model of the GLR/glucagon including both the ECD and the TMD combining on the basis of crystal structures and mutagenesis studies [16] (Figure 1F). The C-terminus of glucagon interacts with the ECD of the GLR, the central part of glucagon interacts with the extracellular extension of TM1 that links the ECD and the TMD domain and the N-terminus (residues 1–6 of glucagon) interacts with the TMD (at a similar depth as most class A ligands).

The first step of receptor activation is believed to be the interaction of the C-terminus of peptides with the ECD of receptors followed by the binding of the N-terminal region of ligand within TMD. This second step is predicted to initiate structural changes in the TMDs (similar to class A) that activate G proteins [24,25]. A conserved positively charged residue (Lys 187 in GLR) is crucial for the binding of peptide ligands at the TMD, as it is believed to interact with the third residue (Gln3 of glucagon, Asp/Glu3 in most secretin-like hormones) of the endogenous peptides [22,26,27]. Other important residues for the binding, not listed here, are also characterized [16,28,29]. There is no available crystal structure of family B GPCRs in an active form yet and therefore, the precise structural changes that occur during activation and interaction with G proteins remain unknown. The intracellular resemblance of the overall folds of class A and B (Figure 1) suggests similar structural rearrangement upon activation, including movement of TM 6, to allow G protein interaction with the intracellular parts of TMDs. Such
movement can be modelled [30] using the increasing repertoire of active structures of class A GPCRs.

Structure-based design of small-molecule modulators of the GLR is now possible after the insight provided by the very recent crystal structure of GLR in complex with the negative allosteric modulator MK-0893 [17], but also by the crystal structure of CRF1R bound to the negative allosteric modulator CP-376395 [18]. The (distinct) location of both allosteric binding pockets (Figure 1 right) has been a big surprise. MK-0893 binds outside the TMD bundle in a position between the intracellular parts of TMs 6 and 7 extending into the lipid bilayer. By contrast, the CRF1R ligand CP-376395 binds to a small cavity generated between TMs 3, 5 and 6 at the intracellular half of the TMD bundle, but in a much more deeper position than any known orthosteric binding pocket of class A GPCRs. In both cases, the allosteric ligands block the helical movements that are required for the activation. Clearly these complexes represent outstanding possibilities for the design of novel small-molecule allosteric ligands targeting the GLR.

Structure-based design of new GPCR ligands

Precise description of key receptor-ligand interactions enables virtual screening of chemical libraries against the crystal structures and homology models using a molecular docking method. Although virtual screens have not yet been described for gut hormone rhodospin-like GPCRs, such efforts have been recently undertaken for several other GPCRs. Thus, the crystal structures of the β2 adrenergic [31], dopamine D3 [32], chemokine CXCR4 [33], serotonin 5-HT1B, 5-HT2B [34] receptors were explored in docking of different chemical libraries. Homology models were also successfully used in virtual screens for a number of GPCRs, among them, the chemokine CXCR3 [33], adenosine A1, A3 [35] and the free fatty acid FFA1 [36] receptors. Overall, results confirm that the use of a GPCR structure can provide ligands having new chemotypes, high affinity, diverse pharmacological profiles (agonist/antagonist), improved subtype selectivity and tailored selectivity patterns (dual modulators, as an example). Furthermore, virtual screening based on docking to homology models has been also probed for poorly-characterized GPCRs such as GPR17 [37], GPR68 and GPR65 [38] that led to identification of first synthetic small molecule ligands, demonstrating possibilities of finding novel small molecule binders without a need of knowing natural ligands in advance. This is particular promising for gut hormone GPCRs for which only few synthetic ligands are available.

In 2011 De Graaf and colleagues [39] made a first attempt to explore the structure-based ligand design for gut hormone GPCRs belonging to Class B. In particular, they conducted structure-based design of allosteric modulators of the glucagon-like peptide 1 (GLP-
1) and glucagon (GLR) receptors using knowledge-based homology models of GLP-1 and GLR receptors constructed on the basis of a validated corticotropin-releasing factor type 1 receptor (CRFR1) homology model. Because much structure-activity relationship of CRF (SAR) and site-directed mutagenesis data were available, this receptor was used to build a comprehensive 3D model of a Class B GPCRs, which, in turn, served as a template to build homology models for poorly-characterized GPCRs. This virtual screening campaign resulted in two GLR novel ligands and one GLP-1 ligand, demonstrating the potential of structure-based approaches to find novel ligands. At present, more solid ground for structure-based ligand design targeting gut hormone GPCRs of Class B is available.

It is appreciated that 3-D structures of GPCRs provide information for drug design that cannot be gained from ligand-based approaches. This involves various ligand binding modes, multiple binding sites, allosteric sites as well as impact of water molecules [40]. Multiple receptor conformations obtained from crystallography, NMR and molecular dynamics simulations could further assist structure-based drug design campaigns by giving possibility to predict ligand activity (biased and unbiased agonists and antagonists). There is a clear trend in GPCR drug discovery in moving from traditional high-throughput screening and ligand-based approaches towards efficient structure-based virtual screening campaigns.

Conclusion

The recent release of GPCR structures including gut hormone GPCR members of classes A and B confirmed some general features predicted from a rhodopsin structure. It also establishes unexpected ligand-binding modes and critical aspects of the receptor activation process. These major advances foster structure-based rational design of new ligands having more predictable activity. At present, structure determination of gut hormone receptors in complex with their ligands (natural, synthetic) and interacting signalling proteins (G-proteins, arrestins...) represents the forthcoming challenge which promises to revolution rational drug design and gut hormone endocrinology. New generations of pharmacological tools such as biased, dual or allosteric ligands are expected.

References


*The authors show that docking of a library of chemical fragments to homology models of receptor subtypes, validation of selected hits experimentally and subsequent fragment-to-lead optimization provide selective ligands even in the absence of crystal structures of receptors.


**Using GPR68 and GPR65, the authors demonstrate that combination of physical and structure-based screening can be used for deorphanization of orphan receptors.**


*A comprehensive review of unappreciated ligand binding features that GPCR structures provide for drug discovery.

Commented [IT1]: In capitals. Should be changed.
Figure 1. Cartoon representation of the structures of the neurotensin 1 (left; PDB id: 4XES) and glucagon (right; PDB ids: 4L6R and 4LF3 for the TMD and ECD, respectively) receptors. Displayed with ball-and-sticks and a van der Waals surface are the endogenous peptides neurotensin (fragment 8-13) and glucagon, and the negative allosteric modulator MK-0893. The β2-adrenergic receptor partial inverse agonist carazolol and the CRF1R negative allosteric modulator CP-376395 are also shown as reference. Horizontal black lines represent the membrane. Helices of the TMD are displayed as cylinders (TM 1: cyan, TM 2: yellow, TM 3: red, TM 4: gray, TM 5: green, TM 6: blue, TM 7: pale-red, helix 8: pale yellow), are shown in purple. The binding mode of glucagon and the relative orientation of TMD and ECD were adapted from previous works [16,22,23].