Dissecting the role of Gpr158 in the brain

Çetereisi, D.

2020

document version
Publisher's PDF, also known as Version of record

CHAPTER 1

Introduction

G protein-coupled receptors (GPCRs) belong to the largest cell membrane receptor family, characterized by a seven transmembrane domain (7TM) structure. GPCRs can be localized in different cell compartments, including endoplasmic reticulum, mitochondria, nucleus and cell membrane[1,2]. Over 800 GPCR genes have been identified from the human genome and they have been divided into five different families based on phylogenetic criteria: Glutamate, Rhodopsin, Adhesion, Frizzled and Secretin families[3,4]. GPCRs are involved in the regulation of diverse functions, e.g., metabolism, vision and the immune system, to name a few[5,6].

GPCRs are important "druggable" therapeutic targets due to their broad range of expression and modulatory roles in physiological functions, together with the accessibility at the cell membrane[7,8]. GPCRs can bind to and be activated by a variety of ligands including peptides, lipids, ions, hormones and neurotransmitters[9]. During G-protein dependent activation, the receptor undergoes ligand-induced conformational changes that are followed by downstream signaling activation through dissociation of G-protein subunits[10]. Different signaling pathways can be induced by specific ligand-receptor pairing events based on the type of Gα subunit ($G\alpha_s$, $G\alpha_{i/o}$, $G\alpha_{q/11}$, $G\alpha_{12/13}$) of heterotrimeric G proteins[11]. For instance, the activation of $G\alpha_s$-coupled GPCRs regulates the level of cAMP, whereas the activation of $G\alpha_{q/11}$-coupled GPCRs induces an increase of intracellular Ca^{+2} [11]. Although many of the ligand(s) and/or signaling mechanism of GPCRs are known, there are GPCRs without known endogenous ligands that are referred to as "orphan" GPCRs[12]. As the endogenous ligand is sometimes hard to pinpoint, or may even not exist[13], the therapeutic potential of orphan GPCRs has been increased by studying their function in biological systems, e.g. by overexpression or knock-down/out, independent of the endogenous ligand[6].

In the subsequent paragraphs, I will first introduce the receptor that forms the focus of my thesis, namely Gpr158. Subsequently, I will shortly introduce topics of importance for studying the physiological role of Gpr158 in absence of a verified ligand in terms of 1) neuronal development and structural plasticity, 2) synaptic transmission and plasticity, and 3) learning and memory. Lastly, I will focus on current strategies of the deorphanization process, i.e. the search for the endogenous receptor ligand.

G protein-coupled receptor 158 (Gpr158)

Gpr158 belongs to the glutamate GPCR family[14]. It contains an epidermal growth factor (EGF)-like domain and a leucine repeat region in the N-terminal which are structural features that belong to adhesion GPCR (aGPCR) family[15,16]. Yet, on the other hand it lacks the extracellular Venus flytrap module (VFTM), which is characteristic of the glutamate GPCR family. Instead, it has 11 cysteine residues near the extracellular domain and other cysteine residues in the extracellular loop 1 and 2 that are most likely involved in ligand recognition[17]. Although formally Gpr158 is known as an orphan receptor due to lack of activity demonstrated with a potency

consistent with a physiological function[6], recent studies have begun to shed light on possible ligand(s) of Gpr158 as well as the activation mechanism of specific G protein-related signaling pathways[18,19]. *Gpr158* is expressed in different organs, such as spleen, liver and lung[15]. Notably, its expression is highly enriched in the brain, particularly in striatum, hippocampus and prefrontal regions, where it mediates multiple functions[15,18,20]. In addition to its role in learning and memory[18,21], recent studies have shown that Gpr158 has a role in depression-like behavior[20,22], neuronal development[21] and synapse formation[23].

Functional characterization of GPCRs in the absence of a verified ligand

1. Neuronal development and structural plasticity

Appropriate neurodevelopment is essential for neuronal communication and signal transmission. Neurodevelopment includes essential processes such as, production of neuronal progenitors, neuronal migration, neuronal differentiation, axon guidance and synapse formation[24,25]. After migration of neurons to their target region they develop neurites. One of the neurites is specified as an axon and extends towards the target cell (axonal guidance) and forms synapses, whereas the rest of the neurites are specified as dendrites[26].

