Structural mechanism of calcium-mediated drug binding and G-protein coupling to the melanocortin 4 receptor

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Nicolas Andreas Heyder

Berlin, 2022

1st reviewer:	Dr. Patrick Scheerer
2nd reviewer:	Prof. Dr. Christian Freund
Date of defense:	02.06.2022

Acknowledgment

First and foremost, I would like to thank my advisor, Patrick Scheerer. Patrick is kind, supportive, and believes in his students. His focus on tackling challenging projects is inspiring and allows the freedom to navigate minor problems freely. Patrick has given me tremendous independence and financial backing in tackling this tour de force. In my first two years, He trusted me to go to the Kobilka lab in Beijing and Tsinghua as a representative of his lab. Then, for four and a half years, he supported me in changing central experimental frameworks and believed in the chance of success when I doubted.

I want to thank all members of the Scheerer lab, whose support, guidance, and kindness have been invaluable. Michal Szczepek invested a great deal of time and effort in guiding me into membrane-protein research, and our journey to China is unforgettable. Gunnar Kleinau spent countless hours drafting, correcting, and thinking about our MC4R manuscripts. Andrea Schmidt built the models of the MC4R complexes and inspired me with her unagitated manner in handling an arbitrary number of tasks with kindness. Furthermore, I would like to thank everyone from IMPB, especially the Schwefel lab that provided us with the bacmid DNA.

I want to give special thanks to my fellow graduate students, David Speck, Dennis Kwiatkowski, Luisa Herder and Monique Gallandi. David and I spent the first COVID-19 lockdown together in the cell culture and expressed the protein used for the here shown structures.

Brian Bauer and Anja Koch are the heart and engine of the lab. Anja cloned the constructs, handled the HEK293 cell culture, and helped me daily in finding everything. After establishing the protocol, Brian managed the G-protein production protocol, thereby allowing me to focus on the receptors and structural experiments.

Tarek Hilal and Sarah Paisdzior are great collaborators; they taught me the fundamentals of cryo-EM techniques and signaling-assay analysis, respectively. Daniel Hilger guided the complex preparation process from afar and inspired me to test the nanoBRET idea. Magdalena Schacherl invested countless hours in teaching me EM data processing, and her mentorship is an endless source of ideas.

Ultimately, I would like to thank Dr. Patrick Scheerer and Prof. Christian Freund for reviewing this thesis.

Selbstständigkeitserklärung

Hierdurch versichere ich, dass ich meine Dissertation selbstständig verfasst und keine anderen als die von mir angegeben Quellen und Hilfsmittel verwendet habe.

Agenda

Abstract

The melanocortin-4 receptor (MC4R) is a hypothalamic regulator of energy homeostasis and appetite. Hence, this G-protein coupled receptor (GPCR) is a prime pharmaceutical target for obesity treatment. During this thesis, the central experimental prerequisites for *in vitro* GPCR studies were established: (1) cell culture techniques for expressing receptors and G-proteins, (2) a fluorescent-based in cell ligand-binding assay for assessing suitable ligands, and (3) a receptor-G-protein complex formation and purification protocol. These biochemical experiments lead to the first cryo-electron microscopy structures of a GPCR-G-protein complex in Germany, the MC4R-G_s complexes with two novel FDA-approved drugs, the peptide agonists NDP-α-MSH and setmelanotide, with 2.9 Å and 2.6 Å resolution, respectively.

Both complexes demonstrate the importance of calcium as a cofactor for agonist binding, precisely adjusted depending on the different ligands. Furthermore, the structures together with signaling data reveal both the agonist-induced origin of the transmembrane helix (TM) 6-dependent receptor activation and the mediator role of TM3 in fine-tuned activation, driven by interactions with the respective agonist and calcium. This interplay proceeds towards the receptor G_s-protein interface, particularly at intracellular loop three, impacting MC4R's G-protein coupling specificity.

These structures uncover mechanistic details of MC4R activation, inhibition, and G-protein coupling specificity, facilitating the development of optimized drugs.

Zusammenfassung

Der Melanocortin-4 Rezeptor (MC4R) reguliert das Energiegleichgewicht und den Appetit im Hypothalamus des menschlichen Gehirns. Dieser peptidische G-protein gekoppelten Rezeptor (GPCR) ist daher von hohem medizinischem Interesse für die pharmakologische Behandlung von Adipositas. Diese Arbeit umfasst die Etablierung der Methoden für in *vitro* GPCR Studien. Dazu zählen Methoden für (1) die heterologe Expression von GPCRs und G-proteinen, (2) die Etablierung eines fluoreszenzbasierten Ligandenbindungsassay und (3) das Protokoll für die MC4R-G-protein Komplexierung. Diese biochemischen Arbeiten erlaubten die ersten strukturellen Analysen eines GPCR-G-protein Komplexes in Deutschland, die Kryo-Elektronenmikroskopstrukturen des MC4R-G_s Komplexes mit zwei kürzlich FDA zugelassenen Medikamenten, die peptidischen Agonisten NDP-α-MSH und Setmelanotide mit Auflösungen von 2.9 Å und 2.6 Å. Die Strukturen zeigen die Bedeutung eines Calciumions als Kofaktor für die Stabilisierung der Agonistinteraktion mit MC4R auf. Die Strukturen und zellbasierte Signalisierungsdaten legen den Ursprung der Liganden induzierten Transmembranhelix (TM) 6 abhängigen Rezeptoraktivierung offen. Des Weiteren, wird die Rolle der TM3 als Vermittler der feinjustierten Rezeptoraktivierung, die durch das Zwischenspiel der Agonisten und dem Calciumion gelenkt wird aufgezeigt. Dieses Zwischenspiel setzt sich bis an die MC4R-G-protein Berührungsfläche fort, insbesondere am intrazellulären Loop 2. Hier beeinflusst dieses die ligandenabhängige G-protein Spezifität.

Die hier beschriebenen Strukturen zeigen die mechanistischen Details der Aktivierung und Inhibierung des MC4R sowie seiner G-protein Kopplungsspezifität des MC4R. Sie können als strukturelle Basis bei der Entwicklung von optimierten Medikamenten zur Behandlung der Adipositas ein strukturelles Fundament beitragen.

1 Introduction

1.1 G-protein coupled receptors

The most prominent eukaryotic cell surface receptors family are G-protein coupled receptors (GPCRs), in humans encoded by 800 different genes (4% of the human genome) [1]. This thesis will focus on the rhodopsin-like or class A GPCR family, the largest family with 700 members. They share a common activation mechanism [2] and a seven-transmembrane helix (TM) topology that connects the extracellular and intracellular environment. GPCRs sense various external stimuli, including light, hormones, odorants, lipids, peptides, small molecule compounds, and whole protein domains [3]. Their ubiquitous role in nearly all aspects of human physiology made GPCRs a prime target for drug development resulting in a share of approximately 30% of all FDA-approved drugs [4]. Receptor stimulation triggers internal signal transduction pathways by heterotrimeric G-proteins. Agonist binding outside the cell stabilizes a receptor conformation that interacts with its cognate heterotrimeric G-protein inside the cell. The GPCR acts as the guanidine exchange factor (GEF) by inducing the GDP to GTP exchange at the G-protein core. This

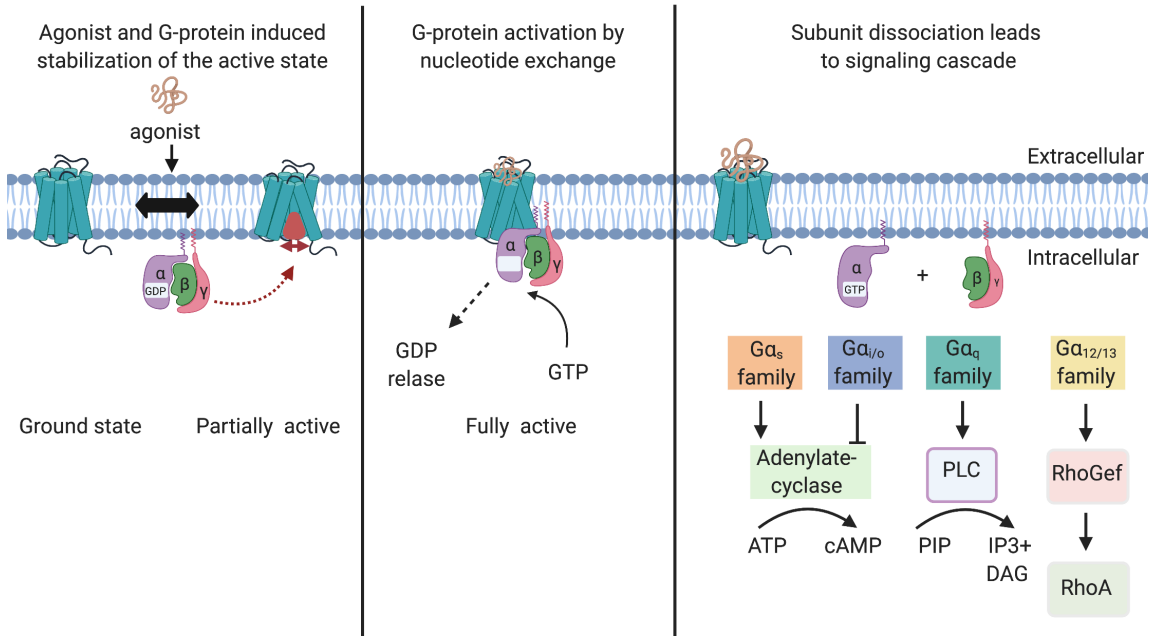

Figure 1.1.1 ***GPCRs are activated by binding agonists and heterotrimeric G-proteins - the activation results in the induction of G-protein-dependent signaling pathways***. The schematic figure was generated using biorender.com.

exchange triggers the dissociation and disassembly of the G-protein into the Gα and Gβγ subunits. The dissociated subunits initiate signal transduction pathways by engaging with downstream effector proteins [5]. The Gα subunit's identity, which interacts directly with the receptor, defines the G-protein mediated (or canonical) signaling pathway (Figure 1.1.1). Receptor kinases can phosphorylate activated receptors at the intracellular loops (ILs) and the receptor carboxy terminus (Ctt). Upon phosphorylation, receptors can be engaged by β-arrestins [6]. This engagement arrests G-protein signaling, can induce receptor internalization, and trigger various G-protein independent signaling (or non-canonical) pathways [7]. Intriguingly, the same intracellular receptor pocket recognizes G-proteins [8], β-arrestins [9] and receptor kinases [10].

The structural investigation of GPCRs is challenging research due to their intrinsic flexibility and low stability in membrane mimicking agents. Structures of GPCR-G-protein allow delineating the molecular mechanism of ligand binding, receptor activation, G-protein recognition and are detrimental for the development of novel GPCR targeting drugs.

1.1.1 Improved experimental access to structures of inactive GPCRs and active receptor-G-protein complexes

In recent years, the number of GPCR structures has grown, driven by advanced membrane-protein biochemistry and structure determination methodologies.

Protein production and purification were facilitated by the standardization of the baculovirus *Spodoptera frugiperda* 9 (*Sf*9) insect cell protein expression system [11], novel detergents such as 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (LMNG) [12], and rapid purification strategy protocols [13].

Structure determination of inactive receptors by X-ray crystallography was enabled by (i) increasing access to synchrotrons, (ii) *in meso* crystallization techniques [14] and (iii) extensive protein engineering (such as fusion proteins [15] and thermostabilizing mutations). Nevertheless, membrane proteins are difficult to grow into well-ordered

crystals due to their inherent flexibility and lack of large polar surfaces that enable stable crystal packing. In addition, the production of protein crystals demands large amounts of pure and homogeneous protein that is stable enough to sustain the crystallization process.

Two historic model systems for the structural understanding of class A GPCRs are rhodopsin and the β2-adrenergic receptor (β2AR). Crystal structures of rhodopsin in the inactive [16] and active conformation stabilized by pH [17] or a G-protein mimicking peptide [8] laid the foundation for understanding canonical GPCR activation.

Crystal structures of β2AR in its inactive [18] and nanobody (Nb) 80 stabilized active state [19] confirmed universal aspects of class A GPCR activation, such as the outward movement of TM6. The structure of β2AR coupled to the heterotrimeric G_s-protein [20] expanded the picture by displaying the fully active receptor state engaging the G-protein.

This first generation of GPCR structures revealed that ligand binding at the extracellular orthosteric site induces structural rearrangements in conserved structural motifs and conformational switches, namely the $C^{6.47}W^{6.48}xP^{6.50}$, $N^{7.49}P^{7.50}xxY^{7.53}$, $P^{5.50}I^{3.40}F^{6.44}$, and $D^{3.49}R^{3.50}Y^{3.51}$ motifs [superscripted numbers according to the unifying Ballesteros & Weinstein numbering for class A GPCRs [21]]. These changes result in the outward movement of the cytoplasmic half of TM6. This translocation allows the Gα subunit of the G-protein to bind the resulting cytoplasmic receptor opening with the carboxy tail, the α5-helix, and induces the GDP release.

Until January 2020, 72 different GPCR were solved by X-ray crystallography (264 inactive and 88 active structures [footnote 1]), including only one, the aforementioned β2AR-G_s structure, of a GPCR in complex with a heterotrimeric G-protein [(20; footnote 2)]. However, due to GPCRs conformational heterogeneity, their structure determination by X-ray crystallography is limited by crystal formation. In contrast, cryo-EM allows the visualization of GPCR-complexes in vitrified aqueous solution and reduces the need for time-intensive receptor stabilization by mutagenesis.

In the advent of the cryo-EM resolution revolution [22], researchers used cryo-EM to determine GPCR-G-protein complex structures.

[1] www.gpcrdb.org/structures; 29.01.2021
[2] In June 2021 – the crystal structure of the dopamine 1-Gs complex was published.

Methodologically, this revolution was driven by the introduction of direct electron counting detectors, the development of cryo specimen preparation methods [23], novel computational approaches, and increased computational power. Therefore, until January 2020, 14 GPCRs were structurally characterized (80 GPCR-G-protein-complex structures) via cryo-EM [footnote 3].

For cryo-EM, a homogenous protein sample is applied onto carbon or gold grids, vitrified in liquid ethane, and imaged with a cryo-electron microscope, equipped with a direct electron counting detector. In addition, the required sample amount is drastically smaller, less than one milligram, compared to a crystallization experiment, allowing to address receptors with lower expression levels. Furthermore, the vitrification of an aqueous solution circumvents the need for a rigidified crystallized sample, and therefore, several states can be captured in one experiment, as shown for the NTSR1-G_i complexes in the canonical and non-canonical conformation [24]. However, a high particle density of a pure and homogeneous sample in random orientations on the grid is required. Hence, cryo-EM experiments require extensive biochemical efforts and specimen preparation and optimization iterations.

1.1.2 GPCR activation - a dynamic process facilitated via conserved microswitches

Most GPCRs are not binary on/off switches but are instead described by a conformational equilibrium (or dynamic process) [25], shown by combinations of spectroscopic [26] and structural studies (*e.g.*, [27,28]). The basal activity level, the ligand-independent GPCR capacity to induce guanidine nucleotide exchange, is modulated by different types of ligands (Figure 1.1.3). Inverse agonists facilitate the reduction of basal signaling below its baseline. The active conformation is stabilized by agonists, resulting in an increased GEF activity. Antagonists compete for binding with inverse agonists or agonists and maintain the receptor equilibrium, the ratio of the inactive and active receptor populations. Biased agonists stabilize a conformation that favors one signaling pathway and acts as an antagonist to induce alternative pathways [29]. Many GPCR studies are centered around the usage of pharmacologically well-characterized ligands. Historically their usage allowed to identify and clone GPCRs using affinity chromatography (*e.g.*, [30]). In addition, ligands shift

[3] www.gpcrdb.org/structures; 29.01.2021

the transient conformational receptor equilibrium to the desired state (*e.g.*, inactive or active) and make them accessible for investigation. GPCRs with diffusible ligands have two agonist affinity states, the low-affinity state in the absence and a high-affinity state in the presence of a cognate G-protein [31]. This model supported the idea of a transmembrane-spanning allosteric relationship from the orthosteric ligand-binding pocket (LBP) to the G-protein binding cavity (GBC).

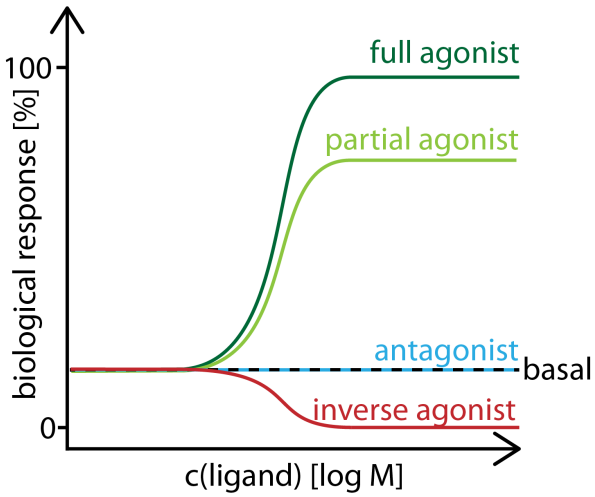

Figure 1.1.3 ***Different ligand types can induce various responses at a GPCR***. Concentration--response representation of different GPCR ligand classes impacting the biological response.

This theory was supported by a β2AR structure complexed with a covalently bond agonist that displayed structural characteristics of an inactive receptor conformation in the intracellular half of the receptor – the low-affinity state [32]. Hence, agonists alone do not stabilize the fully active conformation independently. Here, the presence of an intracellular stabilizing entity is required to lock the fully active state. Such entities include G-protein-derived peptides [8,33], nanobodies [34,35], engineered G-proteins [36], and heterotrimeric G-proteins. Furthermore, the allosteric nature of GPCRs is not limited to the active state. It also accounts for the allosteric coupling of antagonists with a nanobody stabilizing an inactive conformation [37]. Thus, GPCRs are classic allosteric proteins with loose allosteric coupling between the LBP and GBC. Consequently, the structural investigation of fully active receptors requires both the complexation with agonists (except for highly basal active receptors) and a stabilization of the open receptor conformation.

GPCRs have a low sequence similarity in the upper half of the receptor responsible for recognizing diverse ligands. The structural conservation increases towards the G-protein interface, differentiating only a tiny subset of interaction partners. One highly conserved feature in the ligand-binding pocket is the disulfide bridge between EL2 and $C^{3.25}$ at TM3. This interaction is required for receptor stability and agonist binding in various GPCRs [38]. The melanocortin-4 receptor (MC4R) is at this position exceptional with $D122^{3.25}$.

GPCR activation has been associated with highly conserved motifs that connect the orthosteric site with the G-protein binding cavity. The primary description of these microswitches was performed using the rhodopsin system [3]; hence these are described based on the active [33] and inactive [16] rhodopsin structures. Of note, other GPCRs have different side-chain interactions among those motifs, but intriguingly the general mode of action is highly conserved.

At the transmembrane core, the primary interaction interfaces are at the central and intracellular sides of the TM bundle. TM1-TM2, TM3-TM4, TM3-TM5, and TM3-TM6-TM7 comprise these interfaces. TM1 and TM2 do not undergo conformational transitions upon receptor activation. Studies in the neurotensin receptor 1 indicate that TM1 and TM2 are relevant for receptor translation and membrane insertion but not for receptor activation and ligand binding [39]. TM3 is the helix with the most conserved interfaces, connecting to all TMs except TM1 and TM7, and has been assigned a key role in maintaining the scaffold in all GPCR structures for active and inactive conformational states [39]. Interactions in the TM3-TM6 interface are most conserved between residues 3.36 - 6.48, and 3.40 - 6.44 [39]. Here, $W^{6.48}$ is part of the *CWxP*$^{6.50}$ motif in TM6. A proline kink in TM6 mainly characterizes receptors harboring this microswitch. The so-called rotamer toggle switch model [40] claims that $W^{6.48}$ and $F^{6.52}$ are coupled and change during receptor activation. This model has been adapted in recent years by focusing solely on the downward shift of $W^{6.48}$. This residue is at the pivot point of the TM6 outward positioning that is observed in active state structures and allows G-protein binding.

At the TM7 intracellular site is the *NP*$^{7.50}$*xxY*$^{7.53}$ motif that constraints the TM1-TM2-TM7 interface in the inactive state. The central residues are $N^{7.49}$ at TM7 with $D^{2.50}$ at TM2 and $D^{1.50}$ at TM1 [3]. Upon activation, $Y^{7.53}$ rotates towards TM5 and is probably involved in stabilizing the inward conformation of TM5 [17,41].

In addition, the TM3-TM5 and TM6 interface is conserved by the *PIF* motif between $P^{5.50}$, $I^{3.40}$, and $F^{6.44}$. The *PIF* motif is significant for receptor integrity and the stabilization of inactive conformations [42]. MC4R has at TM5 in contrast to most GPCRs an $M^{5.50,}$ and we proposed that TM5 is not kinked in contrast to receptors with a proline at that position [43,44].

At the G-protein interface, the structural hub TM3 forms via the *DRY* motif an interaction with its extension IL2 and connects IL2 with the TM core. In the muscarinic system, this interaction is formed via a salt bridge between $D^{3.49}$ at TM3 with arginine at IL2 [45]. In the case of β2AR, the latter residue is a conserved tyrosine. Nevertheless, in both cases, the tethering of IL2 to TM3 is conserved via the same residue positions.

$R^{3.50}$ at TM3 is conserved in 96% of class A GPCRs [46] and forms an intrahelical contact to the adjacent E/D [3.49]. In many GPCRs, like rhodopsin [3], $R^{3.50}$ tethers to TM6 via the 6.30 position. However, the TM3-TM6 lock is not conserved among all class A GPCRs. For instance, the TM6 in β2AR is linked to TM3 via IL2. Upon receptor activation, $R^{3.50}$ at TM3 forms a hydrogen bond to $Y^{5.58}$ and thereby releases the TM3-TM6 restraint and stabilizes the inward conformation of TM5.

In summary, GPCR activation is governed via conserved microswitches, but their interaction pattern is uniquely adapted for each receptor. Hence, the activation mechanism of each receptor requires thorough investigation.

1.1.3 G-protein coupling specificity

Generally, each of the 800 GPCRs can couple to more than one G-protein and has a unique signaling profile [47,48]. In contrast, the number of G-proteins is limited, and a complete understanding of this GPCR-G-protein specificity is still not fully established.
On a broader scale, gene expression regulates G-protein specificity on the tissue and cellular level for most GPCRs; and modulates the available GPCR and G-protein populations. Nevertheless, most cells express G-proteins ubiquitously [49]. Hence, coupling specificity is expected to be determined at the molecular level, at the GPCR-G-protein interface.
G-protein binding to the GPCR results in the release of GDP; the nucleotide release induces the dissociation of the heterotrimer and drives the downstream signaling cascades (Figure 1.1.1). The comparison of GTP-bound G-protein with receptor-G-protein complexes

illuminate the molecular mechanisms of GDP release. Receptor conformations with an outward tilted TM6 allow the Gα subunit of the G-protein to bind the resulting cytoplasmic receptor opening with the carboxy tail, the α5-helix (first shown in [41]). The 6 Å displacement of the α5-helix into the receptor equals the distance of the helical domain opening during GDP release [49].

This α5-helix insertion loosens interactions of the Gα subunit with GDP, resulting in GDP dissociation and subsequent GTP binding. The α5-helix undergoes a disorder-to-order transition, meaning that the helix is disordered in the GTP bound state and is helical if bound to a GPCR [49]. The α1-helix acts as a connective hub in the inactive G-protein state by linking the α5-helix, H-domain, and GDP [50]. The α5-helix insertion into the GPCR loosens this interaction network. Consequently, the α1-helix undergoes an order-to-disorder transition mandatory for GDP release [49].