Neuronal development processes can be studied in *in vitro* conditions. For example, dendritic growth and branching have been studied in cultured neurons and this revealed the contribution of multiple factors, receptors and molecules in this process[27]. Synapse formation is another important process needed for transferring information between neurons. This process has also been studied *in vitro*, such as in artificial synapse formation assays, in which the effect of the receptor of interest on synapse formation could be investigated by measuring presynaptic protein accumulation on the axons[28].

In neuronal communication, signals between neurons are mediated by the cell soma, dendrites, axon and synapse[29]. There are multiple parameters that can affect this information flow, such as dendritic branching and dendritic spine formation. Alterations in dendritic architecture have been correlated with impaired learning and memory[21,30] and are implicated in many neuropsychiatric, neurological and neurodevelopmental disorders with reduced cognitive capacities[31].

Several GPCRs, especially adhesion G protein-coupled receptors (aGPCRs) members, which are characterized by their long N-terminal region, have been identified as important players in neural development and synapse formation[32,33]. The N-terminal of most aGPCRs contains a GPCR autoproteolysis site-including domain (GAIN) and many aGPCRs members have cell-cell adhesion motifs (e.g. EGF-like domains, leucine-rich repeats, cadherin domains and thrombospondin-like repeats)[16]. Among their multiple function in the nervous system, aGPCRs take roles in different processes of neuronal development. For instance, the

members of the BAI (brain-specific angiogenesis inhibitors) subfamily regulate dendrite morphogenesis (i.e. BAI3)[34], the CELSR subfamily regulates neuronal migration (i.e. CELSR1)[35] and dendritic development (i.e. CELSR2 and CELSR3)[36], and a member of the ADGRG subfamily (i.e. GPR56) regulates cortical development and lamination[37].

Besides neuronal development, GPCRs play a role in spine morphology that might affect structural plasticity (the dynamics of spine formation and morphology), which is underlying mechanism of learning and memory as part of synaptic plasticity[38,39]. Like other receptors and molecules (i.e. glutamate receptors, adhesion molecules, small GTPases), GPCRs regulate spine formation and morphology through different G protein-related signaling pathways or independent pathways (i.e. ß-Arrestin or Gßγ dependent pathways) (reviewed in[39]). For instance, the activation of $G\alpha_q$ coupled group I mGluRs by their agonist induce dendritic spine length growth through increased intracellular calcium release and dendritic protein synthesis as shown in hippocampal tissue slices and neuron cultures[40]. Considering the role of GPCRs in neuronal development and spine morphology, it may not be surprising to see that GPCRs play important roles in synaptic transmission and plasticity as well.

2. Synaptic transmission and plasticity

Synaptic transmission can be defined as information processing between pre- and postsynaptic neurons, which involves presynaptic vesicular release in the synaptic cleft and postsynaptic receptor activation[41]. In this process, neurotransmitter-gated ion channels (e.g., AMPAR, NMDAR) are known to be responsible for fast synaptic transmission, whereas GPCRs (e.g., mGluRs, GABABRs, CBRs) are responsible for slow transmission, frequently modulation of synaptic transmission.

Synaptic plasticity is the ability of synapses to change their strength. Based on the persistency of the change of synaptic strength, the form that lasts from seconds to minutes is referred to as short-term synaptic plasticity, whereas persistent changes for hours to days or lifelong in synaptic strength are referred to as long-term synaptic plasticity[42]. The long lasting increase in synaptic strength is known as long-term potentiation (LTP, thought to be principal cellular model of memory)[43] and the decrease in synaptic strength is known as long-term depression (LTD)[44].