Sixteen human Gα genes are subdivided into four families, $G\alpha_s$, $G\alpha_{i/o}$, $G\alpha_{q/11}$, $G\alpha_{12/13}$, that specify the specific downstream signaling cascade [51] (Figure 1.1.1). The G_s and $G_{i/o}$ family controls intracellular cAMP levels by activation or inhibition of adenylate cyclase, respectively. The $G_{q/11}$ family modulates cytoplasmic Ca^{2+} levels by activating phospholipase C (PLC), and the $G_{12/13}$ family predominantly addresses the RhoGef-Rho pathway. By sequence analysis, efforts to predict coupling specificity or a G-protein barcode that defines which GPCRs are addressed were made [52,53] but did not result in a predictive theory. The aforementioned first generations of active GPCR structures highlighted the importance of the GαC-terminus (GαCT), the α5-helix, which embodies the GPCR-G-protein interface with approximately > 70% of the buried surface area [49]. Early mutagenesis studies have shown that exchanging the last 4-6 GαCT residues can switch coupling selectivity from G_q to G_i [54]. The residues at the -4 and -3 positions (counting from the carboxy tail of Gα) define G_i vs. G_s coupling via a CG or YE motif, respectively. These two residue positions have been predicted by computational means [55] and shown by structural studies in the rhodopsin system (manuscript in preparation, Kwiatkowski *et al.* 2021).

Nevertheless, GPCR-G-protein coupling specificity cannot be adequately predicted based on linear sequence comparisons, and a three-dimensional structural barcode on the receptor side is more likely. Upon GPCR activation, TM3, TM5, TM6, and their extension IL2, and IL3, undergo the largest reorientations at the intracellular site and have been implicated with coupling specificity [47,56–58]. Muscarinic 3 receptor (M3R) couples endogenously via

G_q but switches to G_s by exchanging IL2 and IL3 from the G_s-coupled b1AR [59]. Since the absence of the $G\alpha_s$ Y at the -4 position at the $G\alpha_{i,}$ carboxy terminus is implicated with the lesser extended TM6 for G_i compared to G_s-coupled GPCR-G-protein complexes [60,61]. Ultimately, to discriminate a potential GPCR-G-protein coupling theory, there is a need for more biophysical and structural data of non-described GPCR-G-protein pairs.

A highly interesting receptor-G-protein protein is the MC4R in complex with the G_s-protein, MC4R's primary signaling partner. This receptor is a peptidic class A GPCRs that belongs to the family of melanocortin receptors.

1.2 Melanocortin-4 receptor

1.2.1 Obesity and the melanocortin system

Obesity is defined by a body mass index (BMI) of above 30 kgm^{-2} and is one of the severest global health burdens. In a 2016 study, the WHO estimated that 650 million adults are affected globally (WHO, 2020). Severely increased body weight is associated with comorbidities like type 2 diabetes and cardiovascular diseases [62]. The apparent main driver of the increased obesity prevalence is driven by the consumption of high caloric foods in combination with reduced physical activity. This prominent driver fosters the public belief that body weight regulation is a matter of willpower [63].

However, twin and adaption studies unraveled a 40% - 70% heritability of the BMI [64,65]. Besides health concerns, the stigmatization of people affected by obesity has lasting social and physiological effects [63]. Human and mouse studies have identified neuronal signaling circuits that regulate mammalian appetite and metabolism. Here, the melanocortin pathway is the best-characterized pathway, and genetic disruptions are the most common cause of monogenic obesity [66]. Hence, a better understanding of key players of the melanocortin pathway can facilitate the development of novel anti-obesity drugs.

The melanocortin pathway consists of three main components. (i) Five G-protein coupled receptors, the melanocortin receptors (MCR), are regulated by (ii) the products of the pro-peptide proopiomelanocortin (POMC) and (iii) agouti and agouti-related peptides (AgRP). POMC is a 30 kDa protein that is post-translationally processed by prohormone convertases into a multitude of biologically active peptides, including adrenocorticotropin hormone (ACTH) and the melanocyte-stimulating hormones (MSH) (α-MSH, β-MSH, γ-MSH). Agouti protein [67] and AgRP [68] act as endogenous antagonists on the MCRs, regulating various physiological functions like melanogenesis, sexual drive, and energy metabolism.

This thesis focuses on the melanocortin-4 receptor and its role in appetite regulation and energy metabolism. The characterization of the melanocortin pathway as a central regulator of feeding behavior started in the 1950s and is still ongoing.

1.2.2 The melanocortin system and cloning of MCRs

First experiments with hormone extracts containing ACTH and melanocortin hormones were conducted in the 1950s. Initial research was somewhat subjective and done in dogs that started to stretch their backs after hormone extract administration [69] and rats that were more sensitive to electroshocks [70]. During the 1970s, the expression of MSH was designated to the hypothalamus, and its function in skin pigmentation was revealed [71]. During the 1980s, POMC was cloned from *rat* POMC neurons [72], and the Hruby laboratory extensively investigated melanocortin peptide derivatives (*e.g.*, [73,74]) that laid the foundation for the structure-activity relationship of MSH derived drugs used today (Section 1.3). Nevertheless, it took until 1990 to find the related binding sites of MSH in the brain [75].

Two years later, the first two melanocortin receptors, MC1R and MC2R, were cloned from a melanoma sample binding to α-MSH and ACTH, respectively [76]. MC1R regulates melanogenesis, and both receptors share a high sequence similarity since α-MSH is the cleavage product of the thirteen amino-terminal residues of ACTH. The cloning of two additional brain receptors, MC3R, MC4R [77], and the peripheral MC5R [78], followed.

1.2.3 Melanocortin pathway player and the connection to feeding behavior

Already, in 1986 it was hypothesized that melanocortin peptides inhibit food intake [79]. Genetic disruption of MC4R in mice was directly linked to obesity [80] and reports of MC4R frameshift mutations in humans followed in 1998 [81,82]. MC4R variants are the most common cause of monogenic obesity, and to date, more than 150 residues are associated with naturally occurring MC4R mutations (a systematic summary of the pathogenic mutations has been published in [44]). Furthermore, 4% of investigated patients with a BMI > 30 kgm^{-2} carry an MC4R mutation [83]. To exclude environmental factors, humans with different ethnic backgrounds were analyzed, and MC4R mutations were confirmed as the most prevalent gene mutations in obese patients [84,85]. Hence, it became clear that the genetic disruption of MC4R causes obesity, driven by either the reduction of energy expenditure, increased energy consumption, or a combination of both. A study showed that wild-type mice exposed to a diet with increased fat content display increased

thermogenesis, physical activity, and decreased food intake. In contrast, MC4R knockout mice increase food intake, thus associating MC4R signaling with appetite control [86,87].

In addition, the agouti protein was shown to be upstream of MC1R and acts as an MC1R antagonist resulting in a brighter skin pigmentation [67]. Hypothalamic cells do not express agouti protein, but it acts as an antagonist on MC4R [88]. However, AgRP, in contrast to the agouti protein, is expressed endogenously in the brain and acts as a natural antagonist of MC3R and MC4R [68]. Furthermore, AgRP is reported as both an MC4R inverse agonist [89,90] and an antagonist (*e.g.*, [91,92]).

The melanocortin pathway is one of the critical effectors of leptin, a hormone secreted by adipocytes in relation to the amount of fat tissue and signals the brain about long-term energy storage and starvation [93]. Mutations in leptin [94] and the leptin receptor (LepR) [95] result as well in severe obesity.

POMC is expressed in the pituitary, skin, and hypothalamic neurons [66]. The production of melanocortin peptides from POMC requires the action of several proteases. The first two steps are catalyzed by proconvertase 1 (PC1) and proconvertase 2 (PC2), followed by carboxypeptidase E (CPE). Mutations in the PCSK1 gene encoding PC1 [96] and in POMC [97] were found in obese patients.

MC4R is additionally regulated by its adaptor protein, the melanocortin receptor accessory protein 2 (MRAP2), and its loss of function is also associated with obesity [98,99].

The only prominent player of the hypothalamic melanocortin pathway that does not result in obesity upon genetic disruption is MC3R. This receptor is thought to regulate the balance of energy expenditure and fat storage [100].

1.2.4 MC4R regulation on a tissue level

These studies unraveled an intricate system of MC4R regulation in the paraventricular nucleus of the hypothalamus (Figure 1.2.4). Various peripheral hormones, such as GLP-1, peptide YY (PYY), ghrelin, and leptin, address a set of two neurons in the arcuate nucleus of the hypothalamus. POMC-expressing neurons are activated, leading to the expression of MSH. α-MSH and β-MSH activate MC4R that is expressed in the paraventricular nucleus

of the hypothalamus. MC4R regulates energy expenditure, blood pressure, heart rate, sexual drive, and growth.

AgRP expressing neurons release AgRP that downregulates MC4R. Ghrelin activates, and leptin inhibits the secretion of AgRP. This counterplay, of endogenous agonists and an antagonist (or inverse agonist), results in the fine-tuned mediation of MC4R signaling. In recent years the MC4R agonist setmelanotide successfully underwent clinical trials that addressed patients with upstream defects, namely PCSK1 loss of function, POMC deficiency [101], and LepR deficiency [102]. These studies highlight the central role of MC4R in the melanocortin pathway and thus as a regulator of human satiety. Furthermore, the signaling profile of setmelanotide has been connected to its successful application in treating rare-genetic obesity [102].

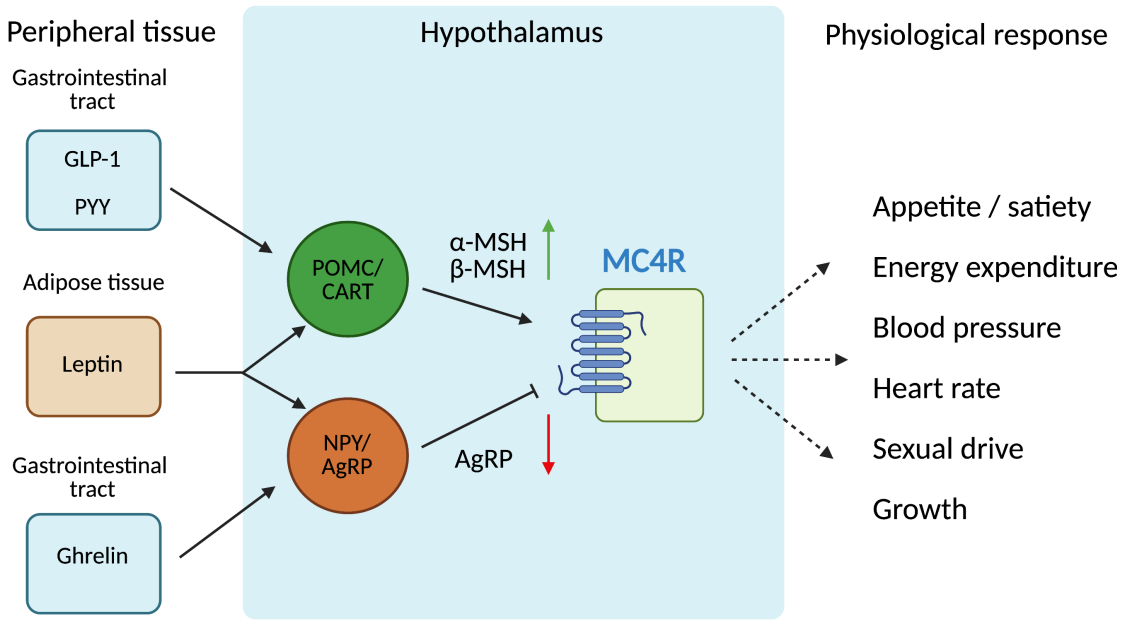

Figure 1.2.4 ***A set of two neurons regulates MC4R and various physiological responses.*** Schematic representation of the melanocortin pathway, emphasizing the up- and downregulation of MC4R by α-,β-MSH, and AgRP, respectively. Figure was generated using biorender.com.

1.2.5 MC4R signaling regulation at the cellular level

The up and downregulation of MC4R's basal signaling by endogenous agonists and the inverse agonist AgRP is further discriminated at the cellular level. MC4R is a promiscuous GPCR coupler; the activation of G_s, $G_{q/11}$, and G_i have been reported.

Historically, MC4R signaling was mainly investigated in heterologous transfected cell systems. Here, the G_s/cAMP signaling is the most prominent signaling pathway. MC4R activation of G_s was directly associated with controlling energy balance, thermogenesis, and peripheral glucose metabolism [103]. MC4R induced G_s signaling results in the PKA-dependent phosphorylation of cAMP-responsive element-binding protein (CREB)[104]. The genetic disruption of CREP in PVN neurons of mice caused obesity due to

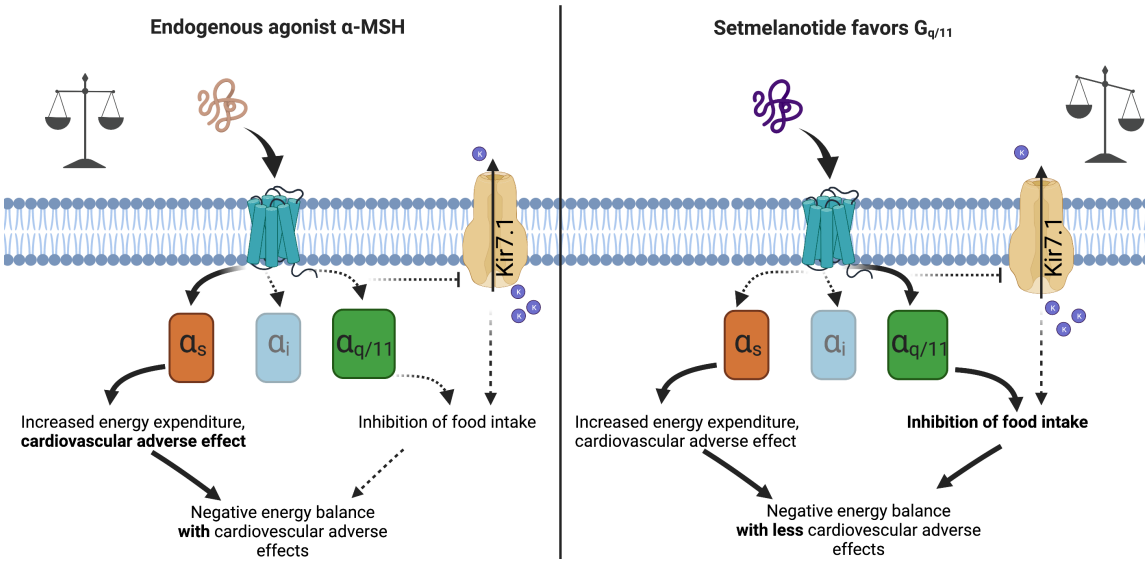

Figure 1.2.5 ***MC4R is a promiscuous receptor that signals via G_s, G_i, and $G_{q/11}$. G_s signaling is associated with cardiovascular adverse effects and $G_{q/11}$ signaling with the inhibition of food intake.*** MC4R signals promiscuously via G_s, G_i, and $G_{q/11}$. α-MSH, the endogenous agonist, is defined as a balanced ligand that activates G_s 50-fold more than $G_{q/11}$ signaling [102]. In contrast, a biased ligand, such as setmelanotide, induces the $G_{q/11}$ pathway in the same fashion as the G_s pathway [102]. Cardiovascular adverse effects are associated with G_s signaling and appetite reduction with $G_{q/11}$ and Kir7.1 signaling. MC4R inhibits Kir7.1 and increases intracellular potassium levels. On the contrary, a balanced ligand displays increased cardiovascular adverse effects as a $G_{q/1}$ biased ligand. Figure adapted from [106] and created with biorender.com.

reduced energy expenditure but did not result in decreased food intake [105]. This data indicates that G_s signaling is not regulating MC4R driven appetite control.

MC4R was connected with G_i [107], arrestin [108], and ERK [109] signaling but not directly related to appetite regulation. Thus, comprehensive investigations of these pathways are elusive. Furthermore, MC4R inhibits the inwardly rectifying potassium channel Kir7.1. Agonist-dependent MC4R stimulation increases, and AgRP-dependent MC4R inactivation decreases the frequency of the Kir7.1. dependent-firing potential of PVN neurons [110]. The addition of PKA inhibitors did not affect the Kir7.1 dependent depolarization. This observation indicates a G_s independent mechanism of regulation of Kir7.1 activity by MC4R. Furthermore, the targeted disruption of Kir7.1 in PVN neurons of mice causes an obese phenotype [111].

A factor that regulates cytoplasmic Ca^{2+} level is $G_{q/11}$ signaling. First indications of increased Ca^{2+} levels after α-MSH stimulation were reported in murine hypothalamic cells expressing MC4R [112]. This observation was later connected to $G_{q/11}$ family (members are G_q, G_{11}, G_{14}, and G_{15}) signaling via the measurement of phospholipase C activation [113]. The local knockout of G_q and G_{11} in PVN neurons confirmed this assumption, and the knockout mice were obese even though G_s and other signaling pathways were not disrupted [114].

The hypothesis that appetite is regulated via MC4R induced PLC signaling is strengthened by the pharmacological profile of setmelanotide (Figure 1.2.5). Compared to α-MSH, this peptidic agonist has a PLC-biased signaling profile [102] and is the first FDA-approved anti-obesity drug targeting MC4R. Clement *et al.* [102] measured PLC activation via the NFAT reporter gene assay, which does not discriminate which $G_{q/11}$ family member drives PLC activation. BRET measurements using carboxy-terminal modified $G\alpha_q$ subunits, which imitate 11 G-protein subtypes, indicate that MC4R couples to G_q, G_{14} but not to G_{15} [115]. In contrast, a recent study that uses the entire length Gα subunits in a BRET assay contrasts these findings with the activation of only G_{15} and not G_q [116]. This study validated $G_{q/11}$ coupling via the $G_{q/11}$ binding domain of the effector p63 Rho guanine nucleotide exchange factor to validate $G_{q/11}$ members' coupling [116]. Hence, *in cell* G-protein coupling assays produce contradicting results, and the cognate G-protein subunit from the $G_{q/11}$ pathway needs to be identified.

All these findings point toward a highly complex regulation of MC4R signaling that includes various pathways in addition to G_s activation. $G_{q/11}$ induced PLC signaling and Kir7.1 are involved in appetite regulation, while G_s signaling relates to energy expenditure and cardiovascular effects.

Further factors potentially mediating MC4R signaling are MC4R dimerization, oligomerization, and the accessory protein MRAP2. However, the discussion of these exceeds the scope of this thesis.

1.3 MC4R ligands and pharmacology

MC4R is down-regulated physiologically by the endogenous antagonist [92,117] or inverse agonist [89] AgRP and is activated by the cleavage products of POMC: α-MSH, β-MSH, γ-MSH and, ACTH. Of them, α-MSH and β-MSH are physiologically relevant in the hypothalamus. These have as well the highest binding potencies (β–MSH> α–MSH >> ACTH > γ–MSH [footnote 4]). In addition, MC4R agonists share the common MC4R recognition motif, HFRW, essential for high-affinity peptide agonist binding. We introduced the standard $H^0x^1R^2W^3$ ligand peptide numbering system (Figure 1.3.1) to simplify the comparison of MC4R agonists, which is indicated as the superscript of the residue (such as for setmelanotide $R1^{-3}$ resembles the first residue of setmelanotide that precedes $H4^0$ with three residues) [118].

Pharmacological trials, except for setmelanotide, failed due to adverse effects partly driven by the high conservation of the MCRs and the above-discussed G-protein specificity. MC1R, MC3R, and MC4R are activated by α-MSH, β-MSH, γ-MSH, and ACTH.

Setmelanotide is 33-fold more potent in activating MC4R than MC1R and 200-fold more potent than MC3R [119] and is therefore subtype-selective. However, the high ligand-binding pocket sequence similarity of MC1R and MC4R; and the requirement of ligands with an HFRW motif for MC1R and MC4R activation complicate subtype-selective drug development.

MC4R primarily targets the G_s pathway, and all reported agonists induce a G_s response. Drug potencies of MC4R stimulated PLC activation indicate that $G_{q/11}$ activation does not

[4] www.guidetopharmacology.org; 15.02.2022

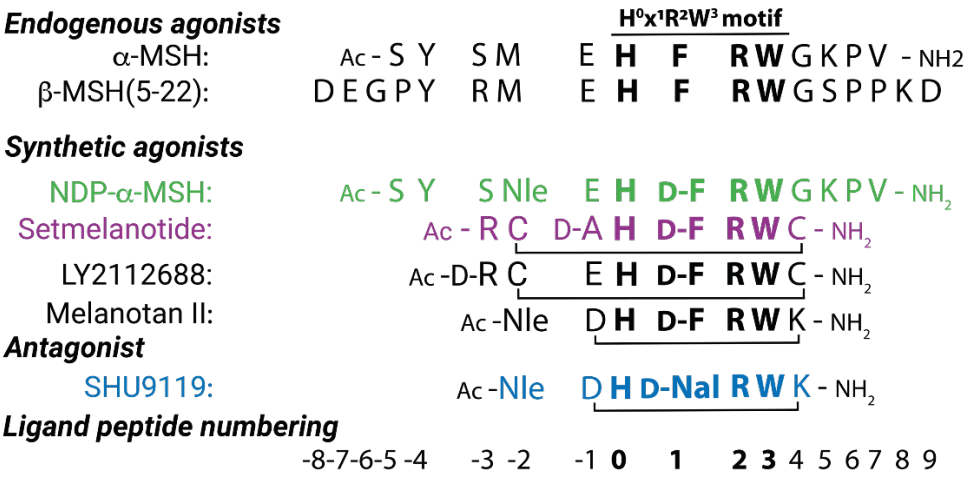

Figure 1.3.1 ***Sequence alignment of MC4R agonists and antagonist SHU9119.*** Based on the MSH *HxRW* motif a ligand unifying numbering system is introduced for comparison and simplification and indicated for each peptide residue in superscript without spacing. A superscript integer with spacing indicates for citations. Abbreviations of non-natural amino-acids are Nal, naphtylalanine; Nle, norleucine. Amino-terminal acetylations are abbreviated by Ac and carboxy-terminal amination as NH2.

exceed the potency for G_s activation (Table 1.3.1). In the context of MC4R drugs, a $G_{q/11}$-biased drug is referenced to the signaling profile of α-MSH. Hence, a drug that induces $G_{q/11}$ and G_s with the same potency is a $G_{q/11}$-biased drug (*e.g.*, setmelanotide). A drug with a more pronounced G_s signaling profile will be referenced in the following as a G_s-balanced drug (*e.g.*, α-MSH).

Table 1.3.1 ***Binding affinities and functional pharmacology of melanocortin ligands relevant for this study***. Superscripts indicate the corresponding literature

ligand	α-MSH	β-MSH	NDP-α-MSH	setmelanotide	LY2112688
	[nM]	[nM]	[nM]	[nM]	[nM]
Ki	51 [91]	20 [120]	1 [91]	2 [121]	4 [122]
EC_{50}(cAMP)	10 [123]; 13 [102]	33 [123]	4 [123]	7 [102]	14 [102]
EC_{50}(NFAT)	489 [123]; 174 [102]	107 [123]	30 [123]	10 [102]	330 [102]

1.3.1 Endogenous ligands

The mature form of α-MSH is acetylated at the amino terminus and aminated at the carboxy terminus [66]. α-MSH induces 17 to 49 times more potent G_s over $G_{q/11}$ signaling and has an affinity of 51 nM (Table 1.3.1). The non-acetylated variant is termed des-acetyl α-MSH. An early study in 1981 showed that acetylation enhances the biological activity of the peptide, and it was hypothesized that this is driven solely by the increased biological stability [124]. However, the impact of acetylation on different signaling pathways has not been elucidated since then.

β-MSH is not acetylated or aminated, and two isoforms act on hypothalamic cells, both the entire length peptide with 22 residues and the more potent shorter version of β-MSH with residues 5-22 [66]. Initial MC4R experiments were conducted mainly in mice that lack β-MSH, and it was thought that only α-MSH played a role in human energy balance, but it was shown later that β-MSH is at least as important α-MSH in humans [125]. β-MSH induces three times more potent G_s over $G_{q/11}$ signaling with an affinity of 20 nM (Table 1.2.5). This subtle $G_{q/11}$ bias has not been thoroughly investigated.

1.3.2 Synthetic MC4R Ligands

This thesis focused on understanding the ligand binding modes of NDP-α-MSH and setmelanotide. In addition, the antagonized SHU9119-MC4R structure will be used in the discussion to delineate the MC4R activation mechanism. Therefore, these three synthetic ligands will be introduced here.

1.3.2.1 NDP-α-MSH

NDP-α-MSH, or afamelanotide, is a high-affinity variant of α-MSH with two single side-chain substitutions. The substitution of L-$F7^1$ to the stereoisomer D-$F7^1$ is the primary driver of the increased affinity [73]. The second substitution is $M4^{-2}$Nle. The introduction of Nle, norleucine, or 2-aminohexanoic acid, makes the peptide resistant to inactivation by chloramine-T, which is necessary for radioactive assays by iodonization [126]. NDP-α-MSH

was optimized for MC1R activation and is an MC1R selective drug that addresses MC4R with an affinity of 1 nM (Table 1.3.1). NDP-α-MSH (branded as Scenesse) was FDA approved in 2019 to treat MC1R-driven melanogenesis. The drug was approved to prevent skin damage from phototoxicity in patients suffering from erythropoietic protoporphyria. Non-severe adverse effects of NDP-α-MSH administration include back pain, nasopharyngitis, or headache [127] but not appetite reduction or cardiovascular side effects. NDP-α-MSH $G_{q/11}$ signaling is in a similar potency range as setmelanotide (Table 1.3.1).

1.3.2.2 Setmelanotide

Setmelanotide is an eight residue-long cyclic peptide. It shares with NDP-α-MSH solely the H^0D-$F^1R^2W^3$ core motif. The precursor of setmelanotide was designed in 2005 with the sequence Ac-RC*EHD-FRWC*-NH2 [119]. This approach combined previous cyclization attempts [128,129] and used β-MSH as the template to increase MC4R subtype selectivity. Setmelanotide originates from this peptide by the $E3^{-1}$D-A substitution. Ultimately, the only reminiscent of β-MSH outside the HFRW motif remained $R1^{-3}$. The peptide is cyclized by a disulfide bond of $C2^{-2}$ and $C8^4$ and has a 2.1 nM binding affinity to MC4R with a 33-fold selectivity over MC1R for the induction of G_s signaling [130]. Setmelanotide has been FDA approved for treating three forms of rare genetic obesity since November 2020 and is the first MC4R targeting anti-obesity drug (brand name: *Imcivree*). The approval includes patients with POMC deficiency, PCSK1 deficiency, a protease cleaving POMC, or LepR deficiency (FDA reference ID: 4707991). Setmelanotide does not induce - in contrast to other MC4R agonists - G_s-related adverse effects like tachycardia or high blood pressure. However, non-severe adverse effects have been reported related to MCR subtype selectivity, such as skin hyperpigmentation [66]. A $G_{q/11}$ biased signaling profile could drive the reduced cardiovascular adverse effects. Setmelanotides biased signaling profile has been shown by addressing PLC measured via an NFAT reporter assay in HEK293 cells [102]. The NFAT reporter assay is a downstream surrogate for PLC activation via $G\alpha_{q/11}$ and $G\alpha_{i/o}$ signaling, and the addition of pertussis toxin does not inhibit the response upon setmelanotide stimulation[102]. Hence, a $G_{q/11}$ bias independent of $G_{i/o}$ signaling of setmelanotide has been shown. This observation aligns with mice experiments directly linking feeding behavior and $G_{q/11}$ signaling [131].

1.3.2.3 SHU9119

SHU9119 is a potent MC4R antagonist with a K_i of 2.9 nM [132] that acts as an MC1R and MC5R agonist and an MC3R partial agonist [133]. SHU9119, a heptameric peptide, is circularized by a lactam ring formed by $D2^{-1}$ and $K7^{4}$ (Figure 1.3.1). The sequence is based on NDP-α-MSH [133], and introducing the D-napthylalanine (D-Nal) moiety instead of $D\text{-}F^{1}$ drives the antagonistic effect. Melanotan II (MTII), a potent MC4R agonist, shares the identical sequence, including the lactam circularization as SHU9119 except for the $D\text{-}F^{1}$D-Nal substitution. The molecular mechanism of D-Nal preventing MC4R activation became not apparent from the SHU9119-MC4R structure [91], and the comparison to an active state structure can shed light on this question.

1.3.3 Ca^{2+} ion - a cofactor for ligand binding

The dependence of melanocortin binding on Ca^{2+} ions (Ca^{2+}) was shown in the early 1980s and was first documented for ACTH binding in 1953 [134]. Thus, the binding regulatory property of this ion has become a paradigm for labs in this field. Nevertheless, the molecular mechanism and the site for Ca^{2+} binding was an open. The role of Ca^{2+} as a potential co-factor of melanocortin binding was shown in 1981 by the Hruby laboratory [135]. Here, the addition of NDP-α-MSH to frog skin induced prolonged skin darkening. In the absence of NDP-α-MSH, frog skins previously darkened were lightened by removing Ca^{2+} from the medium. The skin could be darkened again by the addition of Ca^{2+}. This cycle could be repeated indefinitely. This experiment showed impressively that melanocortin binding to MC1R is Ca^{2+} dependent before the MCRs were cloned. The Rinken lab developed the first radioligands independent binding assay for MC4R. Their fluorescence polarization assay using fluorescently labeled MC4R ligands and MC4R expressing budded baculovirus particles deciphered the Ca^{2+} dependency under in vitro conditions [136]. The antagonist-bound crystal structure of MC4R in 2020 displayed a Ca^{2+} ion in the ligand-binding pocket of MC4R coordinated by three acidic residues at TM2, TM3, and the backbone of the peptidic antagonist SHU9119. Active state structures and subsequent studies can potentially unravel if Ca^{2+} is a permanent co-factor or is part of the complex signaling regulation of MC4R.

1.3.4 Pharmacological description of ligand binding and signaling

The interaction of a ligand with a receptor is either investigated using a direct-ligand binding experiment (*e.g.*, radioligand assays, BRET assays) or via studying downstream signaling (*e.g.*, cAMP assay) induced by receptor activation.

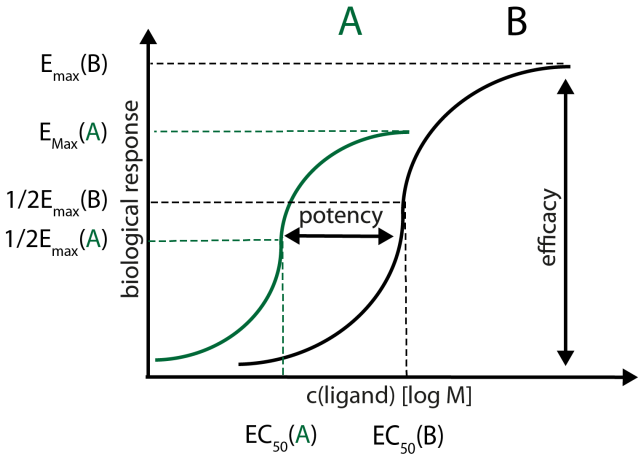

Figure 1.3.4 Potency and efficacy. Potency refers to the amount of ligand needed to produce the half-maximum biological response (EC_{50}). Efficacy describes the maximum ligand response. The biological response is plotted against concentration in a graph to give the concentration-response curve. The maximum response is referred to as the E_{max}.

The direct measurement of ligand binding in a one-site binding event in equilibrium between a ligand L and a receptor R can be described as a kinetic and thermodynamic process. A Hill coefficient of one defines a one-site binding event [137].

$$R + L = RL$$

This reaction is characterized in equilibrium by a constant on-rate k_{on} and a constant off-rate k_{off}.

$$k_{on}[R]\,[L] = k_{off}[LR]$$

The dissociation constant K_D is defined:

$$K_D = \frac{k_{on}}{k_{off}} = \frac{[L][R]}{[LR]}$$

Ligand binding affinities are influenced by non-covalent interactions between the receptor and the ligand. These interactions between the receptor and ligand can also be measured in the change of Gibbs free energy ΔG.

$$K_D = e^{\frac{\Delta G}{RT}}$$

Hence, a direct ligand-binding assay entails details over a ligand-binding event's kinetic and thermodynamic nature.

The investigation of downstream signaling pathways is not a direct measurement of ligand binding. It entails information of the kinetics of the ligand-binding event, but these are not direct measurements. The strength of a ligand inducing a signaling response is measured as potency and defined as the half-maximal effective concentration (EC_{50}). Potency correlates with affinity in the case of a one-site binding event in equilibrium [138]. Hence, if a Hill coefficient determined in concentration-response experiments equals one, a change in potency correlates with affinity [138]. Consequently, residues in mutagenesis studies that display a difference in potency compared to WT protein impact the molecular interactions of the ligand with the receptor. If such a residue is not in the vicinity of the ligand-binding pocket, a change in potency indicates an allosteric coupling of the ligand with this residue. In contrast, residues that are not relevant for ligand binding but are relevant for the active state display a reduced efficacy without changes in potency (Figure 1.3.4). The efficacy describes the maximum response that can be achieved with a ligand and is measured as E_{max}.

1.4 Objectives of this study

MC4R, the central switch in the leptin-melanocortin pathway, is to date the most promising candidate for the treatment of genetically driven obesity. The primary objective of this thesis is elucidating the structures of the active MC4R state - the activation of MC4R results in signaling responsible for appetite reduction. Hence, active-state MC4R structures can foster the development of novel anti-obesity drugs.

Structures of agonist-MC4R-Gs-protein complexes in combination with mutagenesis signaling allow investigation of the main objectives of this thesis:

i. The characterization of the peptide-binding modes of NDP-α-MSH and setmelanotide to MC4R allows deciphering residues of relevance for MC4R activation and agonist binding.

ii. Unraveling the mechanism of MC4R activation, by comparing the active state MC4R-G_s complex structures with the antagonist-bound MC4R structure [91], indicate the amino acids involved in receptor activation.

iii. The MC4R-G_s-protein interface could indicate MC4R-specific sites involved in G-protein coupling. The structures could identify the molecular determinants of G_s-protein coupling to MC4R. Further, comparing MC4R complexes bound with agonists with deviating pharmacological profiles could reveal receptor areas responsible for driving G-protein specificity.

The determination of active-state MC4R structures requires establishing the following experimental prerequisites:

1. Establishing the methods and tools required for the structural investigation of non-visual GPCRs in the lab of Patrick Scheerer. This includes the heterologous production of proteins required for a receptor-protein complex (MC4R, hetero-trimeric G_s protein, and Nb35), establishing a complex formation, and a rapid complex purification protocol.

2. Establishing a ligand-binding experiment to verify the ligands of interest and lay the foundation for developing improved MC4R agonists.

3. Preparation of a stable agonist-MC4R-G_s-protein complex in sufficient amounts that sustains vitrification for single-particle cryo-EM experiments.

1.5 Methodologic introduction

An active state GPCR structure requires three main components, a potent agonist, a functional GPCR in sufficient amounts, and a stable G-protein or G-protein mimetic. Therefore, a laboratory needs to establish a stable ligand binding assay and a heterologous expression system for the receptor and the G-protein. Three years of this doctorate were focused on establishing novel protocols and infrastructure to characterize the MC4R at a structural level. The agonist-MC4R-G_s complex structures were determined by single-particle cryo-EM, which will be described in the last section of this chapter.

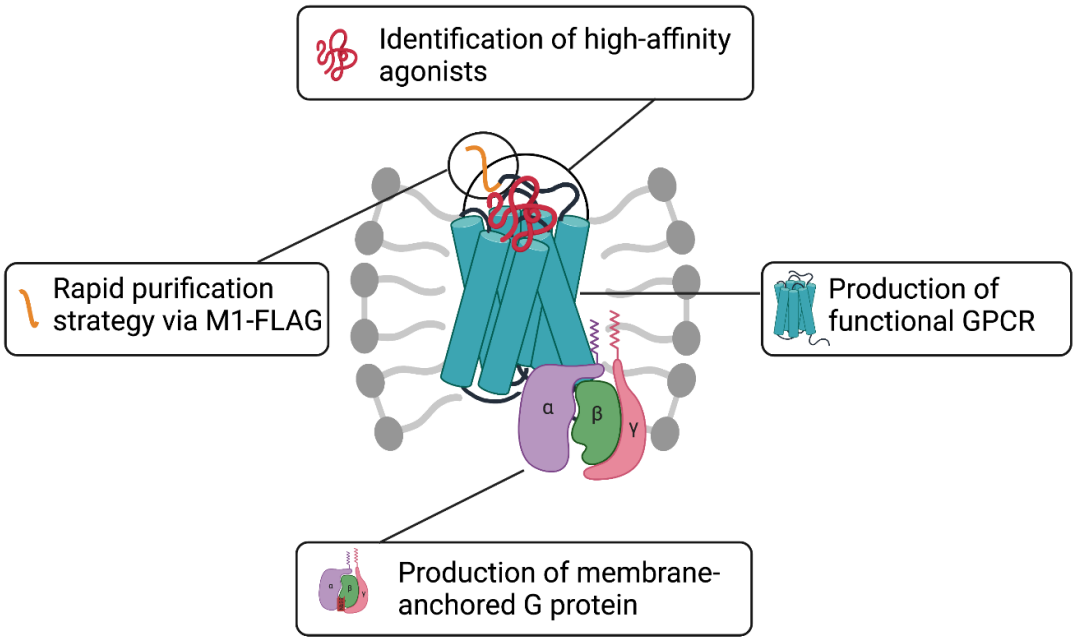

Figure 1.5 ***Experimental prerequisitess for an agonist-GPCR-G-protein complex formation experiment.*** A simplified figure was created using biorender.com.

1.5.1 GPCR ligand-binding assays

GPCRs have a variety of sub-states that can be stabilized by ligands (Section 1.1.3). Agonists stabilize active states, and highly basal active receptors can form GPCR-G-protein complexes without agonists. However, stabilizing a class A GPCR-G-protein sample preparation in a homogeneous active state without an orthosteric agonist has been only fea-

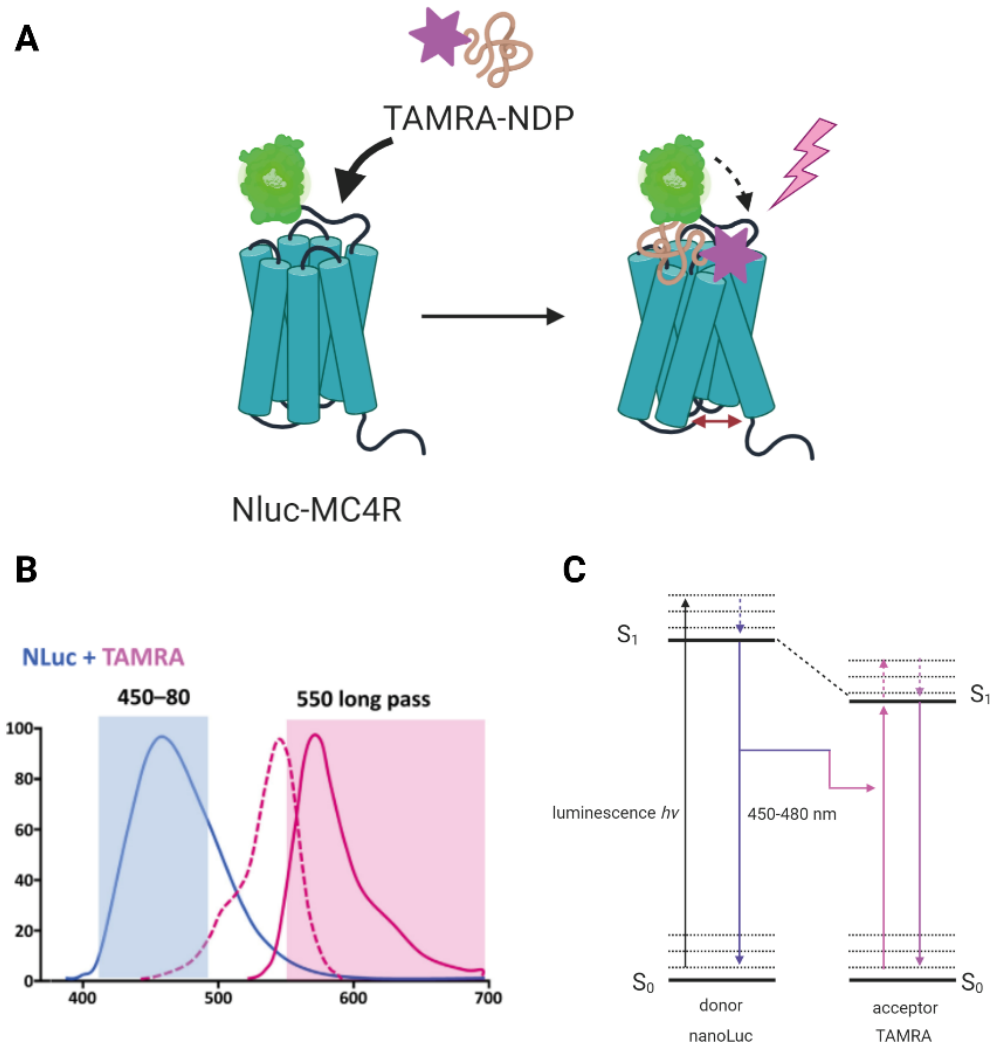

Figure 1.5.1 ***NanoLuciferase (nLuc) based ligand binding assay.*** (**A**) NLuc is fused at the amino-terminus of the GPCR of interest, and a fluorescent (here TAMRA) labeled ligand binds to the orthosteric ligand-binding pocket induces a bioluminescence energy transfer (BRET) effect. (**B**) The overlap of nLuc and TAMRA spectra induces the energy transfer. (**C**) The Jablonski diagram displays the donor's (nLuc) coupling and resonance energy transfer to the acceptor (TAMRA). The BRET ratio is calculated as the quotient of the energy transfer (long pass) by the nLuc luminescence (short pass). The figure was created using biorender.com

sible for the 5HTR1-Gi [139] complex. Hence, a potent agonist and a ligand-binding assay is the first pillar for a structural campaign. To highlight the relevance of ligands in GPCR research, the β2-adrenergic receptor, the model system for ligand binding GPCRs, was purified [140] and cloned [30,141] using the β-blocker alprenolol. Ligand-binding assays utilize in general isotope-labeled ligands (radioligands). Tritiated small molecules were used to assess the amount of functional receptor expression in living cells and ligand properties. Peptidic ligands usually require labeling with the hazardous iodine isotope I-125 (I^{125}). Its use makes experiments difficult, dangerous, and expensive. The nano-luciferase-based ligand binding assays invented by Stoddart and colleagues [142] allowed to assess ligand binding properties of peptidic ligands in living cells without the need of using radiolabeled ligands (Figure 1.5.1).

The method, in brief, a GPCR construct is required at which a nano-luciferase (nLuc) is fused at the amino terminus of the receptor (Figure 1.5.1A). Then, the nLuc-GPCR construct is transfected into HEK293T cells, and a fluorescent-labeled ligand is added. Ligand binding to the GPCR brings the fluorescent-labeled ligand in proximity to the fused nLuc and induces a bioluminescence energy transfer, given a suitable overlap of the nLuc and fluorophore spectra (Figure 1.5.1B). The quotient of the long pass (the energy transfer) by the short pass (the nLuc chemiluminescence) is called the BRET ratio.

$$BRET\ ratio = \frac{long\ pass\ [550nm]}{short\ pass\ [480nm]}$$

The titration of the fluorescent-labeled ligand allows the determination of its dissociation constant (K_D). The increase in the BRET ratio correlates with the amount of fluorescently labeled ligand bound to nLuc-GPCR. Non-labeled ligands can be characterized via the inhibitory constant (K_i) by titrating the non-labeled ligand against a constant concentration of the fluorescent-labeled ligand [A] (in the range of 1-10-fold its K_D) [143] using the Cheng-Prussof equation.

$$K_i = \frac{IC_{50}}{\frac{[A]}{K_d} + 1}$$

The inhibitory concentration IC_{50} is defined as the half-maximal inhibitor concentration that reduces the response (here, the BRET effect) halfway between the baseline and maximum response.

A nLuc based ligand-binding assay can be harnessed to assess (*i)* expression levels of GPCR constructs, (*ii*) the identification of high-affinity agonists, and (*iii*) the functionality of receptor constructs:

i. The expression level of nLuc-GPCR constructs (such as point mutations or truncations) in HEK293 cells is assessed by comparing the level of chemiluminescence to a wild-type (WT) nluc-GPCR construct.

ii. Concentration-response measurements of labeled ligands or titration experiments of non-labeled ligands against labeled ligands allow the identification of high-affinity agonists.

iii. Receptor mutants can impair receptor folding leading to impaired ligand binding. Such effects become apparent by changes in ligand affinity.

Ligands and ligand binding assays have been the pillar of the biochemical and structural investigation of GPCRs. However, GPCRs without known ligands (orphan GPCRs) are hard to stabilize in a defined state and therefore not easily accessible by crystallography and cryo-EM.

1.5.2 Heterologous expression of functional receptor

The structural investigation of integral membrane proteins like GPCRs requires microgram to milligram quantities of purified protein. Therefore, the initial and most crucial step is the expression of properly folded and functional receptors. Despite tremendous advances in all aspects of biochemical research, protein expression is still based on trial and error. Historically, the expression was optimized for crystallization constructs. Here, extensive construct optimizations leading to the stabilization of the required state were performed. The investigation of GPCR-G-protein complexes by cryo-EM reduced the complexity of the construct design because the optimization focuses solely on the expression of functional receptors and not on the formation of protein crystals. Furthermore, cryo-EM requires

smaller amounts of protein, and thereby the investigation of low expressing GPCRs is possible.

Certain post-translational modifications (PTMs) are critical for expressing one GPCR but may not be essential for the other one. For example, some GPCRs require glycosylation for high-affinity ligand binding and G-protein binding (*e.g.*, somatostatin receptor, rhodopsin, and β_2AR) [144]. Other receptors require glycosylation for the proper membrane trafficking (*e.g.*, A1AR, H2R, BB1, and AT2R) [144]. PTMs, like glycosylation, cannot be deduced from the amino acid sequence alone; therefore, either different cell systems are tested, or higher eukaryotic systems are the preferred choice. Furthermore, the lipid composition of the membrane can impact the receptors. For instance, the active state μ-opioid receptor is stabilized in the presence of cholesterol but an inactive state in the presence of ergosterol [145]. In the following different cells systems for GPCR expression are presented.

1.5.2.1 E.coli

Besides its prokaryotic nature, *E. coli* harbors several advantages over other cultures: it is easy to handle, genetic modifications are quickly introduced, and if the receptor is expressed, large amounts of the receptor can be produced at low costs. It is the preferred expression system for isotopic labeling and thus NMR studies. Rapid expressed protein can accumulate in inclusion bodies and receptors can be refolded *in vitro* for NMR studie [146]. The absence of native G-proteins or other GPCR interacting G-proteins hampers the production of functional receptors. The lack of PTMs results in a homogeneous product but limits receptor functionality and membrane trafficking. Furthermore, the lipid composition in *E. coli* strongly deviates from more complex mammalian membranes, and cholesterol cannot be produced by *E.coli* which acts as an allosteric regulator for specific GPCRs. Functional receptor expression in *E.coli* has been achieved for NTSR1 and CB2R, but is only suitable for a small array of receptors.

1.5.2.2 Yeast – Pichia pastoris

P.pastoris is the favored yeast system for the overexpression of GPCRs for structural studies and has yielded two crystallography structures, namely histamine 1 receptor [147] and

α2-adrenergic receptor [148]. The expression system is easy to handle, scale, and manipulate. Yeast cells can facilitate the formation of disulfide bonds, and N and O linked glycosylation. On the downside, yeast cells have the potential of hyper-mannosylation and to glycosylate non-endogenous glycosylation sites. Furthermore, they have instead of cholesterol ergosterol. Yeast cell lines are mainly used for directed evolution methodologies [149,150] to increase GPCR stability but are not an ideal host for GPCR expression.

1.5.2.3 HEK cell lines

From a functional perspective mammalian cell systems are the expression system of choice. PTMs, membrane composition, and abundance of GPCR interacting G-proteins are close to endogenous conditions. Human embryonic kidney (HEK) 293T cells, the standard mammalian cell line, are adherent cells that express GPCRs well, but the available surface area limits cell growth. Thus, HEK cells were not used widely for structural studies. In 2018, the Expi293F system (Thermo Fisher) was adapted for GPCRs by the Lefkowitz lab and coworkers [151] and allowed the structure determination of the angiotensin 2 receptor [152]. The Expi293F cells are derived from the HEK293 cell line but are not adherent. Therefore, the cells are cultivated in flasks, allow the expression of large amounts of GPCRs in mammalian cells, and have the potential to become the preferred expression system in the GPCR field. In addition, HEK cells are the preferred system for cell-based assays assessing ligand binding and cell signaling.