Synaptic transmission and plasticity can be studied at the molecular, cellular and behavioral levels. Among all approaches, brain slice electrophysiology provides valuable information for studying synaptic transmission and plasticity[45]. To date, almost 300 GPCRs have been found expressed in hippocampus of which 20 are currently implicated in synaptic transmission and plasticity[7,39]. For instance, type 2 cannabinoid receptor (CB2) has been identified as an important player in synaptic transmission in hippocampus. The chronic activation of the CB2 receptor was demonstrated to induce excitatory synaptic transmission in hippocampal slices[46],

whereas knock-out of the CB2 gene has been shown to reduce excitatory synaptic transmission and LTP, and to have an impact on dendritic spine morphology in the mouse hippocampus[47]. The absence of G protein-coupled receptor mGluR5 (a member of group I metabotropic glutamate receptors) in mice has been shown to reduce LTP in CA1 and dentate gyrus regions of hippocampus and to induce impairments in spatial learning and memory[48].

3. Learning and memory

Learning and memory are essential processes for animal survival. Learning is the process to obtain new information, and memory is the ability to retain and recall this information later in time, which helps to update our behavior and emotions with previous experiences. Patients with memory problems, e.g., with different forms of amnesia, have been instrumental in classification of different types of memory[49]. These can be divided into different forms based on the behavioral form (e.g., being related to explicit or implicit memory) or their duration (e.g., being memorized shortly or long-term).

The information that can be retained for a short period of time is known as short-term memory (STM), whereas the information that can be stored for a long period of time is known as long-term memory (LTM)[50]. These two forms of memory show differences in the underlying molecular and cellular mechanisms[51]. Different forms of memory-related behavior and duration of memory can be investigated using different behavioral paradigms in experimental animals. Behavioral paradigms can be employed to study the involvement of GPCRs or other molecules of interest in specific brain regions to assess their contribution to the different forms of learning and memory. For instance, the Morris water maze (MWM) has been used to assess long-term spatial learning and memory in rodents. In this paradigm, animals learn to make associations with spatial cues in order to find the location of a hidden platform in a water-filled tank. Long-term spatial memory is usually assessed after a 24 h retention interval after the last training[52]. Although the MWM mostly relies on hippocampus, reversal learning involves different brain regions (e.g. orbitofrontal cortex (OFC), medial prefrontal cortex (mPFC), dorsal and ventral regions of striatum and amygdala) and can be investigated by changing the location of hidden platform in the tank[53]. Another hippocampus-related learning paradigm that has been used to investigate context-specific long-term memory (long-term aversive memory) is contextual fear conditioning (cFC), in which animals make associations between the foot shock and context[54]. Similar to cFC, the passive avoidance (PA) test paradigm has been used to investigate safety memory, in which animals associate a foot shock and a dark environment[55]. It has been shown that safety learning and memory in cFC and PA tasks are associated with G protein-coupled CB1 and CB2 receptors in the hippocampus[56].

Deorphanization of GPCRs

Despite an array of strategies to decipher the functional role of orphan receptors *in vivo*, deorphanization, i.e. identification of natural/endogenous ligand(s) of a particular receptor, remains of interest to enable specific targeting of the GPCR at hand. Among all identified GPCRs, approximately 100 remained as orphan GPCR[57]. From a historical perspective, deorphanization of GPCRs started in 1988 by identifying the ligands of 5-hydroxytryptamine (5HT-1A) and the D2 dopamine receptors[58,59]. Two basic approaches are currently applied in the deorphanization process: the forward pharmacology approach and the reverse pharmacology approach[60,61]. In the forward pharmacology approach, the deorphanization process interrogates a known ligand using a receptor library. In contrast, reverse pharmacology interrogates a known receptor using a potential ligand collection (Figure 1). Although, both approaches have their own merits, reverse pharmacology has been considered the most widely used approach for the deorphanization of GPCRs[13].

Different deorphanization strategies can be followed within the reverse pharmacology approach. These are grouped into three overarching strategies: library-based, tissue extract-based (also called orphan receptor strategy) and information-based approaches[62]. In this thesis, I used the tissue extract-based approach for the deorphanization of Gpr158 and hence I will focus on this strategy. In the tissue extract-based strategy, tissue extracts (in this case from brain) are tested in a cell-based functional screening assay, in which the receptor of interest is heterologously expressed. The generation of high-quality tissue extracts, complex and wide-range fractionation of these extracts and choosing and optimizing high quality functional screening assays are the main difficulties of this approach[60]. Additionally, the nature of the ligand (i.e. purine, mono- or trace amine/amino acid, peptide, lipid) is another important, and potentially complicating, factor to consider in the tissue extract-based deorphanization process.