1.5.2.4 The baculovirus insect cell system

The baculovirus insect cell system has been the system of choice for the expression of GPCRs developed over the last years and has been used for this thesis. Baculoviruses have been exploited for heterologous protein expression since 1983 [153]. Insect cell lines grow in solution and are suitable for large scale expressions in contrast to adherent cells. The baculovirus insect cell system is the dominant expression system in GPCR structural biology, and this success is based on the scalability, a high similarity with mammalian PTMs, and that GPCRs are expressed mainly as functional receptors [144].

The expression of the target protein is achieved by the infection of a recombinant *Autographica californica* virus, a multiple nuclear polyhedrosis virus. The polyhedrin gene, a redundant but highly expressed late virus gene, is replaced by the gene of interest. Baculoviruses target insect cells only and are therefore safe for humans. Most PTMs that are found in mammalian cells is carried out by the insect cell background [154].

A range of different systems has been developed to generate recombinant viruses [155]. Either they can be generated by transposition in *E.coli* (Bac2Bac system) or by heterologous recombination in insect cells (flashback, UltraBac system). Commercial baculoviruses are available, but adapted bacmids can be produced in-house with comparable protein expression yields [156].

The first virus generation (P0) requires two amplification steps (P2) to acquire sufficient multiplicity for protein expression. The generation of highly potent viruses is difficult to achieve because assessing the quantity of the virus does not correlate with the potency of the generated viruses [155]. The most accurate determination of suitable viruses lies in monitoring the cell density, viability, and if possible, the fluorescence of an eGFP fusion construct. Virus generation of P1 and P2 requires approximately 72 hours, the cell count should not exceed two-fold, and the viability should be in a range of 60-90% (according to the Kobilka lab [footnote 5]). P1 viruses can be stored at -80°C, the generation of one good P1virus can last for several years of experiments. P2 viruses are used for protein expression, and the expression time varies between 24 to 52 hours. The optimal expression time can be assessed by comparing the yields of test purifications. Here, the cell count should not exceed two-fold upon infection, and the viability should be higher than 90% (according to the Kobilka lab [footnote 5]). Baculoviruses can be used to infect a variety of different cell lines. The most common GPCR expression systems are *Sf*9 cells and *Trichoplusia ni* (*Tni*) cells. The latter is used preferentially to express membrane attached proteins *e.g.,* heterotrimeric G-proteins [157] or soluble proteins *e.g.*, scFv16 [158].

Insect cells have a different lipid composition than mammalian cells. They have low levels of cholesterol, have very high phosphatidylinositol content, and no phosphatidylserine [144]. Depending on the target protein, there have been reports of protein misfolding [159]. Using too potent virus can induce early cell lysis, protease release, protein degradation, or

[5] The baculovirus *Sf*9 protocol was adapted from the Kobilka lab, Stanford, and optimized by us for our target GPCRs. Here special thanks to Daniel Hilger, Betsy White, Tong Sun Kobilka, and Brian Kobilka for their focus on detail.

removalof amino terminal tags essential for receptor purification [160]. Up to date, more than 90% of all published GPCR structures have utilized the baculovirus insect cell system footnote 6.

1.5.3 Agonist GPCR-G-protein complexes

The pharmacological relevance of GPCRs cannot be overstated, and 45% of all approved GPCR targeting drugs are agonists [4]. In addition, agonist-bound GPCR structures allow to unravel the agonist binding modes and thereby the molecular mechanism of receptor activation.

The interpretation of agonist binding modes demands active state GPCR structures, and the structural elucidation of active state receptors was initiated in the rhodopsin system in 2008 [41]. Receptors can be stabilized in an active conformation by agonists, thermostabilizing mutations (*e.g.*, [161]), pH [41], the addition of G-protein mimicking nanobodies [19,34], engineered G-proteins (*e.g.*, [162]), and heterotrimeric G-proteins.

The Nobel prize structure of Brian Kobilka of the β2-adrenergic receptor (β2AR) in complex with the nucleotide-free heterotrimeric G_s in 2011, stabilized by nanobody 35 (Nb35) [20], is the hallmark of receptor complex research of the last twenty years. Until June 2021 [163], no other complex structure with a heterotrimeric G-protein has been solved by crystallography, and it took six years, till 2017, the year in which this thesis was started, until the cryo-EM revolution [22] hit the GPCR field with the GLP1-G_s cryo-EM structure [164]. The underlying principles of sample preparation for G_s coupled complex preparations have not changed, but each receptor requires an individual adaption. The most critical and demanding steps are the expression and purification of the receptor and the heterotrimeric G-protein in insect cells (Section 1.1.4). In addition, the cognate receptor G-protein pair needs to be identified. Here, published data can assist [115] but ultimately, *in vitro* complexation experiments with different agonists and G-proteins need to be performed.

Furthermore, the complexation strategy needs to be adapted for each receptor G-protein pair. The complex is formed by adding a GDP-bound heterotrimeric G-protein to a receptor embedded in the *Sf*9 insect cell membrane (*e.g.*, [164]) or detergent purified receptor (*e.g.*, [20]).

[6] www.GPCRdb.org/structures; 15.07.2021

The addition of an agonist in 100-fold to 1000-fold excess of its K_D stabilizes an active conformation of the receptor and favors G-protein binding. The receptor induces GDP release, and GTP binding would initiate complex dissociation under physiological conditions. Under *in vitro* conditions, GDP is removed from the equilibrium by adding apyrase, a nucleotide hydrolase that hydrolysis GDP to GMP. The resulting nucleotide-free GPCR-G-protein complex is trapped in a stable low energy state [20]. Another predicament for GPCR structure analysis is the stabilization of G_s. Therefore, a nanobody (Nb) library of the β2AR-G_s complex was generated by llama immunization [165]. This library was screened by nucleotide hydrolysis experiments to verify that the G-protein was functionally relevant. Furthermore, they assessed the impact of nanobodies on complex flexibility using negative stain EM [166]. Their data indicated that Nb35 stabilizes the $G\alpha_s$-Gβ interface, does not alter the positioning of the α-helical domain and the addition of GTPγS induces particle dissociation. Hence, Nb35 stabilizes a nucleotide-free complex but remains its biological activity. The first structures of G_s-coupled GPCR complexes without Nb35 were released in 2021 [167]. The short-chain Fab fragment 16 (scFv16) stabilized setmelanotide-MC4R-G_siN complex [92] and the mutation stabilized Gs-coupled to GLP1 [167]. GPCR-G_s protein complexes have been based on the stabilizing effects of nucleotide removal and G-protein stabilization by Nb35. These stabilize the G-protein in a defined state that displays subtle differences at the αN- and α5-interface among different GPCRs but resembles high structural conservation.

1.5.4 Brief introduction in single-particle EM

1.5.4.1 Electron microscopy

Electron microscopy (EM) is an imaging procedure that uses in contrast to light microscopy radiation with significantly smaller wavelengths. The smaller wavelength results in higher resolution, what in turn allows for the visualization of smaller objects. EM was developed in 1931 at the Technische Hochschule Berlin [168] and laid the foundation for microscopes used today (Figure 1.5.4.1).

Electrons combine the characteristics of particles and waves [169] and, if used for imaging, allow the determination of significantly higher resolutions compared to visible light. The

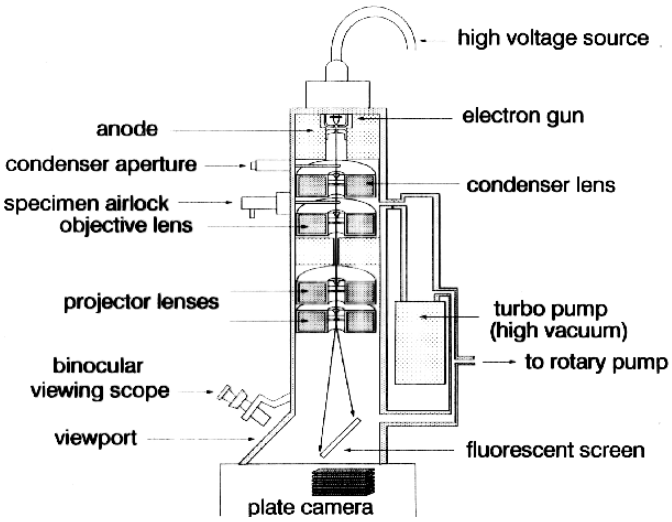

Figure 1.5.4.1 Electron microscope set-up. Schematic drawing of a transmission electron microscope. It is adapted from a drawing by Robert C. Wagner, University of Delaware [footnote 7].

resolution *d*, according to Abbe, is inverse proportional to the wavelength λ if the aperture *A* is constant.

$$d = \frac{A}{\lambda}$$

In addition, there is a direct correlation of the electron acceleration voltage U_B and the wavelength λ of the electron [169].

$$\lambda = \sqrt{\frac{150}{U_B}}\text{Å}$$

Hence, the maximal resolution *d* directly depends on the electron acceleration voltage U_B.

$$d \sim \sqrt{\frac{150}{U_B}}\text{Å}$$

The most common EM method for the structure determination of biological molecules is based on the transmission of electrons (TEM) through an object. Electrons can be diffracted by the object [170] and are phase-shifted to the non-diffracted zero beam. Here, the scattering is distinguished as elastic scattering at the atom nucleus and non-elastic scattering at the

[7] www.udel.edu/Biology/wags/b617/ b617.htm; 15.07.2021

electrons of the atom-shell. The latter is characterized by an energy transfer from the electron beam to the shell electrons that induces an energy reduction and a subtle change in direction. An energy filter can be utilized to separate non-elastic scattered electrons and thus to increase the image contrast. However, electrons do scatter at the air molecules; hence the microscope column must be permanently kept at a high vacuum. The resulting dehydration of biological samples by this vacuum is avoided by the cryo-EM techniques [171]. Here, the specimen is kept frozen and hydrated throughout the imaging process.

1.5.4.2 Contrast-transfer-function

A biological sample is comprised of a multitude of atoms. For example, a GPCR-G-protein complex consists of more than 10,000 atoms (PDB ID: 3sn6). Therefore, the electron beam is scattered at a multitude of points and not at one defined point. The resulting measurement signal is not one single but a superimposition of various contrasts. This complex signal is separated into single components by applying a Fourier transformation. There is a linear correlation between the measured spectrum and the input signal in an ideal situation. An inverse transfer function generates the actual image. The contrast-transfer function (CTF) generates the relationship between the object and the image. The CTF is dependent on the spherical aberration of the objective lens, the defocus, and the spreading and divergence of the electron beam [172].

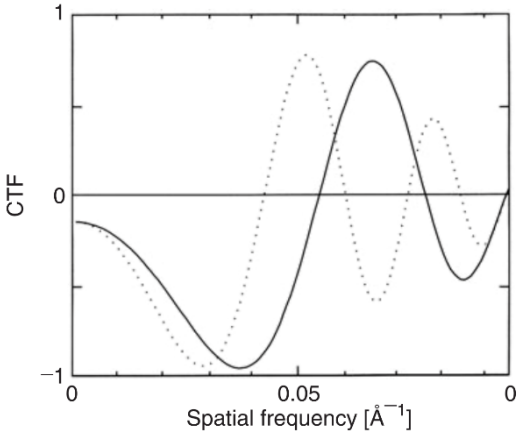

Figure 1.5.2.2 Contrast-transfer function. Electron-optical contrast-transfer function for a mixed-phase / amplitude object, for two defocus values -1 µm (solid line) and -1.5 µm (dashed line). An increase in defocus results in the zero-crossing being closer to the coordinate origin. The figure is taken from [172].

At the zero-crossing of the oscillating CTF is no contrast transfer, meaning no interpretation of the input signal is possible (Figure 1.5.2.2). To assess the maximum number of spatial frequencies, defocus series are recorded [173]. A defocus series consists of varying defocus value measurements for the same image frame; used for the data grouping. Experimental CTFs do not follow an ideal sinusoidal function. With an increase in spatial frequency, the amplitude of the CTF decreases and thereby high-frequency information. The impact of all responsible factors (*e.g.*, drift, detection system, and the electron beam) are summarized in a dampening function [173]. The underrepresented high-frequency information is one of the limiting factors of the resolution in EM. Especially for protein samples, the high spatial frequency information is vital since it represents the fine structure. Low spatial frequency information represents basic structures and shapes. Molecular mechanisms of, *e.g.*, drug binding or enzyme catalysis require high-resolution structures, hence high spatial frequency, to understand the underlying biological mechanism. This general problem of EM is circumvented by collecting a vast number of single images and the *in-silico* summation.

1.5.4.3 Single-particle reconstruction

To reconstruct a three-dimensional structure from single images is a computationally demanding process. The single-particle approach utilizes the projections of isolated single particles to calculate the structure. This requires that the object is represented in varying spatial orientations (projection directions). It is impossible to isolate a defined particle and rotate it freely in all directions in a conventional electron microscope. The object plane can be tilted on a fixed axis, generating several projection directions of one particle, but this is insufficient for high-resolution 3D reconstruction. It is commonly applied for cryo-electron tomography, using sub-tomogram averaging, but the resolution here is limited, yet.

The single-particle reconstruction method relies on the statistical distribution of a homogeneous sample in random orientations on the sample grid. Projections from all spatial directions allow the reconstruction or back projection of the three-dimensional structure. *In silico* ab initio structure determinations aims to gather three Euler angles and two translations for each projection image [174].

The first generation of ab initio structure determination methods uses a set of 2D class averages, which are the product of 2D classification. The success of the 3D orientation

search correlates directly with the quality of the input data. Hence, the used 2D averages must represent homogenous particle groups. Here, the 3D reconstruction strongly depends on the selected particles. This computational approach is based on the central section theorem [175]. That states that Fourier transforms of 2D projections of a 3D object are central sections through the 3D sections of the Fourier transform of the object. Fourier transforms of any two projections intersect along the so-called common line [176]. The common line is the basic principle of the angular reconstitution that takes advantage of the mathematical solution for orienting three projections [176]. The selection of proper 2D classes is the common selection method for *ab initio* reconstruction, but to reduce the bias introduced by the 2D class selection 3D *ab initito* reconstruction methods have been developed in recent years and are implemented, for instance, in cryoSPARC [177]. However, the basic principle remains: a 3D reconstruction is calculated by the back projection of two-dimensional particle images into three-dimensional space.

2 Methods

Methods shown here are in accordance with Heyder *et al.*, 2021[118].

2.1 Constructs and protein preparation for cryo-EM

2.1.1Protein expression of human melanocortin-4 receptor

Wild-type human MC4R (UniprotKB-P32245) was modified to include an N-terminal hemagglutinin signal sequence, followed by a FLAG-tag epitope (DYKDDDK). The C-terminal eGFP, followed by polyhistidine (His-) and rho-1D4 tags, is removable by HRV-3C protease cleavage (construct name: MC4R-eGFP), and the construct was inserted into a pOET3 vector. For the production of MC4R-eGFP, recombinant baculovirus was generated by co-transfecting *Sf*9 cells with pOET3_MC4R-eGFP and linearized BAC10:1629$_{KO}$ [156,178] using Trans-IT Insect (Mirus Bio). *Sf*9 cells were cultured in SF900 II serum-free medium (Invitrogen) at 28 °C for virus generation. A 1 l preparation of *Sf*9 cells at 2 x 10^6 cells ml^{-1} were infected with 10 ml of P2 virus MC4R-eGFP virus. Cultures were grown at 27°C, harvested by centrifugation 48 h post-infection and stored at -20°C.

2.1.2 Protein expression and purification of Gαsβ1γ2 and Nb35

Bovine $G\alpha_s$-short subunit (UniprotKB-P04896-2) in pFastbac vector and rat $G\beta_1$ (UniprotKB-P54311) and bovine $G\gamma_2$ (UniprotKB-P63212) subunits in pFastbacDual vector were previously used and described [19]. Heterotrimeric $G\alpha_s\beta_1\gamma_2$ protein (G_s) was expressed in *Tni* insect cells, maintained in ESF 921™ serum-free insect cell culture media (Expression Systems) at 28°C. The virus was prepared using the Bac-to-Bac™ baculovirus expression system (Thermo Fisher Scientific). The cells were infected with both $G\alpha_s$ and $G\beta_1\gamma_2$ virus, based on small scale titrations and harvested after 48 h post-infection and stored at -20°C. G_s was purified as described previously [19].

According to previously described methods, the single-domain antibody Nanobody-35 (Nb35) was expressed in *E. coli* strain WK6, extracted, and purified by immobilized metal (Ni-NTA) affinity chromatography [19].

2.1.3 Complex formation and purification

MC4R-G_s-Nb35 complexes with both agonists NDP-α-MSH and setmelanotide (from now on abbreviated by agonist) complexes were formed in *Sf9* membranes. *Sf9* cell pellets containing MC4R-eGFP were resuspended in 20 mM HEPES pH 7.5, 50 mM NaCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 25–µM tris–(2-carboxyethyl)–phosphine (TCEP), 25 U/ml apyrase (New England Biolabs) 2.5 mg/ml leupeptin (Enzo Life Sciences, Inc.), 0.16 mg/ml benzamidine (Sigma-Aldrich) and 1 µM of the respective agonist (in-house peptide synthesis). For 1l of cell pellets, G_s pre-incubated with Nb35 was added and incubated overnight at 4°C. The membrane sample containing agonist-MC4R-G_s-Nb35 complex was collected by centrifugation at 46.000xg and carefully resuspended in 20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 25 µM TCEP, 2.5 mg/ml leupeptin, 0.16 mg/ml bezamidine, 1 µM agonist and 1% n-dodecyl β-D-maltoside (DDM), 0.1% cholesteryl hemisuccinate (CHS) (Anatrace, Inc.). After 2 h, the solubilized protein was separated from insoluble remains by centrifugation at 46.000xg. The supernatant was diluted twofold with 20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 5 mM $CaCl_2$, 25 µM TCEP, 2.5 mg/ml leupeptin, 0.16 mg/ml benzamidine, and 1 µM agonist. ANTI-FLAG M1 resin (Sigma-Aldrich) was added and incubated for 2 h at 4°C rotating. M1 resin was collected by centrifugation (500 x g, 5 min) and loaded into a wide-glass column and washed for 5 column volumes with wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1mM $CaCl_2$, 25 µM TCEP, 1 µM agonist) with 0.1% DDM, and 0.01% CHS. Followed by a 1 h incubation in wash buffer with 0.8% lauryl maltose neopentyl glycol (LMNG), 0.08% CHS (Anatrace, Inc.) and 0.02% DDM. Subsequently, LMNG/CHS concentration was lowered stepwise to 0.01% LMNG, 0.001% CHS for 1 h. Elution of the complex was initiated by addition of 20 mM HEPES pH 7.5, 150 mM NaCl, 25 µM TCEP, 1 µM agonist, 0.01% LMNG, 0.001% CHS, 5 mM ethylenediamine tetraacetic acid (EDTA) and 0.2 mM DYKDDDDK peptide (GenScript Biotech). C-terminal eGFP was removed by adding HRV-3C protease (in-house purified), incubated at 4°C overnight. After concentration the agonist-MC4R-G_s-Nb35 complex was loaded onto a Superdex 200 Increase 5/150 GL (Sigma-Aldrich). Receptor containing fractions were concentrated to 5 mg/ml and directly vitrified. 1 l of MC4R eGFP expressing cells yielded 0.25 mg complex.

2.2 Structural biology methods

2.2.1 Negative stain EM of NDP-α-MSH-MC4R-eGFP-G_s-Nb35

Magdalena Schacherl supported me in the negative stain sample preparation and data processing. Jörg Bürger in the group of Thorsten Mielke at the Max Planck Institute for molecular genetics measured the initial negative stain specimen.

NDP-α-MSH-MC4R-eGFP-G_s-Nb35 at a 20 µg/ml concentration in a buffer containing 0.1% DDM/CHS complex was plotted on glow discharged QUANTIFOIL R 2/1 on Cu 300 mesh + 2 nm C grid with 2% uranyl acetate. Micrographs were recorded by a Tecnai Spirit 120 keV microscope with a pixel size of 2.67 Å. CTF correction was done using CTFfind [179]. Particles were selected using gautomatch [180] with a particle diameter of 160 Å. Particles were extracted at full binning size with a particle box size of 90 pixels with RELION [181]. The initial model reconstruction was done after several rounds of 2D classifications and *in silico* purification using RELION 3.01 [182].

2.2.2 Cryo-EM sample preparation and image acquisition

Tarek Hilal carried out cryo-EM sample vitrification and the image acquisition in the FZEM facility of the Freie Universität Berlin. He supported and instructed me to use cryoSPARC for data processing. He processed with me the initial cryo-EM densities used for model building. Andrea Schmidt supervised and supported me by model building and refinement of the complexes.

Vitrification of NDP-α-MSH and setmelanotide-MC4R-G_s-Nb35 complex was conducted immediately after sample preparation at a concentration of 1.2 mg/ml and 5 mg/ml, respectively. 3.8 µl of the sample was applied to glow-discharged holey gold grids (UltrAuFoil R1.2/1.3 300 mesh, Quantifoil Micro Tools GmbH), blotted for 4 s and plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV (Thermo Fisher Scientific) set to 10°C and 100% humidity.

Images were acquired using a FEI Titan Krios G3i microscope (Thermo Fisher Scientific) operated at 300 kV equipped with a FEI Falcon 3EC detector (Thermo Fisher Scientific) running in counting mode at a nominal magnification of 96,000x, giving a calibrated pixel size of 0.832 Å/px. Movies were recorded for 40.78 s accumulating a total electron dose of

40 $e^-/Å^2$ fractionated into 33 frames. EPU 2.8 was utilized for automated data acquisition with AFIS enabled using a nominal defocus between -0.8 and -2 µm.

A total of 5618 micrographs were collected for NDP-α-MSH-MC4R-G_s-Nb35 and 7583 micrographs for setmelanotide-MC4R-G_s-Nb35. These were used for further image processing. Further details are given in Table S1.

2.2.3 Cryo-EM image processing

The complete data analysis was conducted within the cryoSPARC v2.15 framework.

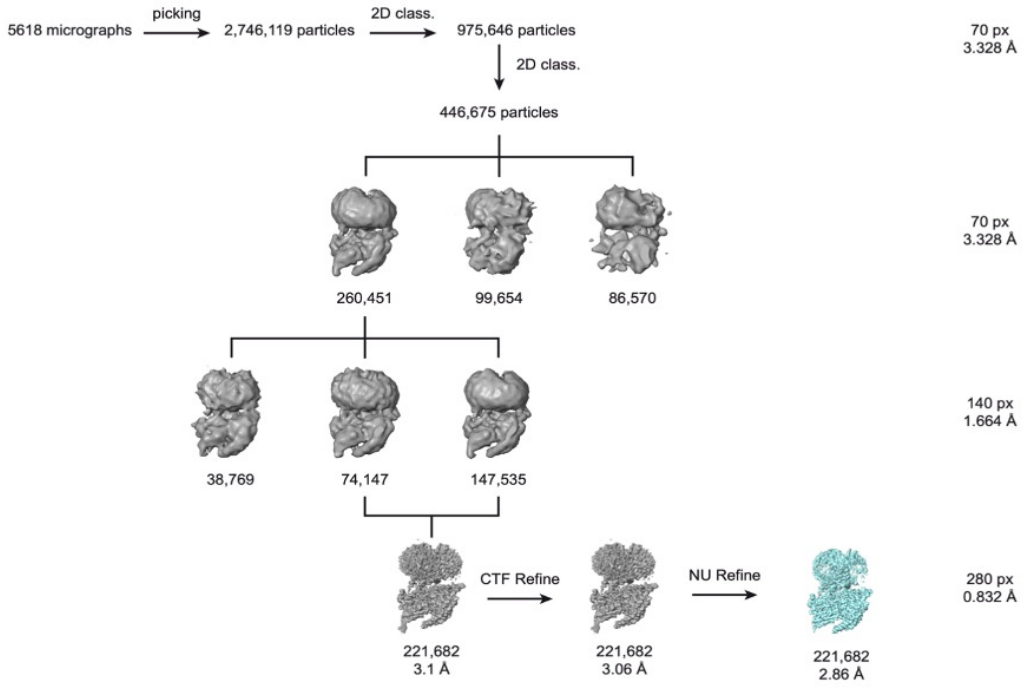

Figure 2.2.3.1 ***Cryo-EM refinement sorting scheme of the active NDP-α-MSH-MC4R-G_s-Nb35 complex dataset.*** After template-based particle picking, 2,746,119 particle images were extracted and Fourier-cropped to a box size of 70 px (pixel size 3.328 Å). After two subsequent runs of reference-free 2D classifications, 446,675 particles were selected and subjected to heterogeneous 3D refinement yielding 260,451 particle images. Re-extraction of particle images Fourier-cropped to 140 px box size (1.664 Å pixel size) was followed by another round of heterogeneous refinement. The reconstruction of the two classes was virtually identical by visual inspection; therefore, particle images were combined and re-extracted unbinned with a box size of 280 px (0.832 Å/px). Homogeneous refinement generated a reconstruction with a resolution of 3.1 Å, which was improved to 2.86 Å by CTF refinement followed by non-uniform (NU) refinement.