Besides the above considerations, there are multiple factors that need evaluation prior to the deorphanization process. First, the receptor should be expressed in a heterologous system with no endogenous expression, and should be located on the cell membrane and be functionally active. The selection of the type of functional assay in terms of activated intracellular signaling pathways is another critical and important consideration. Currently, there are multiple functional assays commercially available that can be optimized based on the biological relevance of the receptor[63]. Among these, cAMP assays are used to investigate $G\alpha_s$ and $G\alpha_{i/o}$-coupled GPCR signaling, inositol phosphate accumulation assays are used for $G\alpha_{q/11}$-coupled GPCR signaling and reporter gene assay are a general tool used for all G protein-related signaling measurements[63].

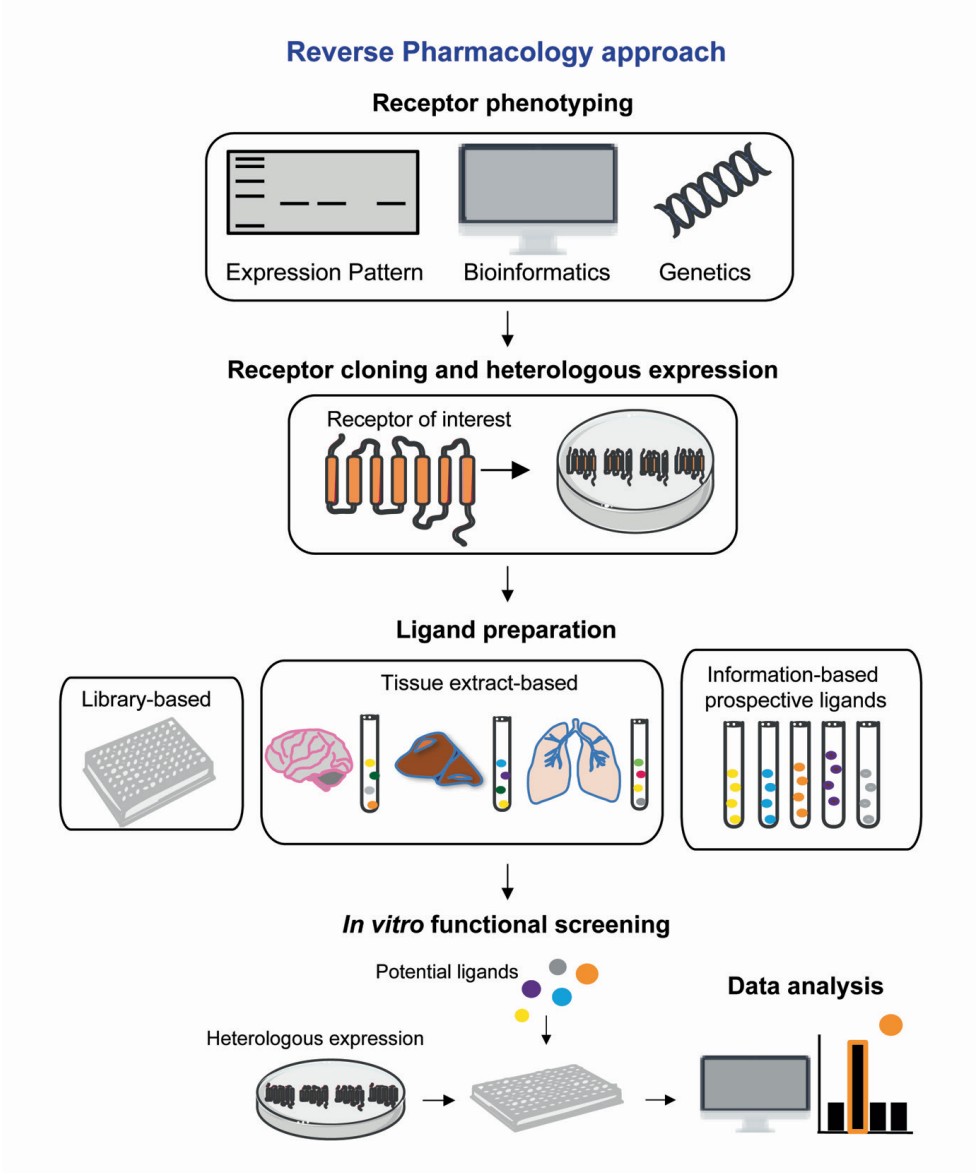

Figure 1. Workflow of the reverse pharmacology approach in the deorphanization process. Reverse pharmacology aims to identify the ligand(s) for a receptor of interest using library-, tissue- or information-based ligand sources. This figure is adapted from Kotarsky and Nilsson, 2004 [60] and Tunaru *et al*., 2017 [61]. Some images are adapted from the smart servier medical art website (https://smart.servier.com).

Experimental approaches to understand Gpr158 function as applied in this thesis

In this thesis, different large-scale screening analyses were applied in addition to more dedicated analyses to explore the function of, as well as a possible ligand for, Gpr158.

First various -omics technologies can be of help. The term 'omics' refers to the comprehensive analysis of biological systems using high-throughput methods, including genomics, proteomics, interactomics and metabolomics. Omics disciplines describe the same biological system at different levels of organization[64,65], where true understanding of the biological system requires the integration of this data[66]. Furthermore, different types of screenings have been used, in which either the content or number of parameters (high-content screening, HCS) or the throughput (High throughput screening; HTS) comes in large numbers. Although HCS and HTC have been used to simultaneously screen multiple parameters in cell-based assays (i.e. intracellular signaling, cell morphology)[67,68], HCS and HTS can be performed at the behavioral level as well[69]. The next sections provide a general overview of the methods and methodologies applied throughout this thesis