NDP-α-MSH-MC4R-G_s-Nb35 dataset

Image analysis of the NDP-α-MSH-MC4R-G_s-Nb35 dataset started with movie alignment and dose-weighting using "Patch motion correction" followed by "Patch CTF estimation". The processing pipeline is visualized in Figure 2.2.3.1. Initial particle picking was done with the "Blob picker" using a particle diameter of 180 Å. Particle images were extracted with a box size of 280 px, Fourier-cropped to 70 px (3.328 Å/px). After reference-free 2D classification, selected class averages were used for template-based particle picking with a 170 Å particle mask. A total of 2,746,119 particle images were subjected to two cycles of 2D classification to clean the dataset. Ab initio reconstruction of particle images belonging to shiny classes was applied to generate a reference model for 3D classification, after which 260,451 particle images were selected for further processing. Homogeneous refinement after re-extraction of the particles, Fourier-cropped to 140 px (1.664 Å/px) generated a 3D reconstruction of 3.45 Å global resolution. Another round of heterogeneous refinement was applied to finally select 221,682 particle images for unbinned extraction (280 px, 0.832 Å/px). Iterations of homogeneous refinement and Global CTF refinement were applied to correct for higher-order aberrations yielding a final reconstruction of 2.86 Å resolution after non-uniform (NU) refinement [183]. Using NU-refinement, masking off the all-helical domain was not necessary to yield a high-resolution map.

Setmelanotide-MC4R-G_s-Nb35 dataset

Processing of the setmelanotide-MC4R-G_s-Nb35 data was done as described for the NDP-α-MSH-MC4R-G_s-Nb35 dataset using the previously generated templates for picking of 4,330,500 particle images. The processing pipeline is visualized in Figure 2.2.3.2. After a single round of 2D classification, 4,267,612 particle images were subjected to two iterative rounds of 3D classification with the NDP-α-MSH-MC4R-G_sαβγ-Nb35 reconstruction as reference filtered to 30 Å. Micrographs with local motions above 10 px or estimated resolutions worse than 4 Å were discarded, leaving a total of 797,185 particle images for another round of 3D classification. Homogeneous refinement of 431,973 particle images after re-extraction with a box size of 280 px (0.832 Å/px) yielded a resolution of 2.82 Å that could be improved to 2.77 Å by CTF refinement. After a final 3D classification, 370,621 particles were selected for NU refinement resulting in a 2.58 Å reconstruction.

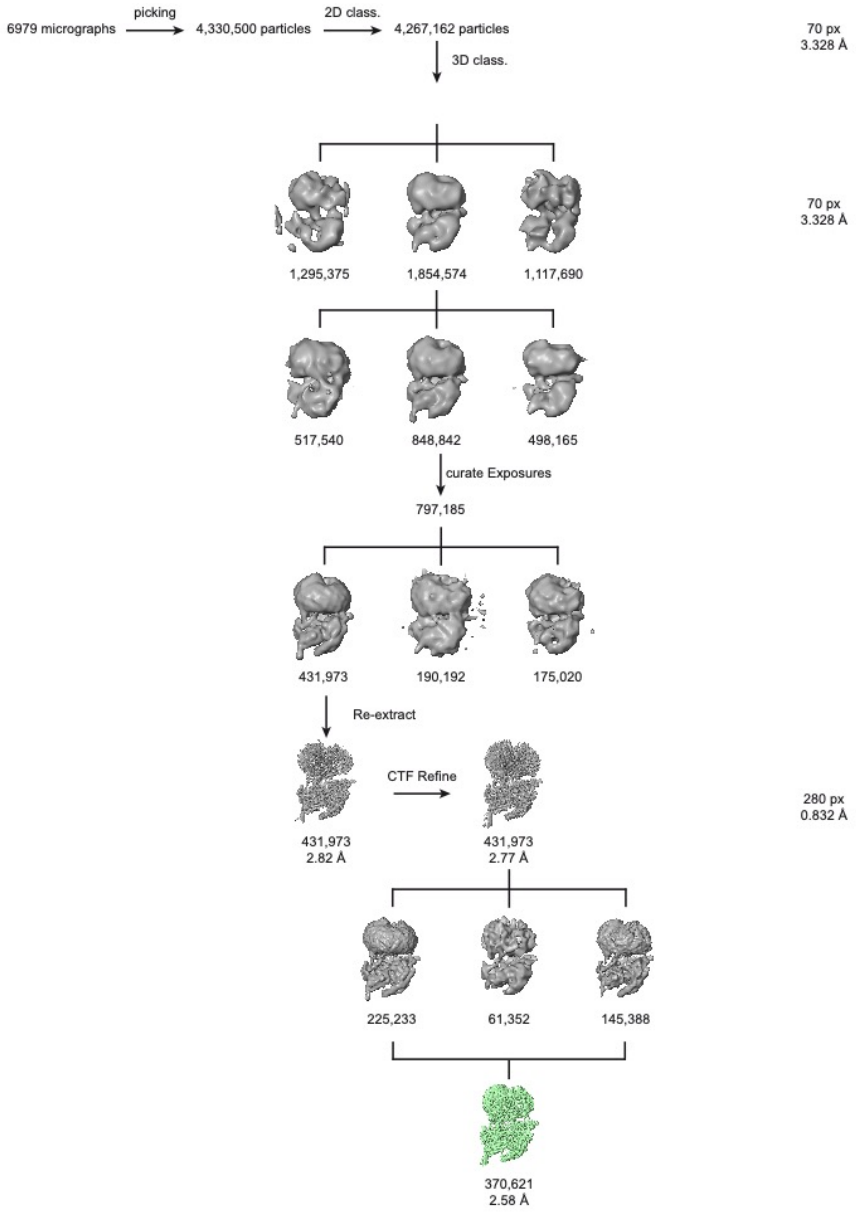

Figure 2.2.3.2 ***Cryo-EM refinement sorting scheme of the active setmelanotide-MC4R-G_s-Nb35 complex dataset.*** Initially, 4,330,500 particles were picked and extracted by Fourier-cropping with a box size of 70 px (3.328 Å/px). 2D classification was followed by two iterative 3D classifications using the NDP-α-MSH-MC4R-G_s-Nb35 filtered to 30 Å as template yielding 848,842 particle images. Micrographs were curated omitting resolutions above 4 Å and high local motion, leaving 797,185 particles for another iteration of heterogeneous refinement. 431,973 particles were re-extracted at full resolution (0.832 Å/px) and subjected to homogeneous refinement. The resulting 2.82 Å reconstruction was improved by CTF refinement to 2.77 Å. Another heterogeneous refinement was conducted to select 370,621 particle images for a final NU refinement, yielding a 2.58 Å reconstruction of the setmelanotide-MC4R-G_s-Nb35 complex.

2.2.4 Model building and refinement

The models of the NDP-α-MSH-MC4R-G_s-Nb35 complex as well as setmelanotide-MC4R-G_s-Nb35 complex were derived from the inactive MC4R structure (PDB ID: 6w25 [91]) together with the G_sαβγ-Nb35 complex of the β2-adrenergic receptor-G_s complex (PDB ID: 3sn6 [19]) as initial models. Both MC4R complexes were built and adjusted manually using the program COOT [184]. Model building for the MC4R ligands NDP-α-MSH and setmelanotide were started *de novo* using COOT. Local-refined, as well as overall cryo-EM maps were used to add water molecules. After every round of manual refinement real-space refinement was performed with the program PHENIX [185] using geometric restraints, a global minimalization protocol and B-factor refinement. Both models were refined with isotropic B-factors in reciprocal space using REFMAC5 [186] of the CCP4 (Collaborative Computational Project, number 4) software suite [187]. The refinement was carried out in the resolution range of 233 - 2.88 Å and 233 - 2.6 Å for the NDP-α-MSH-MC4R and setmelanotide-MC4R complexes, respectively (Table S1).

Structure validation was performed with the programs PHENIX, MolProbity [188], SFCHECK [189], and OneDeep of the Protein Data Bank [187]. Potential hydrogen bonds and van der Waals contacts were analyzed using the programs HBPLUS [190] and LIGPLOT 1.45 [191]All structure superpositions of backbone α-carbon traces were performed using the CCP4 program LSQKAB. All molecular graphics representations in this work were created using the PyMol Molecular Graphics System Version 1.3 (Schrödinger, LLC, New York, NY) and UCSF Chimera [192].

Coordinates have been deposited in the Protein Data Bank (PDB) [193] with identification codes 7pif, and 7piu. Cryo-EM maps have been deposited in the Electron Microscope Databank (EMDB) with identification codes EMD-13454 and EMD-13453.

2.3 Functional characterization of MC4R

2.3.1 Saturation and competition binding assay using nanoLuc Luciferase assay

Wild-type MC4R was modified to include an N-terminal hemagglutinin signal sequence, followed by Luciferase (nanoLuc™ Luciferase; Promega) [143] and cloned into pMT4 vector. Human embryonic kidney 293 (HEK293T) cells grown in DMEM/F-12 (Thermo Fisher Scientific) medium (supplemented with L-Glutamin, HEPES, phenol red, sodium pyruvate pH 6.9-7.3) (Thermo Fisher Scientific, Gibco) were transiently transfected using FuGENE® HD transfection reagent (Promega). 20.000 cells per well were seeded in white corning assay 96 well plates. After 24 h, medium was removed and replaced with 75 µl ligand serial dilutions in Opti-MEM reduced serum media (Thermo Fisher Scientific) without phenol red and incubated for 2 h.

For saturation experiments, TAMRA-NDP-α-MSH (TAMRA-NDP) labeled with the fluorophore 5-carboxy-tetramethylrhodamine (TAMRA) was titrated from 1 µM to 1 pM. For competition experiments, each well contains 20 nM TAMRA-NDP, and the competing ligand was titrated from 1 µM to 10 pM for setmelanotide. Non-specific binding was measured by the addition of 20 µM NDP-α-MSH, to saturate the ligand-binding pocket with the non-fluorescent ligand. After 2h, 25 µl Furimazine (Promega) was added and incubated for 15 min. Luminescence and resulting BRET (bioluminescence resonance energy transfer) was measured in Spectramax ID3 (Molecular Devices) and CLARIOstar Plus (BMG LABTECH) plate reader with 460 nm (short-pass filter) and 610 nm (long-pass filter).

The BRET ratio was calculated by the quotient of long-pass divided by short-pass [143]. GraphPad PRISM 8 (GraphPad Software Inc.) was used for analysis by sigmoidal dose-response (variable slope) for dose-response measurements and one site - Fit K_i for competition experiments.

2.3.2 Functional characterization by NanoGlo, HiBiT, AlphaScreen and PLC activation assays

Signaling data were measured and analyzed by Sarah Paisdzior, Sabine Jyrch, and Heike Biebermann. The Scheerer lab selected the variants, assay conditions and analyzed the data together with the Biebermann lab.

2.3.2.1 Cell lines, cloning, and reagents

HEK293 cell line was purchased from ATCC. Cells were authenticated by single nucleotide polymorphism (SNP) analysis and regularly tested for mycoplasma contamination. Cultivation took place in L-glutamine containing minimal essential medium MEM (Merck Biochrom) supplemented with 5% fetal bovine serum FBS (Thermo Fisher Scientific, Gibco) and 1% non-essential amino acids NEA (Merck Biochrom) at 37°C and humidified air containing 5% CO_2. For cAMP accumulation assays, 1.5 x 10^4 cells per well were seeded in poly-L-lysine coated (Merck Biochrom) translucent 96 well plates (Falcon) and incubated for 24 h. Identically, determination of total and cell surface expression (NanoGlo®HiBiT assay, Promega) were performed in white opaque, poly-L-lysine coated 96 well plates (Corning #3917).
MC4R cDNA was amplified from genomic DNA and cloned into eukaryotic expression vector pcDps. The receptor was N-terminally tagged with the hemagglutinin (5'-YPYDVPDYA-3') epitope (HA) for cAMP measurements and luciferase-based assays. For NanoGlo®HiBiT assays, MC4R was cloned into pBiT3.1-N (Promega) using EcoRI/BamHI restriction sites, resulting in HiBiT protein tag N-terminally spaced by eleven amino acids. All single point mutations were incorporated into the expression vectors using site-directed mutagenesis. Cloned constructs were sequenced and verified with BigDye-terminator sequencing (PerkinElmer Inc.) using an automatic sequencer (ABI 3710 XL; Applied Biosystems). α-MSH, NDP-α-MSH and 3-Isobutyl-1-methylxanthine (IBMX) were purchased from Sigma-Aldrich.

2.3.2.2 Transfection

For determination of cAMP accumulation, HEK293 cells were transfected 24 h after seeding. Cells were transfected with 45 ng plasmid DNA and 0.45 µl Metafectene (Biontex) per well

in MEM without supplements. For the NanoGlo®HiBiT assay, transfection was performed as described previously elsewhere [194]. In short, low amounts of HiBiT-tagged receptor mutants were transfected (0.45 ng/well) and carrier DNA (pGEM-3Zf(+), Promega) was added to 45 ng DNA/well in total in advanced MEM (Thermo Fisher Scientific, Gibco) to ensure comparable transfection conditions to the other performed assays. Transfection for both HiBiT-assay (cell surface and total expression) was carried out simultaneously to ensure comparability.

2.3.2.3 Determination of total and cell surface expression using NanoGlo®HiBiT assay

The level of receptor expression on the cell membrane and total cell expression was determined using the NanoGlo®HiBiT detection system (Promega). The assay was performed according to the manufacturer's protocol (rapid measurements) and has been described elsewhere [194]. In short, 48 h after transfection, media was changed in Opti-MEM reduced serum media (Thermo Fisher Scientific, Gibco) without phenol red (50 µl/well) to remove background noise. Cell surface expression was determined by injection of 50 µl of HiBiT extracellular substrate in the appropriate buffer supplemented with LgBiT. Total expression analyis was carried out similarly, with 50 µl/well of HiBiT Lytic substrate combined with LgBiT in the appropriate buffer containing detergents to lyse the cells (Promega). After orbital shaking for 3 min at 300 cycles per minute, plates were incubated for 10 min at room temperature. Luminescence was measured using a plate reader (Mithras LB 940, Berthold Technologies). As background control, cells transfected with empty vector pcDNA3 were used, and values were subtracted from the sample emissions.

2.3.2.4 Determination of G_s activation by measurement of cAMP accumulation using AlphaScreen assay

Ligand-induced activation of MC4R was determined by using the AlphaScreen™ assay (Perkin Elmer Life Science) according to the manufacturer's protocol and described elsewhere [195]. In brief, 48h post-transfection, cells were challenged with either α-MSH or NDP-α-MSH (1 µM to 0.1 nM) in stimulation buffer (138 nM NaCl, 6 mM KCl, 1 mM $MgCl_2$*$6H_2O$, 5.5 mM glucose, 20 mM HEPES, 1 mM $CaCl_2$ * $2H_2O$, 0.1% BSA, pH 7.4) containing 1 mM IBMX for 40 min for at 37°C and 5% CO_2. Incubation was stopped by freezing cells at -80°C for 10 min before cAMP measurements. The determination of cAMP

accumulation was performed following the manufactures' instructions (Perkin Elmer Life Science) and measured with a plate reader (Mithras LB 940, Berthold Technologies).

2.3.2.5 Determination of phospholipase C (PLC) activation via NFAT reporter gene assay

As a read-out system for $G_{q/11}$ signaling, activation of phospholipase C activation was investigated using NFAT responsive element (NFAT-luc, pGL4.33, Promega) located in the promoter region of the gene encoding firefly luciferase and luciferase is expressed upon second messenger activation. After incubation for 48h after transfection, cells were stimulated with NDP-α-MSH or setmelanotide (1 mM to 0.1 nM) in MEM without supplements for 6 h at 37°C and 5% CO2. The stimulation was stopped by media exchange with 1X passive lysis buffer (PLB, Promega) and freezing the cells at -80°C for 10 min. Luciferase activity was then measured and provided information about the activation of the respective second messenger by transferring 10 µl lysate into a white opaque 96 well plate. Injection of 40 µl firefly luciferase substrate (Promega) and measurement of luminescence was performed with a plate reader (Mithras LB 940).

2.3.2.6 Statistical analysis

Statistical analysis was performed using GraphPad Prism 6. Appropriate tests were carried out and indicated for each individual data set. Statistical significance was set at $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$ and $****p \leq 0.0001$. Concentration-response curves of each experiment were analyzed by fitting a non-linear regression model for sigmoidal response in GraphPad PRISM 6 (GraphPad Software Inc.) to determine EC_{50} values. Statistics for all functional data are given in Tables 3.5.1 - 3.5.3.

3 Results

3.1 Sample preparation and complex formation for cryo-EM

3.1.1 NanoLuc ligand binding assay

GPCR research relies on the usage of well-characterized ligands. In the past, radioligand assays were used predominantly. Tritiated small molecules and I^{125} labeled peptidic ligands are commercially available. However, for MC4R, using a radiolabeled peptidic ligand is hampered by the high costs and radioactive properties of I^{125} labeled peptides. Therefore, I established a nanoLuc (nLuc) based ligand binding assay [196] and adapted the assay for MC4R (Figure 3.1.1). In brief, the nLuc protein was fused at the receptor amino terminus (Figure 3.1.1A), and HEK 293T cells were transfected with nLuc-MC4R and seeded in a 96-well plate (Figure 3.1.1B). Bioluminescence resonance energy transfer (BRET) is observed depending on the relative distance of 5-carboxytetramethylrhodamine (TAMRA) labeled NDP-α-MSH (TAMRA-NDP) and nLuc-MC4R.

Titration of TAMRA-NDP from 10 µM to 0.1 pM is plotted as a concentration-response measurement with a K_D of 1.44 nM (Figure 3.1.1C). Setmelanotide binding was determined by competing the agonist setmelanotide against 20 nM TAMRA-NDP with a resulting K_i of 3.1 nM (Figure 3.1.1D). NDP-α-MSH (K_i = 0.7 nM) and setmelanotide have a significant higher affinity towards MC4R compared to the endogenous agonists α-MSH (K_i = 72 nM) and β-MSH (K_i = 30 nM) (Figure 3.1.1D). Hence, the two high-affinity agonists are the suitable tool compounds for the MC4R-G_s-protein complexation experiment, and the ligand-binding assay confirms the biological activity and stability of the acquired peptides.

Expression screening via transient infection in HEK293 cells is less time-consuming and, hence, easier reproducible than the baculovirus expression system. The nanoLuc assay was adapted to screen MC4R constructs for functionality by concentration-response measurements (exemplary graph Figure 3.1.1.E) and expression levels (Figure 3.1.1F). The luminescence signal correlates linearly with the expression of nLuc-MC4R and nLuc-MC4R mutants. More than 40 constructs were screened (not shown). Single side-chain and carboxy-terminal truncations were analyzed. Expression levels increased

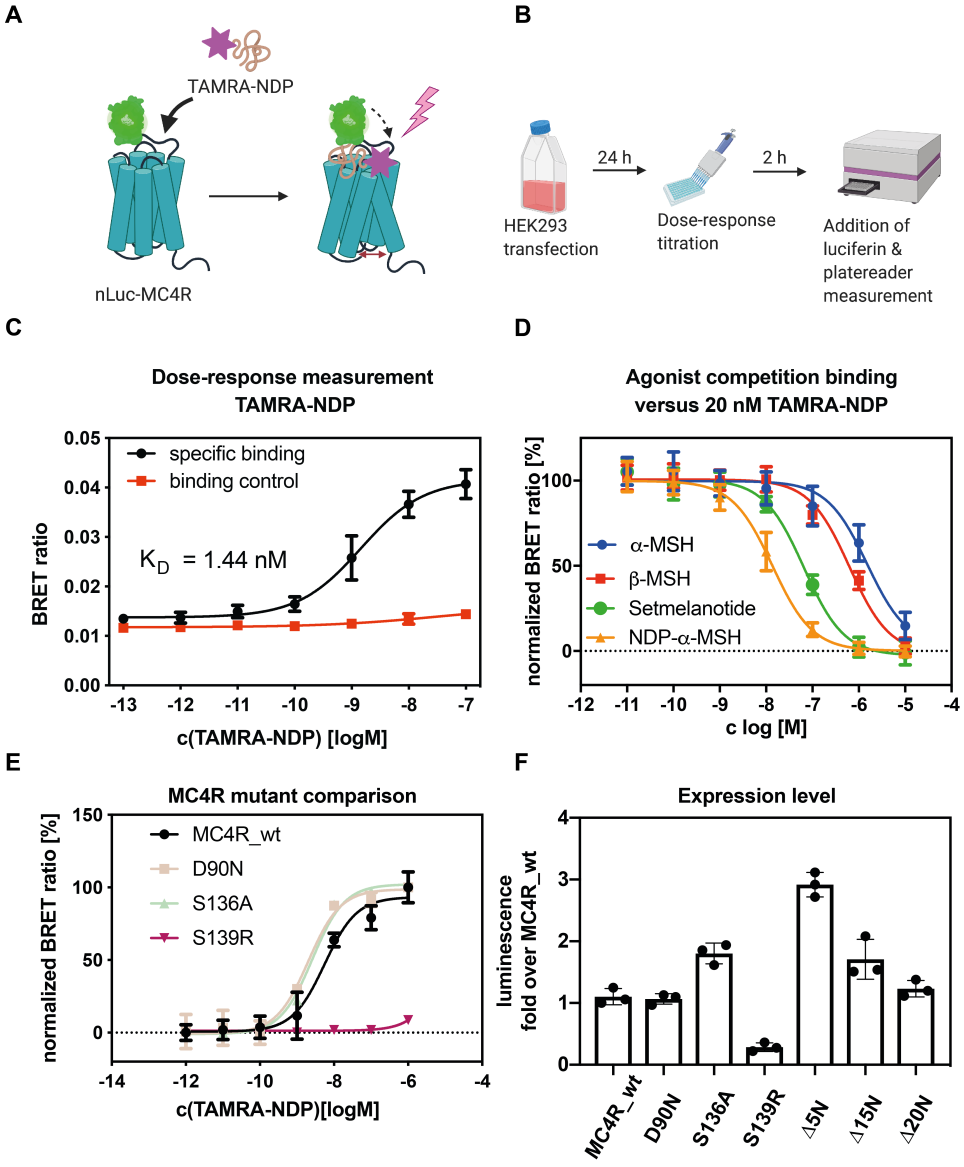

Figure 3.1.1 ***NanoLuc luciferase assay set-up and data.*** (**A**) Schematic representation of nanoLuc Luciferase (nLuc) based MC4R ligand binding assay. (**B**) Ligand binding assay workflow, HEK293T cells were infected 24 h prior, followed by media exchange against ligand titrations. Ligand binding equilibration is ensured by 2 h incubation time with the subsequent addition of the nLuc substrate luciferin and measurement of the short-pass filter (460 nm) and long-pass filter (610 nm) using a fluorescent plate reader. The BRET ratio is the quotient of long-pass by short-pass. (**C**) Concentration-response measurement of TAMRA-NDP-α-MSH (TAMRA-NDP). The addition of 20 µM non-fluorescent labeled NDP-α-MSH is used as non-specific binding control. (**D**) Competition binding experiment of α-MSH (blue), β-MSH (red), setmelanotide (green), and NDP-α-MSH (yellow) binding against 20 nM TAMRA-NDP. (**E**) Normalized concentration-response measurement of TAMRA-NDP binding to a wildtype nLuc-MC4R (MC4R_WT) and three nLuc-MC4R mutants (D90N - yellow, S136A - green, and S139R - magenta). (**F**) Expression comparison of nLuc-MC4R mutants by comparing the luminescence signal in the absence of TAMRA-NDP.

with amino-terminal truncated constructs until twenty residues. In contrast, C-terminal truncations reduced the expression levels of MC4R variants. Ultimately, the construct optimization was not implemented into the final MC4R construct. High expression yields of the MC4R wild-type construct, due to the successful optimization of the baculovirus expression system, made construct optimization obsolete. All structural experiments were performed with a wild-type MC4R construct.

3.1.2 Protein production

Five proteins are required to stabilize a G_s-coupled GPCR in the active state. The receptor of interest, the heterotrimeric $G\alpha_s\beta_1\gamma_2$, and Nb35 (constructs shown in Figure 3.1.2). During the first third of this doctorate, the insect cell culture workflow in the laboratory was established to produce milligram amounts of MC4R in *Sf*9 insect cells (Figure 3.1.2.1) and heterotrimeric G_s in *Tni* cells (Figure 3.1.2.2). However, the successful expression of receptors and G-proteins requires the production of potent viruses. Therefore, we adapted and extended a cell culture pipeline that focuses on virus titrations in the P1, P2, and test expression in a $\log^3$ range (such as 1:50, 1:150, and 1:450 virus volume per cell suspension volume).

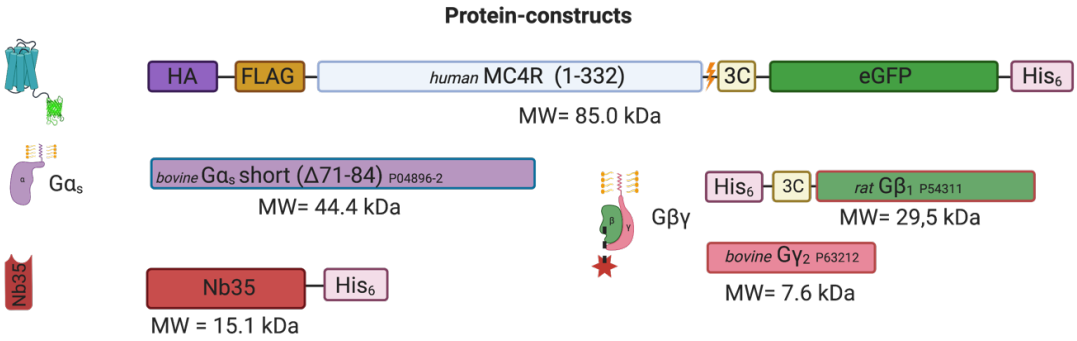

Figure 3.1.2 ***Protein construct used for MC4R-G_s-Nb35 complex studies.*** Five proteins comprise the final complex; the following constructs are schematically shown: Hemagglutinin signal peptide (HA) - M1-FLAG tagged MC4R-eGFP with a carboxy-terminal His_6 tag (His_6), the heterotrimeric G_s-protein ($G\alpha_s$, His-tagged $G\beta_1$ and $G\gamma_2$), and His tagged nanobody35 (Nb35).

3.1.2.1 Expression and purification of MC4R

The most critical step to produce MC4R is generating a potent virus. Initial experiments indicated that a potent MC4R virus allows the cells to duplicate only once in 48 hours post-infection, and the cell viability should be in a range of 60-90% (data not shown). More than 40 MC4R constructs were screened, and the carboxy-terminal extension with eGFP increased the expression of MC4R by more than times (data not shown). The high expressions allowed the use of full-length wild-type MC4R without any modifications inside the receptor sequence.

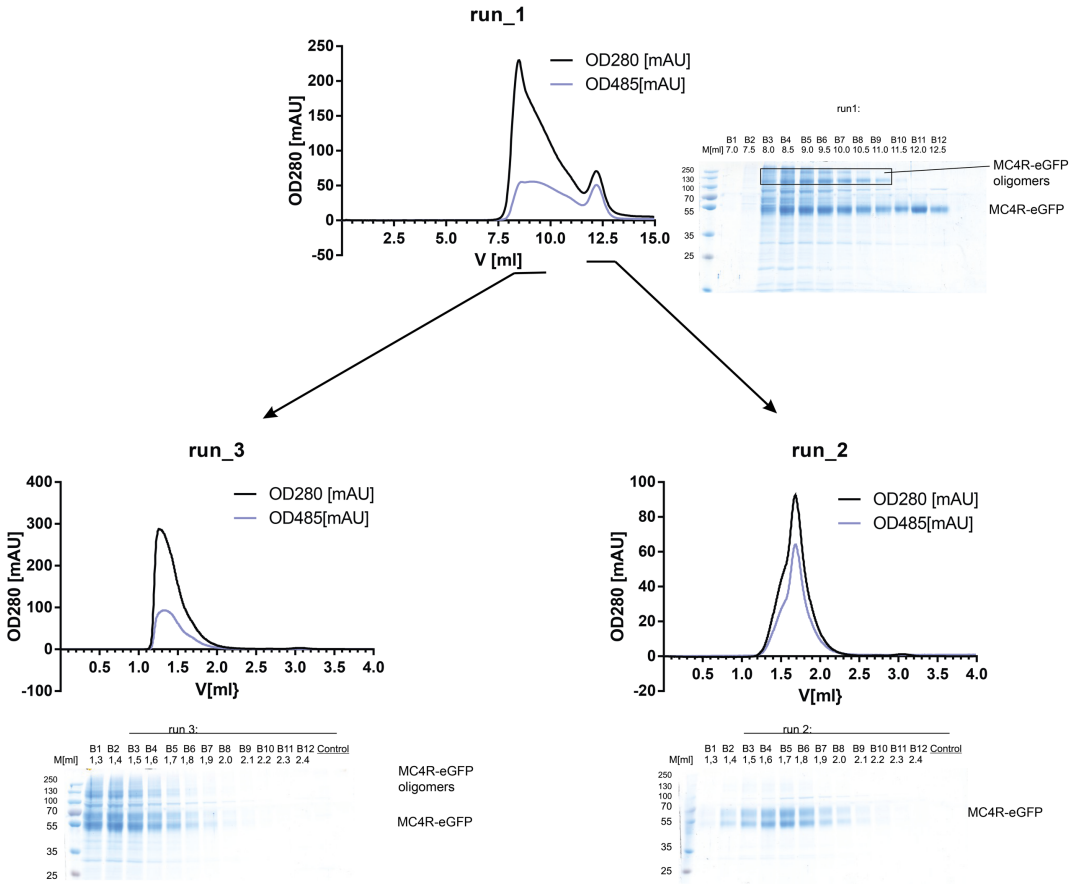

Figure 3.1.2.1 ***Analysis of the multimerization state of detergent purified MC4R-eGFP.*** SEC chromatograms of detergent purified MC4R-eGFP indicate that the receptor forms dimers or multimers. (**run_1)** The product of an M1-FLAG purification of 1 l insect cell suspension infected with MC4R-eGFP virus was applied on a preparative column (Superdex 200 Increase 10/300GL). Fraction B3-B5 (**run_3**) and B10-B12 (**run_2)** of run_1 were applied separately to the analytical column (Superdex 200 Increase 5/150GL) (indicated by black arrows). Each run was analyzed by gel electrophoresis. Signals for multimeric receptors are not visible in the gel of run 3. An SDS gel artifact is the double band around the 55kDa signal in the gel of run_2 and run_3.

The purification of MC4R-eGFP resulted in large proportions of receptor accumulated in the size-exclusion chromatography (SEC) fractions following the void volume of the preparative column (Superdex 200 Increase 10/300GL) (Figure 3.1.2.1 run_1).
Gel chromatography of this SEC run displays signals above the 55 kDa marker protein that are MC4R oligomers and contaminations. The SEC signal at 12.5 ml elution volume resembles monomeric MC4R-eGFP. 75% of the protein, according to the 280 nm absorption (Figure 3.1.2.1_run_1), is lost in the peak with the maximum at 8 ml elution volume. A subsequent run (run_2) with the analytical column (Superdex 200 Increase 5/150GL) has the maximum signal at 1.7 ml elution volume, which is in an expected elution volume for an 85 kDa protein in a detergent micelle. Due to the large amount of protein accumulated in the multimeric state, the complexation of detergent purified MC4R-eGFP was not further pursued. However, the large-scale production of MC4R-eGFP is possible via the baculovirus-*Sf*9 system.

3.1.2.2 Expression and purification of G_s and Nb35

Heterotrimeric G_s is membrane-anchored via a geranylation at the $G\gamma_2$ subunit. Therefore, the presence of a detergent is a necessity for all purification steps. $G\beta_1$ is His_6 tagged (Figure 3.1.2), the $G\alpha s$ and $G\gamma_2$ subunits are co-purified via nickel affinity chromatography (Figure 3.1.2.2B). Hence, the yield of stoichiometric expressed $G\alpha_s$ and $G\beta_1\gamma_2$ indicates properly folded heterotrimeric G_s; the excess purified $G\beta_1\gamma_2$ is separated by ion-exchange chromatography (Figure 3.1.2.2C-D). The purification of 1 l of infected *Tni* insect cells yields 3.5 mg of G_s. The purification protocol was optimized for the final G_s protein yield, not purity. The impurities of Figure 3.1.2.2D do not cause concern since the receptor-G_s-protein complex is purified via the FLAG tag at the receptor amino terminus.
Nb35 was expressed in the periplasm of *E.coli* strain WK6 and is purified via nickel resin; the purification yields 2mg/l of infected expression culture (data not shown).

3.1.3 MC4R-G_s-Nb35 complex formation

3.1.3.1 Optimization of complex purification

A key learning in the process of the sample optimization was that the receptor-G_s-protein

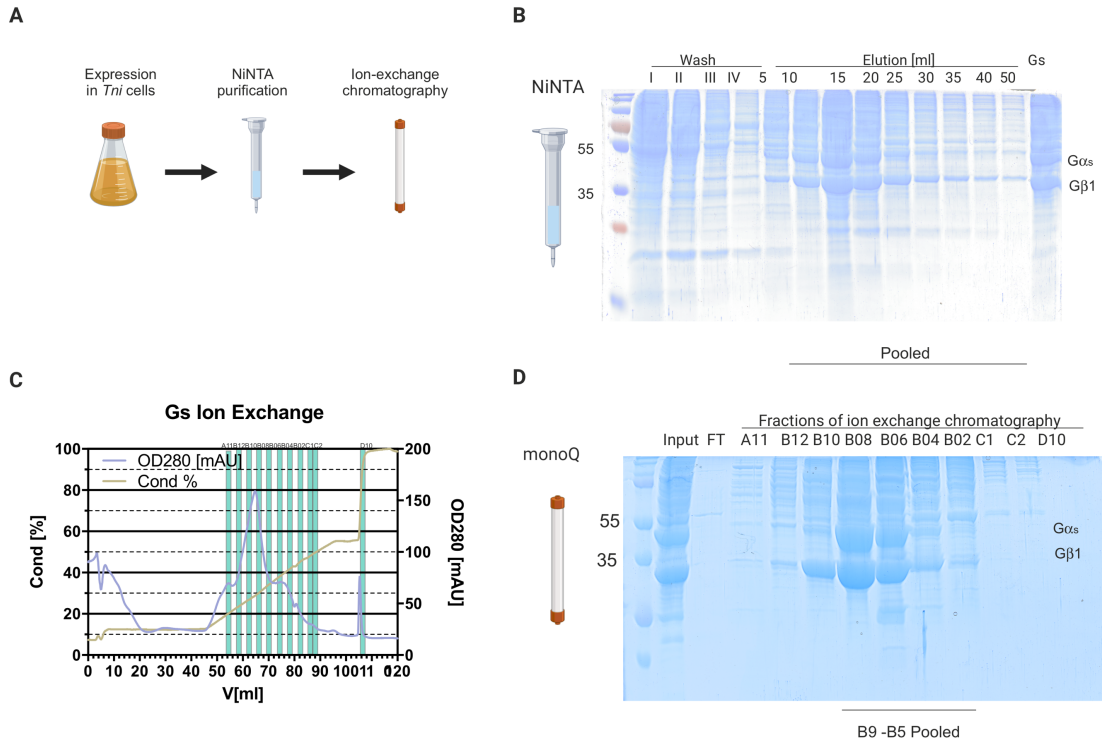

Figure 3.1.2.2 ***Purification of heterotrimeric*** G_s**.** (**A**) Scheme of G_s purification was created by biorender.com. The co-infection of *Tni* cells with P2 virus of $G\alpha_s$ and $G\beta1\gamma2$ was purified using Ni-NTA agarose (NiNTA), followed by ion-exchange chromatography using a monoQ column. Highlighted fractions (cyan rectangles) of the ion-exchange chromatogram (**C**) were analyzed via SDS PAGE (**D**). Fractions B04 - B08 were pooled and used for GPCR complexation experiments.

complex formation at the *Sf9* insect cell membrane reduced the strong tendency of MC4R-eGFP to oligomerize (Figure 3.1.3.1B-C). The complex formation was initiated by adding detergent-purified G_s-protein, *Sf*9 insect cells infected with MC4R-eGFP virus, Nb35, 1 µM agonist, and a low salt buffer for membrane disruption. The addition of nucleotide hydrolase apyrase stabilizes the complex as a nucleotide-free agonist-bound MC4R-eGFP-G_s-Nb35 complex (scheme shown in Figure 3.1.3.1A). The complex extraction of the membrane was optimized using fluorescent-size-exclusion chromatography (F-SEC) (data not shown). F-SEC was used to test a variety of buffer conditions (pH, salt concentration, monovalent, and divalent cations) and detergents (NG310, NG311, NG322, DDM, CHS, CHAPS, CHAPSO, n-OG, DM, GDN, and LMNG) (data not shown). The optimization of detergents and buffer conditions by F-SEC did not substantially improve protein yield and monodispersity compared to published protocols [197,198]. The highest solubilization yield was achieved using 1% n-dodecyl-β-D-maltoside (DDM) and 0.1% cholesteryl hemisuccinate (CHS) in a buffer containing 20 mM

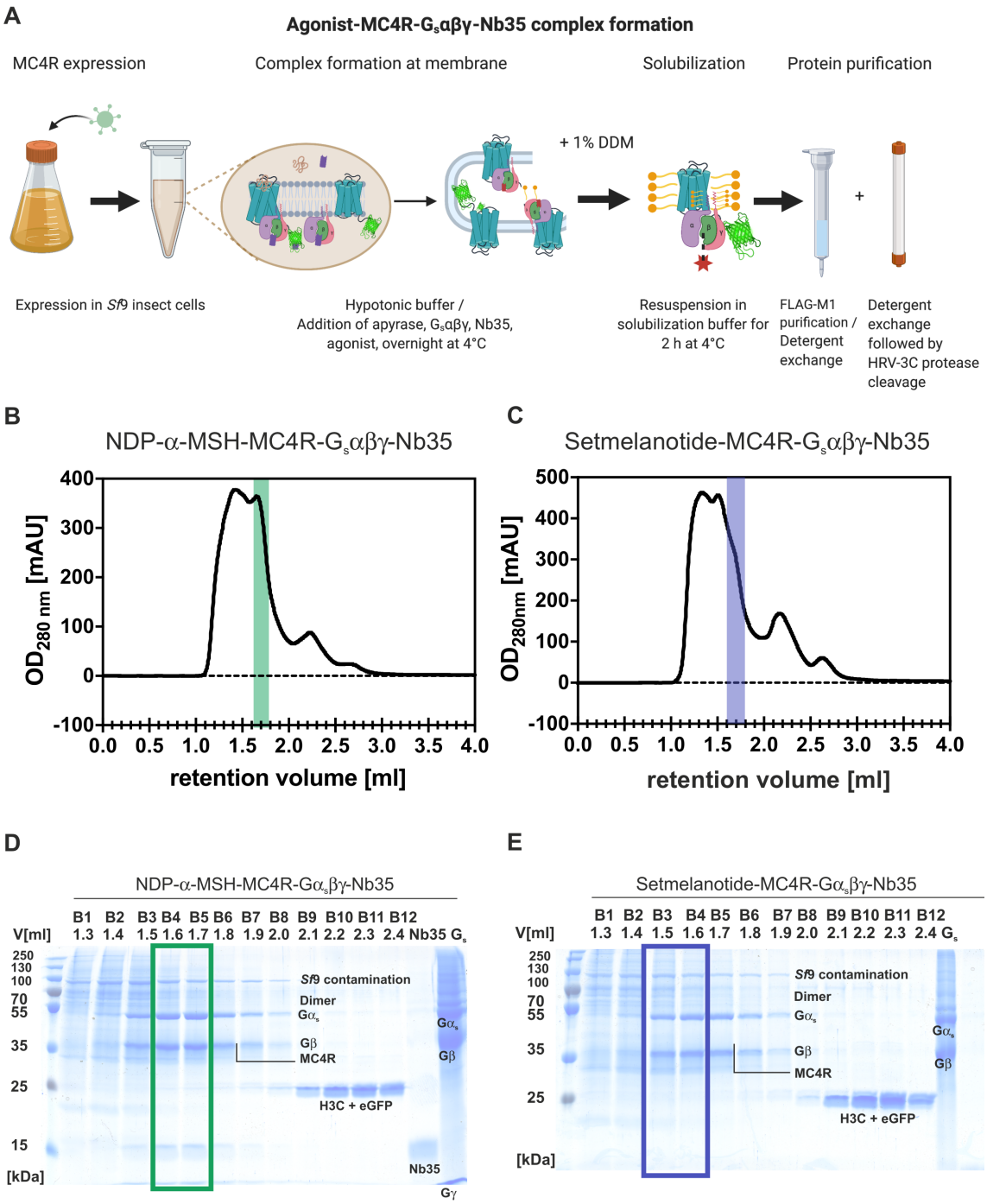

Figure 3.1.3.1 ***Experimental complex formation workflow and the biochemistry data of the complexation experiment.*** (**A**) Workflow for the assembly of the MC4R-G_sαβγ complexes stabilized by NDP-α-MSH or setmelanotide. The agonist-MC4R-G_s-Nb35 complexes were separated from HRV-3C protease and remaining eGFP by size-exclusion chromatography for the complexes with (**B**) NDP-α-MSH and (**C**) setmelanotide. Subsequent gel chromatography of (**D**) NDP-α-MSH-MC4R-G_s-Nb35 and (**E**) setmelanotide-MC4R-G_s-Nb35 were applied to confirm the stoichiometric ratio of MC4R and G_s. G_s-protein and Nb35 were used as controls in the last two lanes. For both complexes, fractions B4 and B5 were concentrated to 5 mg/ml and directly vitrified. Figure **A** was created by biorender.com.

HEPES pH 7.5, 150 mM NaCl, 25 µM TCEP and 1 mM $CaCl_2$. The complex was extracted from membranes by solubilization with 1% DDM and 0.1% CHS. Stable complex formation of M1-FLAG affinity and size exclusion chromatography (SEC) products was confirmed by negative stain EM (Figure 3.1.3.2).

3.1.3.2 MC4R negative stain EM data

Negative stain blotting was done at a concentration of 20 µg/ml, 40 µg/ml, and 80 µg/ml concentration of the NDP-α-MSH-MC4R-eGFP-G_s-Nb35 in a buffer containing 0.1% DDM/CHS on QUANTIFOIL R 2/1 on Cu 300 mesh + 2 nm C grids with 2% uranyl acetate. 20 µg/ml yielded a sufficient particle concentration without overcrowding. However, the presence of detergent micelles resulted in non-interpretable micrographs. The detergent micelle concentration was reduced by washing the grid after sample application three times with buffer without detergent (an example micrograph of a washed negative stain grid is shown in Figure 3.1.3.2B). 296 micrographs were recorded with a Tecnai Spirit 120 keV microscope with a pixel size of 2.67 Å. 54 micrographs were discarded after Ctf correction with CtfFind [179] and the visual inspection of the micrographs. 87.187 particles were picked using gautomatch [180] with a particle diameter of 160 Å. Particles were extracted at full binning size with a particle box size of 90 pixels with RELION [181]. Several rounds of 2D classifications and *in silico* purification yielded 2D classes containing 43,303 particles (Figure 3.1.3.2C). The initial model reconstruction in RELION resulted in a negative stain density indicating the shape of a GPCR-G-protein complex (Figure 3.1.3.2D-E).

The 2D classes resemble the shapes of two attached round blobs, which reflect the shapes of a receptor embedded in a detergent micelle and the G-protein (Figure 3.1.3.2C). In addition, an additional intense bright circular signal was apparent that presumably resembles the eGFP attached to the carboxy terminus of MC4R. The initial 3D reconstruction resulted in a negative stain EM density (Figure 3.1.3.2D-E) that embedded the atomic model of the β2AR-G_s-Nb35 [20] and GFP [199]. Taken together, the 2D classes and the initial negative stain EM density confirmed the successful complex formation of NDP-α-MSH-MC4R-eGFP-G_s-Nb35.

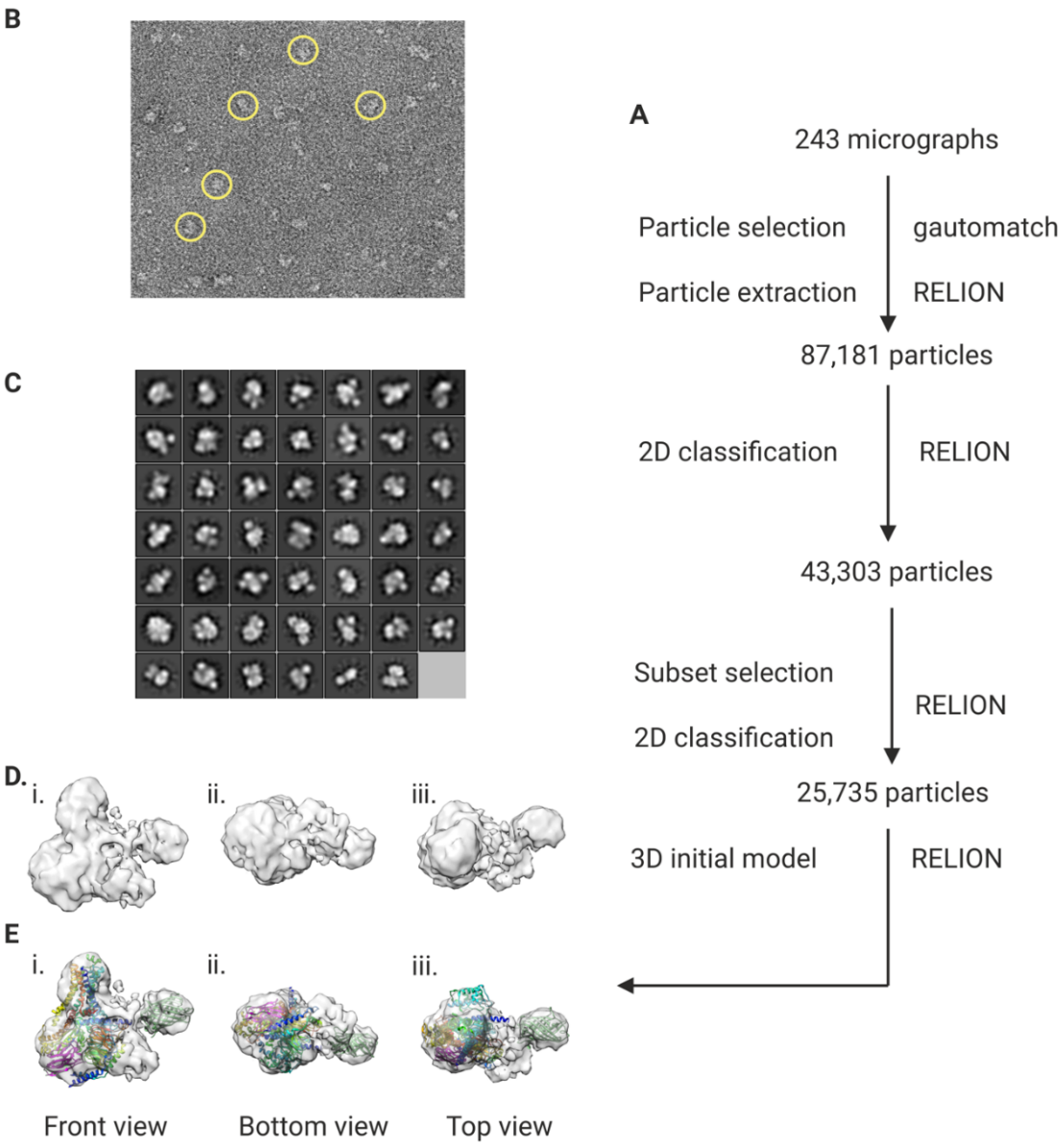

Figure 3.1.3.2 Negative stain EM data and workflow of the NDP-α-MSH-MC4R-eGFP-G_s-Nb35 complex. (A) Flow chart indicates for negative stain EM data processing strategy for MC4R-eGFP in complex with G_s, Nb35, and NDP-α-MSH. **(B)** Particles that are suspected to be receptor G_s-protein complex are highlighted in yellow on a representative micrograph. **(C)** 2D class averages are generated in RELION. **(D)** Negative stain EM density of the initial reconstruction in (*i*) front, (*ii*) bottom, (*iii*) top view, and **(E)** negative stain EM density with the fitting of the atomic models of β2AR-G_s-Nb35 (PDB ID: 3SN6), and GFP (green) (PDB ID: 1GFL) in (*i*) front, (*ii*) bottom, and (*iii*) top view.