Quantitative proteomics

Within neurosciences, quantitative proteomics involves the identification and quantitation of biochemical preparations (e.g. homogenate, synaptosome) of specific brain tissue at a given time and condition of the animal. In recent years, mass spectrometry (MS)-based proteomic approaches have become the method of choice, where the so called "bottom-up" approach quantitatively analyzes the peptides resulting from digestion of the proteins in the samples to compare. The general workflow of the "bottom-up" approach includes sample preparation, protein digestion, peptide fractionation and protein identification and quantification through MS and data analysis.

In recent years, several quantitative strategies have been employed using label-free[70] and label-based methods[71]. In contrast to label-based methods, label-free quantitative methods omit the incorporation or coupling of stable isotopes to proteins or peptides. Currently, several label-free quantitative methods are being used for performing quantitative proteomics analysis at the proteome level, among which the discovery-oriented data independent acquisition (DIA) has rapidly emerged as a powerful technique for label-free quantitative proteomics[72,73]. The 'sequential window acquisition of all theoretical mass spectra' (SWATH) is a DIA method coupled with peptide spectral library match that allows the large-scale quantification of proteins across multiple samples with good precision, reproducibility and proteome coverage[72,74,75].

In chapter 3, I used a SWATH analysis for the detection of differential regulated proteins in the hippocampus of *Gpr158* KO mice.

Metabolomics

Metabolomics is described as the inclusive study of all metabolites (metabolome) that are present in a given biological system or sample[76]. When the analyzed metabolites are lipids and sugars, the corresponding omics studies are referred to collectively as lipidomics and glycomics, respectively. The human metabolome includes small molecules within a mass range of 50–1500 Da with very different chemical and physical properties[77]. The chemical diversity of the metabolome creates a chemical signature that uniquely defines the state of a biological system at a given time and condition. However, this chemical diversity makes them difficult to analyze using a single analytical platform or approach. Other challenges include the lack of unique spectral signatures for many metabolites and the detection of low abundant metabolites.