3.1.3.3 Sample optimization for cryo-EM experiments

Several cryo-EM experiments were required to find the optimal sample preparation strategy. To verify if the receptor-G_s complexes were sufficiently stable for vitrification experiments, the thermal stability of the receptor complex was assessed by SEC (Figure 3.1.3.3). Therefore, the agonist complex samples were split into three aliquots and were incubated for 10 min at either 4°C, 25°C, or 42°C. The resulting absorption curves did not show significant changes relative to the applied temperature; hence the complexes were sufficiently stable for initial vitrification experiments.

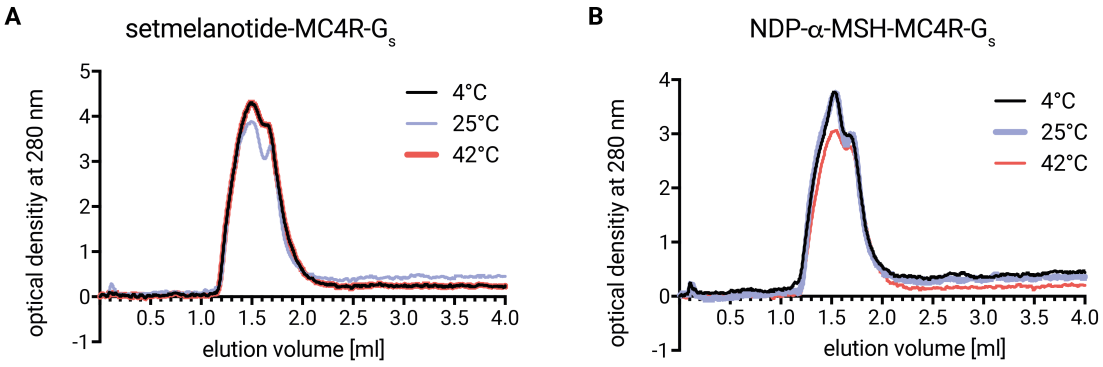

Figure 3.1.3.3 Thermal stability assessment by size exclusion chromatography of (**A**) setmelanotide-MC4R-G_s complex and (**B**) NDP-α-MSH-MC4R-G_s complex. 50 µl of purified samples were incubated at either 4°C, 25°C or 42°C and applied on a Superdex 200 Increase 5/150 GL (Sigma-Aldrich) column. The absorption at 280 nm is plotted over the elution volume.

Although, the thermal stability of the NDP-α-MSH-MC4R-eGFP-G_s complex initial cryo-EM experiments in a buffer containing 0.1% DDM/CHS did not yield sufficient 2D classes (data not shown). The resulting micrographs displayed large amounts of DDM micelles [200]. Therefore, the complex washing protocol at the M1 resin was adapted, and the detergent was exchanged against 0.01% LMNG/CHS. The resulting micrographs were less heterogeneous, but LMNG detergent artifacts were visible (data not shown). The LMNG artifacts were removed by introducing the washing steps of a one-hour rapid dilution step by exchanging buffers with 0.1% DDM/CHS against a buffer with 0.8% LMNG/CHS. A serial dilution to a final concentration of 0.0075% LMNG/CHS and 0.00025% GDN followed [201]. However, removing detergent artifacts did not result in micrographs with a homogeneous particle distribution.

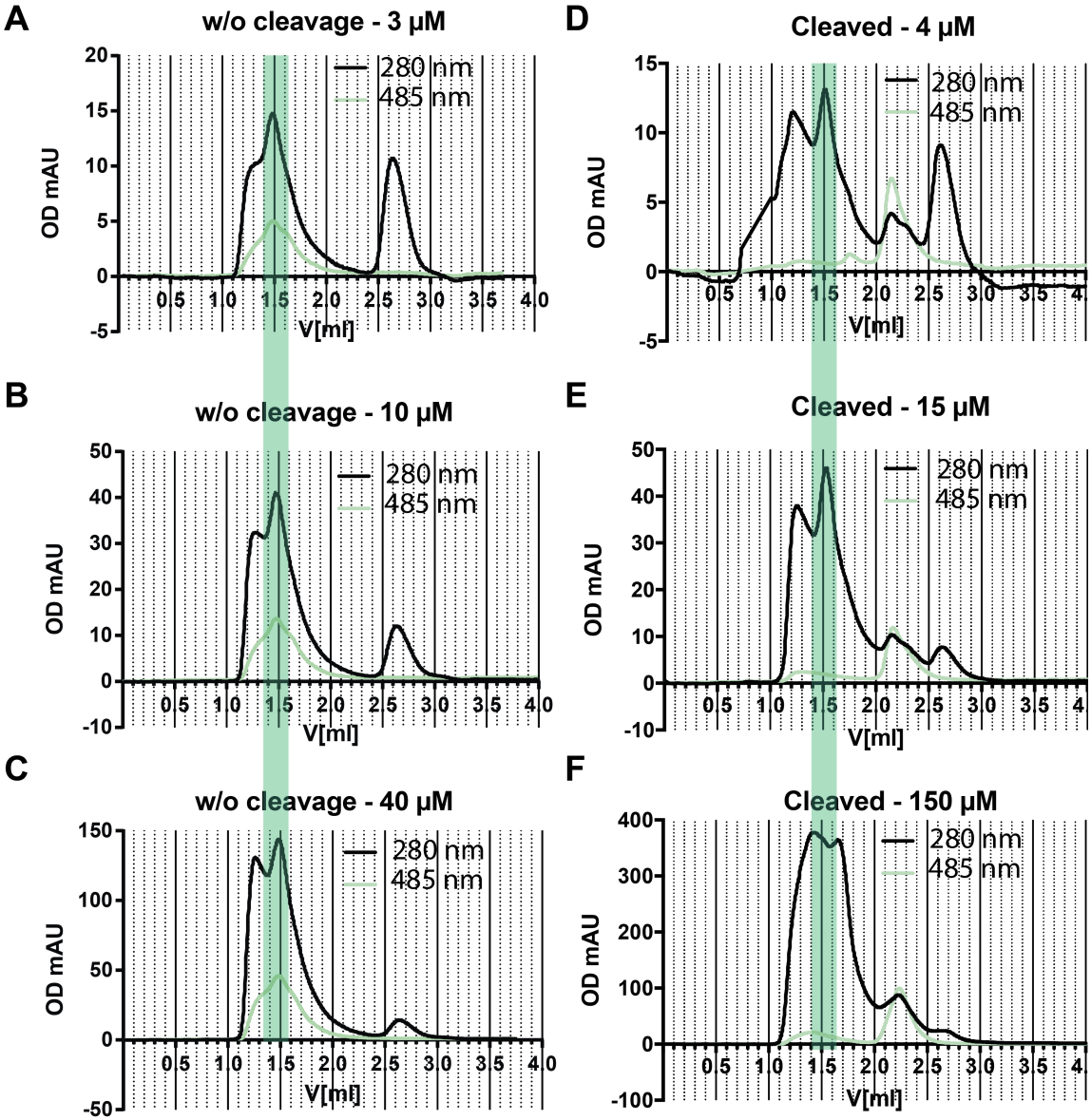

Figure 3.1.3.4 Multimerization comparison with and without proteolytical cleavage of eGFP. SEC runs of (**A-C**) NDP-α-MSH-MC4R-eGFP-Gs-Nb35 without an overnight incubation with H3C protease in varying concentrations and (**D-F**) NDP-α-MSH-MC4R -Gs-Nb35 after overnight incubation with H3C protease. Fractions containing the complex elute at 1.5 ml (highlighted with a green rectangle), and the 1.2 ml peak elutes aggregated receptor. The GFP signal - measured at 485 nm (green) - peaks (**A-C**) at 1.5 ml and in (**D-F**) 2.15 ml elution volume, indicative of the successful proteolytical cleavage of eGFP. The removal of eGFP of MC4R-eGFP by H3C cleavage allows the concentration of the complex sample to ~350 mAU (**F**). The concentration of the sample without the H3C cleavage does not exceed the 150 mAU signal (**C**), and further concentration results in a signal increase around 1.2 ml but not at 1.5 ml (not shown). Concentrations given in the header of the diagrams relate to the overall sample concentration measured prior to the SEC runs.

An additional observation was that protein purification was limited at a concentration threshold of 1 mg/ml since the NDP-α-MSH-MC4R-eGFP-G_s started to aggregate during the concentration step. Sequential SEC runs of the MC4R-eGFP construct resulted in an upper threshold of the complex-containing peak at 1.5 ml elution of 150 mAU. Further concentrating increased the 1.2ml elution-volume peak (Figure 3.1.3.4C). Removing eGFP after affinity purification by H3C protease reduced the aggregation tendency (Figure 3.1.3.4F) and allowed the concentration of the complex containing peak to 350 mAU. SEC and gel electrophoresis confirmed a high protein yield and the stoichiometric ratio of MC4R and G_s (Figure 3.1.3.1B,D). The grid preparation by vitrification of an NDP-α-MSH-MC4R-G_s complex without eGFP resulted in micrographs displaying homogeneous particles (see example micrograph in Figure 3.1.4.1A) distributed in random orientations (Figure 3.1.4.1B). The same complex formation and purification strategy were applied for the setmelanotide-MC4R-G_s complex (Figure 3.1.3.1C,E).

3.1.4 Structure determination of agonist-MC4R-G_s complexes by cryo-EM

3.1.4.1 Optimization of specimen preparation and data processing

NDP-α-MSH-MC4R-G_s complex after removing eGFP was vitrified on Quantifoil R1.2 R1.3 300 Cu and UltrAufoil R1.2/1.3 300 Au grids. The grids were imaged in a Titan Krios microscope, and 4818 micrographs (Quantifoil) and 5625 micrographs (UltrAufoil) were recorded. The micrographs of the UltrAufoil grid yielded significantly more particles than the Quantifoil micrographs (data not shown). The data processing was done using the UltrAufoil micrographs solely. As a result, the micrographs display non-aggregated particles in random distributions (Figure 3.1.4.1A-B).

The data of these micrographs were processed using (i) the cryoSPARC suite and (ii) both RELION 3.01 [182] and cryoSPARC. This comparison examined if the faster cryoSPARC suite yields comparable data quality as RELION. A difference between both software packages is that RELION, in contrast to cryoSPARC, offers Bayesian polishing [202].

The data processing pipeline (i) using cryoSPARC is described in the methods Section 2.2.3. The *in-silico* purification resulted in (i) 221,682 particle coordinates of preselected particles were used for the refinements using cryoSPARC. They yielded a final

resolution after non-uniform (NU) refinement [183] of 2.86 Å. The cryoSPARC selected particle coordinates were used for refinement strategy (ii).

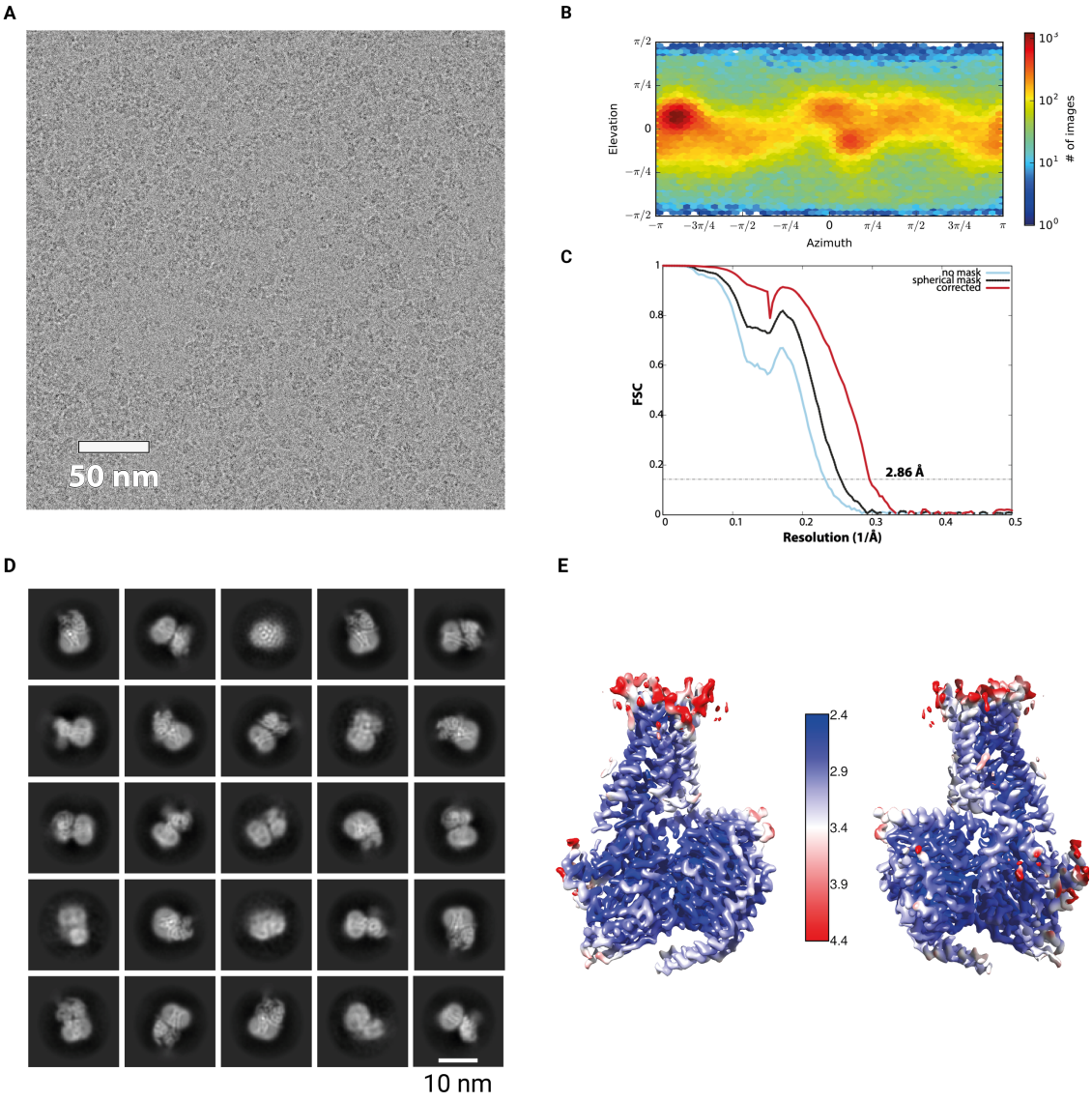

Figure 3.1.4.1 ***Cryo-EM data analysis of the active NDP-α-MSH-MC4R-Gs-Nb35 complex dataset***. (**A**) Representative cryo-electron micrograph. The scale bar corresponds to 50 nm in the image. (**B**) Distribution of projection directions as estimated during homogeneous refinement with program cryoSPARC. (**C**) Global resolution estimation by Fourier shell correlation calculations (FSC = 0.143 cutoff) after "gold standard" refinement. The light blue curve was calculated without masking, the black curve by applying a spherical mask, and the red one after phase randomization using a soft mask. (**D**) Representative 2D class averages confirm random distribution of projection directions. (**E**) Representation of local resolution estimation determined with cryoSPARC. The final cryo-EM density is colored according to the local resolution ranging from dark blue (2.4 Å) to red (4.4 Å).

Data processing using RELION

Movies were motion-corrected using MotionCor2 [203] and Ctf corrected using Gctf [204] with a pixel size of 0.832 Å. 221,682 unbinned particles were extracted with a particle diameter of 190 Å. After two rounds of 2D classification, all particles were selected for high-resolution 3D refinement (Refine3D) in RELION, yielding a resolution of 3.8 Å without masking. Receptor masking with a contour level at 0.013 and post-processing resulted in a resolution of 3.6 Å. Five rounds of CTF refinement (CTF refine), Refine3D, Bayesian polishing (Polish), Refine3D result in a 3.5 Å resolution without masking and the post-processing with masking at 3.0 Å resolution. During the last two iterations, the resolution did not improve. The polished and CTF particles were subjected to NU refinement using cryoSPARC, yielding a resolution of 2.9 Å. Two additional Polish and NU refinement iterations did not significantly improve the resolution. The visual inspection of the maps resulting from data processing strategy (i) and (ii) did not display differences in quality (data not shown). The CPU-based Polish jobs exceed 20 hours for ~200.000 particles. The RELION based pipeline, including rounds of Polish, followed by Refine3D, required ten days of computational time compared to less than a day for the GPU-based cryoSPARC suite.
Consequently, data processing using RELION did not yield a higher resolution map than cryoSPARC for the NDP-α-MSH-MC4R-G_s dataset, and the setmelanotide-MC4R-G_s dataset was processed using cryoSPARC. The processing pipeline for the NDP-α-MSH-MC4R-G_s dataset is visualized in Figure 2.2.3.1, and the processing pipeline for the setmelanotide-MC4R-G_s dataset is shown in Figure 2.2.3.2 and described in detail in Section 2.2.3.

The 2D classes generated by both datasets already revealed secondary structure elements of the receptor and the G-protein (Figure 3.1.4.1D and 3.1.4.2D). The processing of the micrographs resulted in cryo-EM maps with a global resolution of 2.9 Å for NDP-α-MSH-MC4R-G_s and 2.6 Å for setmelanotide-MC4R-G_s (Global resolution estimation by Fourier shell correlation calculations (FSC = 0.143 cutoff) after "gold standard" refinement are shown in Figure 3.1.4.1C and 3.1.4.2C).

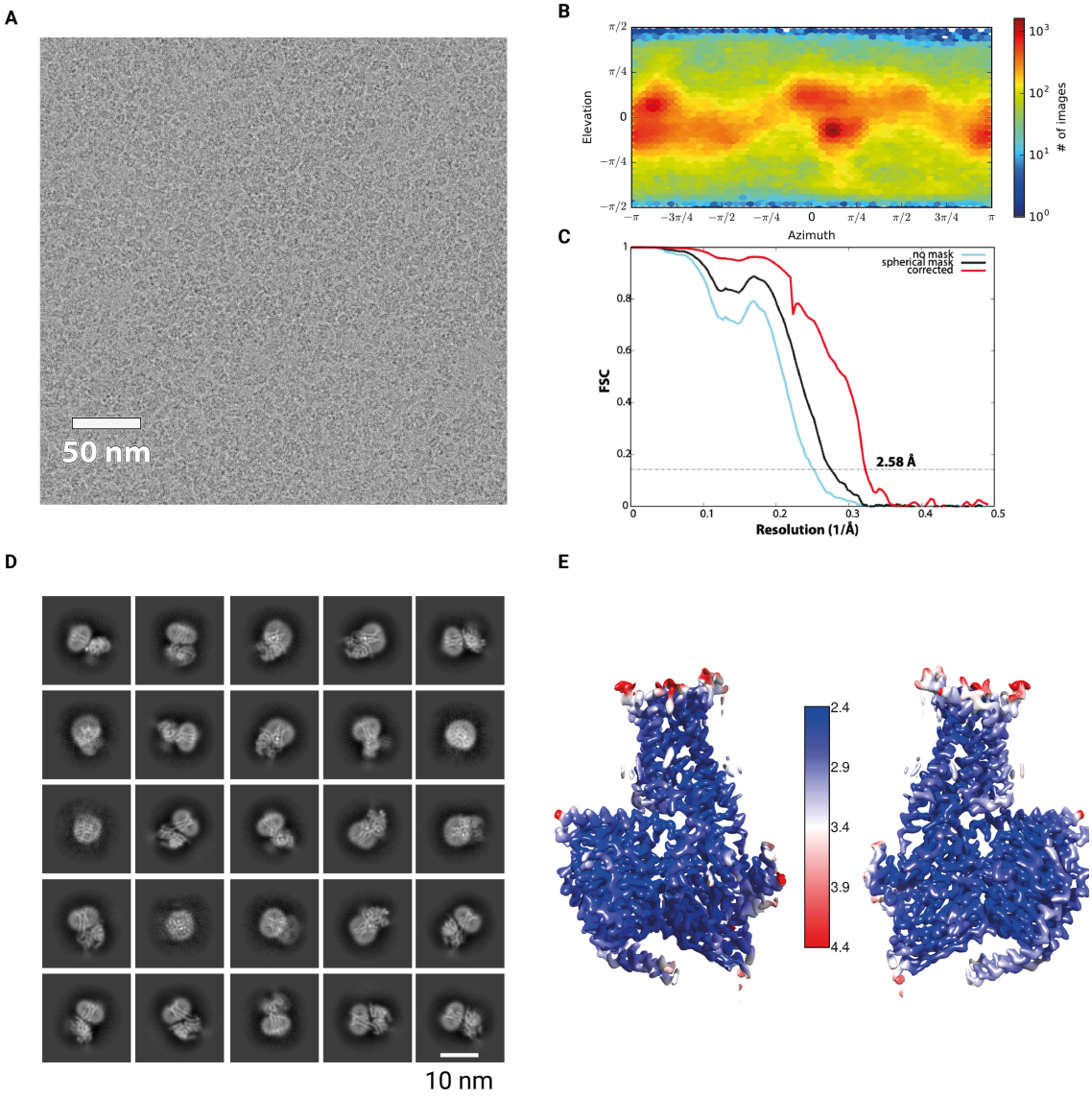

Figure 3.1.4.2 ***Cryo-EM data analysis of the active setmelanotide-MC4R-G_s-Nb35 complex dataset.*** (**A**) Representative cryo-electron micrograph. The scale bar represents 50 nm in the image. (**B**) Distribution of projection directions as estimated during homogeneous refinement with cryoSPARC. (**C**) Global resolution estimation by Fourier shell correlation calculations (FSC = 0.143 cutoff) after "gold standard" refinement. The light blue curve was calculated without masking, the black curve by applying a spherical mask, and the red one after phase randomization using a soft mask. (**D**) Representative 2D class averages confirm the random distribution of projection directions. (**E**) Representation of local resolution estimation determined with cryoSPARC. The final cryo-EM density is colored according to the local resolution ranging from dark blue (2.4 Å) to red (4.4 Å).