Metabolomics strategies are often divided into targeted (metabolite profiling or biology-driven) and non-targeted approaches. Metabolic strategies for both MS and NMR platforms are usually applied for discovering biomarkers and providing insight into mechanisms of biological processes[78]. Other applications include protein–ligand explorations and high-throughput ligand affinity screening[79].

LC-MS/MS-based lipidomics are often for qualitative and quantitative analysis of lipid composition. Similar to metabolomics, lipidomic strategies can be divided into non-targeted and targeted approaches. Among non-targeted approaches, the LC-based IDA method is widely used, where a list of candidate precursor ion peaks with a MS/MS spectrum is generated and user-defined information-dependent criteria to identify lipids are used[80].

In chapter 5, I used a LC-MS/MS analysis for the detection of the compound that activated Gpr158 in a heterologous high-throughput GPCR assay (see below).

High-content (Cellomics) and High-throughput screening

High-content screening (HCS) on cell cultures (Cellomics) is a discipline for quantifying cell-based phenotypic parameters using microscopy-based imaging and informatics methods. Image-based HCS can be used for analyzing protein localization and morphological changes induced by the state of the cells (e.g. after a physiological stimulus). In addition, molecular interference of expression like RNAi or molecular stimulation by small-molecule libraries can be used[81]. In neurosciences, HCS has been used to investigate neurite outgrowth and morphology [68], as well as cellular phenotyping in neurological disease[82]. Also, the impact of schizophrenia risk genes on neuronal morphology has been investigated using HCS combined with proteomics[83]. Furthermore, HCS has been used for investigating the external factors on neuronal development process in neuron-astrocyte co-culture assay (e.g., van Deijk *et al*., 2017[84]).

On the other end of the spectrum, high-throughput screening (HTS) allows for the quantification of cell-based signaling, which is a key process in drug discovery[81,85]. This technology has been used in the GPCR drug discovery field, where the effect of compound(s) or biomolecule(s) of interest on G-protein signaling is investigated using functional screening assays[63].

In chapters 2, and 4, I used HCS using Cellomics to analyze neuronal morphology and synapse formation. In addition, in chapter 5 I used HTS to screen the G-protein signaling for an endogenous ligand of Gpr158 derived from mouse brain sample.

(Automated home cage) behavioral phenotyping

For compounds or receptors, it is of interest to screen for specific behavioral phenotypes elicited. There are many conventional behavioral paradigms to assess learning and memory in rodents. For instance, the Morris water maze and radial-arm maze tests are well known for the assessment of spatial learning and memory[86], whereas fear conditioning and inhibitory avoidance tests are well known paradigms for the assessment of associative memory[87,88]. Although conventional behavioral paradigms give the opportunity to compare experimental data with existing literature, the animal-experimenter interaction increases the stress level of animals that might affect behavioral performance[89].

Automated behavioral tests provide a high rate of repeatability while significantly shortening experimental time and data analysis. In automated home-cage behavioral phenotyping, home-cage behavior (i.e. drinking, eating, grooming, resting and walking) of mice can be automatically measured and analyzed[90]. In addition, more sophisticated measurements including anxiety[91] and attention as well as discrimination and reversal learning performance can be investigated in the automated home-cage without animal-experimenter interaction[92,93].

In chapters 2 and 3, I used either classical tests for learning and memory (MWM, cFC, PA) and anxiety (open field), or automated home cage behavior of discrimination and reversal learning to analyze mice devoid of *Gpr158*.

Thesis outline

The primary aim of this thesis was to explore the function of Gpr158. The experimental chapters provide the functional characterization of Gpr158 at the molecular, cellular, physiological and behavioral levels by using *Gpr158* KO mice or by *in vitro* gene knockdown. The functional characterization of Gpr158 was carried out using both conventional and high-throughput automated home cage behavioral phenotyping, quantitative proteomics, high-content screening (Cellomics), cellular based high-throughput functional screening and metabolomics.

In chapter 2, I examined the role of *Gpr158* in learning and memory using hippocampus-dependent behavioral paradigms and electrophysiological measurements in CA1 pyramidal neurons of *Gpr158* KO mice. Furthermore, I investigate the effect of *Gpr158* gene deletion on neuronal development both *in vitro* and *ex vivo*. Our data revealed that *Gpr158* KO mice show spatial learning and memory deficits in the MWM task and a deficit in the acquisition of extinction memory using the PA paradigm. In the absence of *Gpr158*, CA1 pyramidal neurons revealed normal basal excitatory and inhibitory synaptic transmission. However, Schaffer collateral stimulation yielded reduced post-synaptic currents. Additionally, intrinsic excitability of CA1 pyramidal neurons was found increased in our study. Regarding neuronal morphological measurements, we found that *Gpr158* deficiency induces morphological alterations, i.e., reduced dendritic architecture and complexity, in CA1 pyramidal neurons in *Gpr158* KO mice and primary hippocampal neurons. This morphological impairment in CA1 pyramidal neurons was consistent with the observed spatial memory acquisition in *Gpr158* KO mice.

In chapter 3, I aimed to explore whether *Gpr158* has a role in different learning paradigms in addition to its role in spatial learning and memory (see chapter 2). For this aim, I used the CognitionWall test to determine the discrimination and reversal learning performance of *Gpr158* KO mice. Although acquisition of discrimination learning was not affected in *Gpr158* KO mice, they showed a deficit in the reversal phase of the CognitionWall test. Furthermore, in this chapter I performed an analysis to explore the interactome profile of Gpr158 in hippocampus and striatum. We found the same protein interactors (Gβ5 and RGS7) as previously reported[15], but also Gαz as a new interactor for Gpr158 in both regions, and a few region-specific putative interactors for Gpr158. By using a quantitative proteomic approach (SWATH-MS), we reported that gene deletion of *Gpr158* reduces the levels of RGS7 and Gβ5 and increases the level of KIF1 in the hippocampal synaptosome samples, and putatively affects proteins related to neurodevelopment. As such, the SWATH-MS data support the role of *Gpr158* in neuronal development as shown in chapter 2. Collectively, in this chapter I showed that *Gpr158* plays a role in multiple learning forms and the absence of *Gpr158* induces proteomic changes in the hippocampal synaptosome fractions.

In chapter 4, I investigated the function of *Gpr158* at the synapse. For this aim, I optimized and performed a mixed co-culture assay[28], and a synapse density assay for HCS. We measured accumulation of the presynaptic marker synapsin in primary hippocampal axons that were in contact with *Gpr158*-overexpressed HeLa cells. In line with the recent findings of Condomitti et al.[23], we found that Gpr158 induces synapse formation in primary hippocampal axons. In addition, I found that the knock-down of *Gpr158* in primary hippocampal neurons decreased the synapse density of primary hippocampus neurons and a specific reduction of the density of the post-synaptic protein PSD-95, whereas density of the presynaptic protein

synapsin remained the same. Together these findings indicate that Gpr158 is an important player in synapse formation in primary hippocampal neurons.

In chapter 5, I used a tissue extract-based deorphanization strategy for Gpr158. Using a luciferase reporter gene assay in HEK293T cells, I found that Gpr158 is coupled to the $G\alpha_{q/11}$ subunit and induced the activation of the NFAT response element (RE) through a non-classical $G\alpha_{q/11}$ pathway in these cells. As input material for the tissue-extract based deorphanization, I analyzed multiple protein and lipid fractions from synaptosome-enriched P2 microsome forebrain samples. Using multiple approaches (proteomic, metabolomic and cell-based high-throughput functional assays), we showed the presence of a small lipid-like Gpr158-activating compound. Interestingly, MS analysis and chemical structure prediction pointed to a tissue-induced O-formylation of dodecyl *N*-Dodecyl-D-Maltoside (DDM), the most widely used detergent for tissue sample preparation, that was not present in buffer samples that were run in parallel. Taken together, we provided valuable information for identifying a true activator, but false natural ligand, highlighting the possible pitfalls in GPCR deorphanization.

In chapter 6, the findings of each experimental chapter are discussed in light of current literature. Furthermore, the strengths, limitations and future research directions of this thesis are discussed.