3.1.4.2 Model building and refinement

The high-resolution cryo-EM maps allowed us to generate receptor models for the entire receptor structure, except already known flexible domains at the N and C-terminus, IL1, IL3, and EL3. The models were derived from the inactive MC4R structure (PDB ID: 6w25 [91]) together with the $G_s\alpha\beta\gamma$-Nb35 complex of the β2-adrenergic receptor-G_s complex (PDB ID: 3sn6 [19]) and manually refined in COOT. The coordinates of the ligands were built *de novo* using COOT. Real-space refinement was performed with the program PHENIX [185] using geometric restraints, a global minimalization protocol and B-factor refinement. In addition, both models were refined with isotropic B-factors in reciprocal space using REFMAC5 [186]

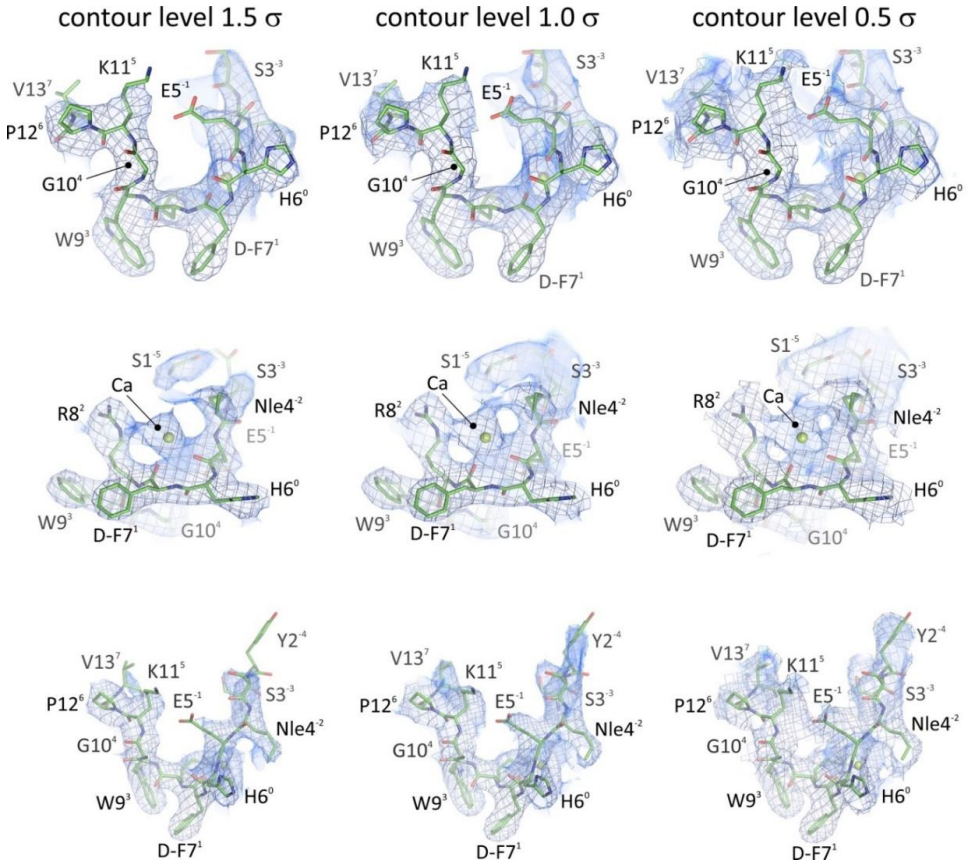

Figure 3.1.4.3 ***Quality of the cryo-EM density map of the NDP-α-MSH ligand of NDP-α-MSH-MC4R-G_s-Nb35 complex.*** Three different views (top to bottom) on the NDP-α-MSH ligand (green color) together with the coordinating calcium ion (lemon green color). All figures show cryo-EM densities of the ligand (light blue colored mesh/volume) contoured at three different contour levels (1.5 (left), 1.0 (middle), and 0.5 σ (right)). NDP-α-MSH is depicted as sticks and the calcium ion as sphere representation.

The G-protein was modeled except the SWIII, αG, α5, and β6α5 loops and the $G\alpha_s$ α-helical domain (AHD). The AHD is highly flexible following nucleotide hydrolysis and is not represented by high-resolution data, and the EM-map could not be used to model the AHD. However, the density for Ca^{2+} and the peptidic ligands, setmelanotide, and NDP-α-MSH could be modeled unambiguously (Figure 3.1.4.3 and 3.1.4.4). The assignment of the vacant electron density at the peptide-TM2-TM3 interface to a Ca^{2+} ion is discussed in Section 4.1.3.

The final model of the NDP-α-MSH-MC4R complex includes the following amino acids (based on the final overall cryo-EM map); MC4R: 40-108; 118-230, 239-316; NDP-α-MSH: 1-13; $G\alpha_s$-protein: 13-47; 194-236; 249-280; 293-306; 322-380; $G\beta_1$: 3-340; $G\gamma_2$: 9-63; Nb35: 1-128; and 115 water molecules.

The final model of the setmelanotide-MC4R complex includes the following amino acids (based on the final overall cryo-EM map); MC4R: 40-107; 117-230, 240-316; setmelanotide: 1-8; $G\alpha_s$-protein: 14-47; 193-236; 248-280; 293-310; 318-351; 355-380; $G\beta_1$: 4-340; $G\gamma_2$: 9-63; NB35: 1-128; and 88 water molecules. The quality of the models is represented in the statistics. Both models have poor rotamers below 1%, clash scores below 3.6, and a Molprobity score below 1.6. The Ramachandran plot gives 0.21% for the NDP-α-MSH, and 0.11% disallowed for the setmelanotide stabilized complex (Table S1).

The side chains of NDP-α-MSH are well defined by the cryo-EM map at the $H6^0F7^1R8^2W9^3$ motif, and the cryo-EM map is less defined at the flexible ends of the peptide (such as the side chain of Y2) (Figure 3.1.4.3**)**. Setmelanotide is smaller, less flexible, and better resolved at 2.6 Å. Although the setmelanotide disulfide bond was probably damaged by radiation, the density of the disulfide bond has a lower signal-to-noise level than the peptide backbone (Figure 3.1.4.4). The high-resolution maps, especially of the setmelanotide-MC4R-G_s complex, allowed the confident placing of water molecules (Table S1). Overall, the structures allow the interpretation of agonist action, receptor activation, and G_s-protein binding.

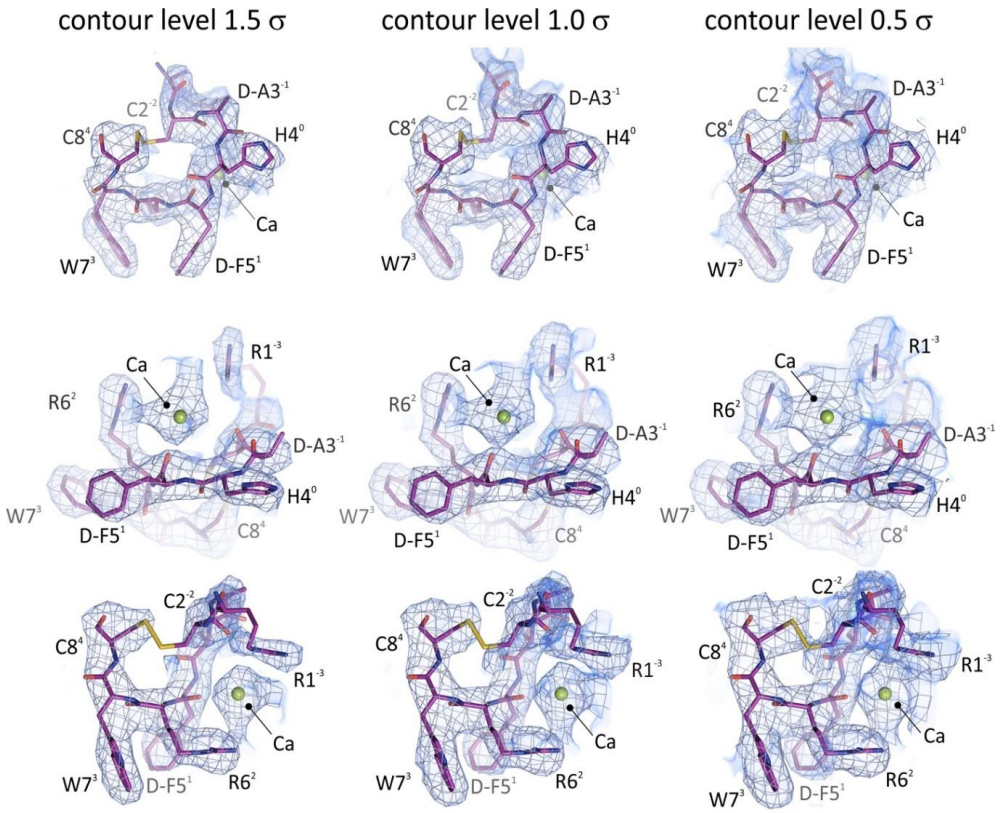

Figure 3.1.4.4 ***Quality of the cryo-EM density map of setmelanotide of the setmelanotide-MC4R-G_s-Nb35 complex.*** Three different views (top to bottom) on the setmelanotide ligand (purple color) and the binding calcium ion (lemon green color) are shown. All figures show the cryo-EM densities of the ligand (light blue colored mesh/volume) contoured at three different contour levels (1.5 (left), 1.0 (middle), and 0.5 σ (right)). Setmelanotide is depicted as sticks and the calcium ion as sphere representation.

3.2 Overall description of the Agonist-MC4R-G_s structures

3.2.1 Overall complex structures

The models generated from both cryo-EM maps (Figures 3.2.1.B-C) encompass for each model the heterotrimeric G_s-protein, $G_s\beta_1\gamma_2$, bound to the intracellular G-protein binding cavity (GBC) of MC4R, which is formed by the TM6 opening upon receptor activation (Figure 4.1). G_s binding to MC4R is stabilized by nanobody 35 (Nb35) at the interface of $G\alpha_s$ and $G\beta\gamma$. MC4R consists of the GPCR typical 7 TM bundle linked by three intracellular loops (ICLs) and three extracellular loops (ECLs). The cytoplasmic helix 8 follows directly after TM7 and parallel the membrane plane. All seven TM and the connecting loops form the ligand-binding pocket (LBP) or orthosteric site at the extracellular site. Here, both agonists face the $H^0x^1R^2W^3$ motif (Section 1.3) between the ELs and inside the seven-transmembrane bundle of the receptor (Figure 3.2.1D). The ligands interact with MC4R through extensive hydrophobic, van der Waals, and polar interactions, with residues in the TMs as well as EL2 (Table 3.5.1). Specific structural MC4R characteristics - discussed in the next section - result in a wide-opened extracellular LBP for large peptidic ligands.

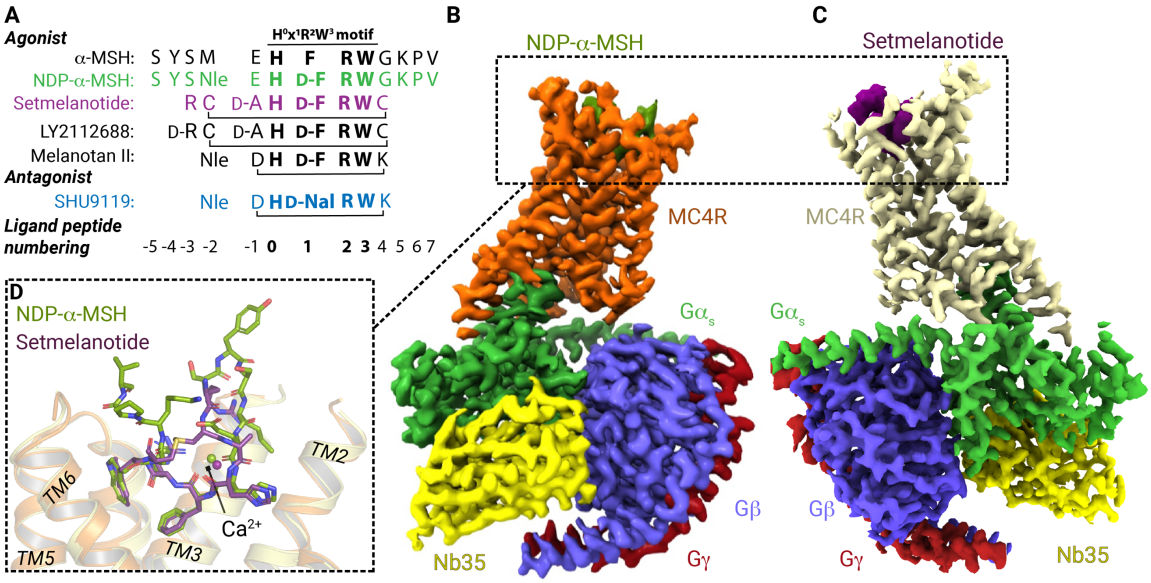

Figure 3.2.1 NDP-α-MSH and setmelanotide bound cryo-EM complex structures of MC4R-G_s. (**A**) Sequence alignment of MC4R agonists and antagonist SHU9119. (**B-C**) Cryo-EM densities of MC4R-G_s complexes stabilized by bound agonists and Nb35 displayed from mirrored perspectives; Gβ, dark blue; Gγ, red; Nb35, yellow. (**D**) Superposition of both structures, focusing on peptide agonist (stick representations).

3.2.2 MC4R forms a wide-opened orthosteric site

MC4R has a large, wide-opened extracellular LBP to accommodate large peptidic ligands, supported by three MC4R-specific sequence characteristics.

(i) MC4R EL2 is only four amino acids long and does not include a conserved class A GPCR at position 3.25 [205]. Therefore, MC4R's orthosteric site is not stabilized by a disulfide bridge between EL2 and TM3. The residue at TM3 that engages in that disulfide bond, $C^{3.25}$, is $D122^{3.25}$, a central player in ligand and Ca^{2+} binding (Figure 3.3.2).

(ii) Secondly, MC4R has two disulfide bonds at the extracellular receptor site. The disulfide bridge between $C271^{EL3}$-$C277^{EL3}$ displays a helical conformation of the EL3 to TM7 transition. This cystine has been shown for two lipidic GPCRs (Figures 3.2.2A-B). Functional studies have confirmed the relevance of this disulfide bond [206]. The second observed MC4R disulfide bridge connects $C40^{Ntt}$ in the N-terminus and $C279^{7.30}$ in EL3 (Figure 3.2.2C). The 39 residues upstream of C40 were not modeled due to inherent flexibility resembling missing cryo-EM density. This disulfide bridge connects TM7 with TM1 via the N-terminus, stabilizing the following residues and the wide-open orthosteric site. This cysteine bridge between the N-terminus and the EL3 was investigated in several studies on MCRs by mutagenesis of the potentially involved cysteines [206,207]. The results and conclusions on the existence of this disulfide bridge were not always in line. However, the complex structures here clarify the existence of this disulfide bond.

(iii) MC4R has no proline in TM2, unlike most other class A GPCRs (Figure 3.2.2D). Instead, TM2 displays at $S99^{2.59}$ a straight helix and a ~6 Å shift of the extracellular TM2 towards the membrane compared to β2AR and other structures (Figure 3.2.2D).

An apparent separated density for Ca^{2+} is found in the cryo-EM densities of the agonist bound structures (Figures 3.1.4.3 - 3.1.4.4). Ca^{2+} binding to MC4R was first described in the antagonist SHU9119-MC4R crystal structure [91] (Figure 3.3.4). Moreover, its relevance has been indicated in proceeding studies (Section 1.3.3). The cofactor calcium and the extended extracellular vestibule define a unique LBP, which appears to be adapted to integrate signals from ligands with differences in size.

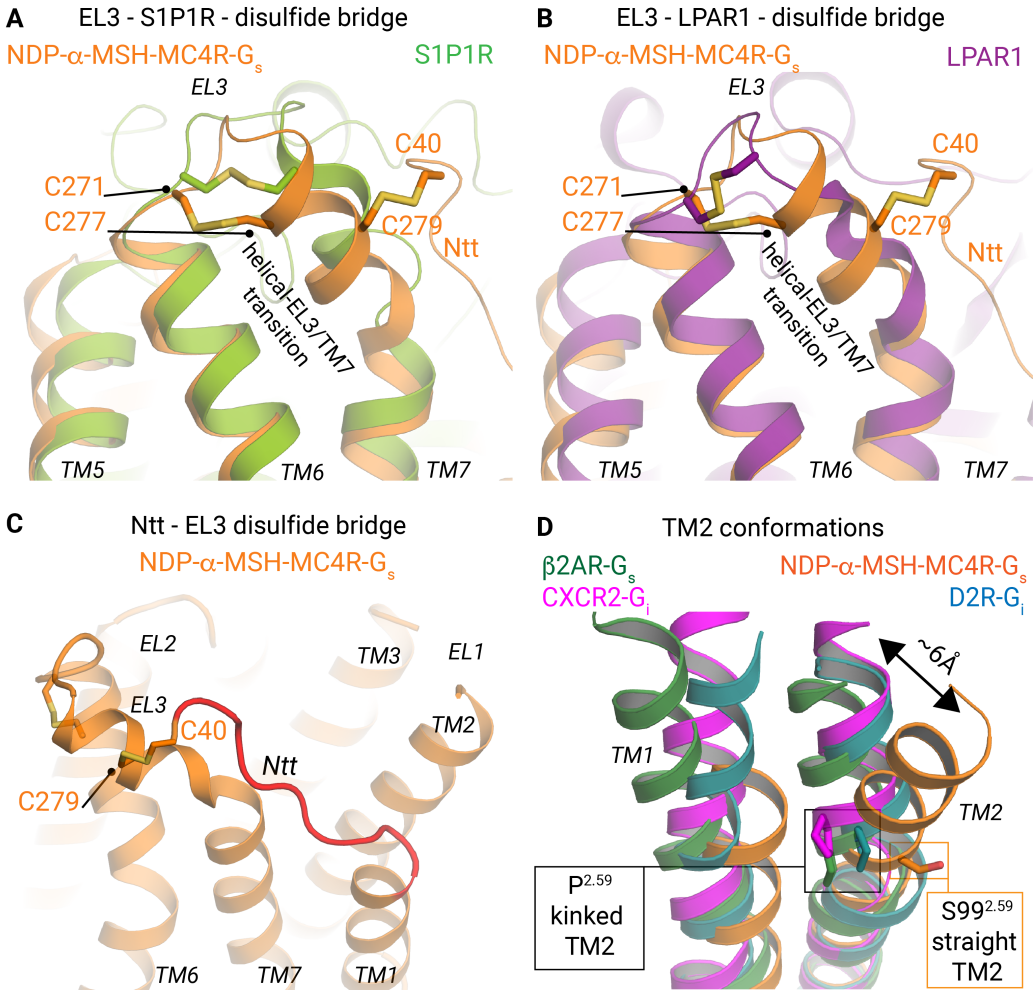

Figure 3.2.2 ***MC4R specificities in the extracellular region.*** The active state MC4R structure (orange) shares a disulfide bridge in the EL3, which is involved in forming a helical transition to TM7 with (**A**) the sphingosine 1-phosphate receptor 1 (S1P1R, light green, PDB ID: 3v2y [208]) and (**B**) the lysophosphatidic acid receptor 1 (LPAR1, deep purple, PDB ID: 4z34 [209]). (**C**) The MC4R has a second disulfide bridge between the receptor N-terminus (Ntt) and EL3 that constrains the extracellular helix close to TM7. (**D**) MCRs have no proline in TM2 at position 2.59, contrary to most other class A GPCR, which usually causes a slight rotation of TM2 toward the extracellular region. The superposition of the active state MC4R structure with GPCR-G-protein complex structures (β2AR, forest green color, PDB ID: 3sn6 [20]; D2R, blue color, PDB ID: 6vms [210]; CXCR2, magenta color, PDB ID: 6lfo [211]) illustrates deviations in TM2 positioning (5.9 Å distance measured between position 2.65 in MC4R and β2AR).

3.3 Ligand binding to MC4R

3.3.1 MC4R agonist binding features

Comparing the antagonized MC4R structure in complex with the antagonist SHU9119 [91] with the two MC4R structures bound to the ligands NDP-α-MSH and setmelanotide shows that all ligands bind deeply in the extracellular receptor vestibule between the TMs. Furthermore, the binding of the ligands results in different shaped orthosteric ligand-binding sites (Figure 3.3.1).

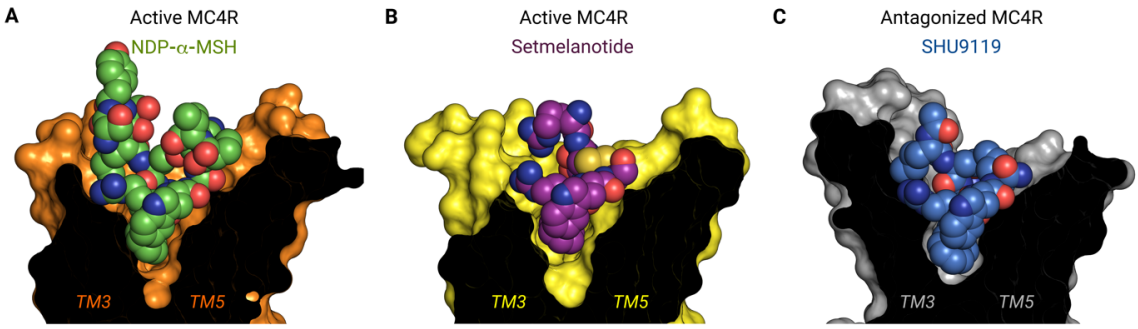

Figure 3.3.1 ***Peptidic MC4R ligands bind into the extracellular ligand-binding pocket.*** Ligand-MC4R structures of **(A)** NDP-α-MSH (carbon atoms; green), **(B)** setmelanotide (carbon atoms, magenta), and **(C)** SHU9119 (blue carbon atoms) bound to the MC4R ligand-binding pocket (sphere representation, clipped surface, orange, light-yellow, gray, respectively).

These ligand-dependent shapes are caused by varying primary ligand sequences (Figure 3.2.1A). SHU9119 and setmelanotide are cyclic and more compact than the linear peptide NDP-α-MSH (Figure 3.3.2).

All three ligands share a common central amino acid motif, $H^0x^1R^2W^3$ (Section 1.3). Most relevant interactions within the LBP are formed via this motif. The x^1 position is located at the bottom part of the LBP (Figures 3.3.2A-B). In contrast to the endogenous ligands, the high-affinity agonists NDP-α-MSH and setmelanotide have the stereoisomer D-phenylalanine (D-F^1) instead of an L-F at position x^1.

Both agonists form similar interactions with the $H^0x^1R^2W^3$ motif, including hydrogen bonds between H^0 and $T101^{2.61}$ in TM2, W^3 with $S188^{EL2}$ in EL2, and $H264^{6.54}$ in TM6

(Figures 3.3.2C-D). In addition, they share similar hydrophobic contacts, H^{0} to $F284^{7.35}$, $L288^{7.39}$, $F51^{1.39}$; D-F^{1} to $I129^{3.32}$; and W^{3} to$Y268^{6.58}$, $I194^{5.40}$, and $L197^{5.43}$ (Table 3.5.1). In both agonist-bound structures, a D-F^{1} at position x^{1} stabilizes Ca^{2+} through the main chain (Figures 3.3.2C-D) while the side chain points into the core of the receptor. Here, D-F^{1} is surrounded by hydrophobic amino acids, namely $C130^{3.32}$, $L133^{3.35}$, $I185^{4.61}$, $L197^{5.34}$, $F261^{6.51}$, and $L288^{7.39}$. The hydrophobic residues $I291^{7.42}$ and $L133^{3.35}$ are in close vicinity of $W258^{6.48}$ of the highly conserved *CWxP*$^{6.50}$ motif in TM6 that is in contact with ligands in several class A GPCRs (*e.g.*, muscarinic acetylcholine receptor-G_{11} complex [45], 5-HT2A serotonin receptor-G_q complex [56]). In contrast, the MC4R agonists do not interact with $W258^{6.48}$ (Figures 3.3.2A-B).

3.3.2 Unique features of both agonist-binding modes

Table 3.5.1 summarizes all ligand contacts to the receptor with the distance threshold of 3.9 Å, and in the following section, unique agonist characteristics are described. NDP-α-MSH deviates from the endogenous agonist α-MSH by two substitutions, M4^{-2}Nle, and L-F7^{1}D-F. Nle4^{-2} is close to $I104^{2.64}$ and $D122^{3.25}$. In contrast to the setmelanotide-complex, $D122^{3.25}$ is directed towards Nle4^{-2}, and only one acidic head group faces the Ca^{2+} (Figure 3.3.2).

NDP-α-MSH's D-F7^{1} points into the receptor core surrounded by $D126^{3.29}$, $I129^{3.32}$, $L288^{3.36}$, and Ca^{2+}. The backbone oxygen of D-F7^{1} points towards Ca^{2+} and is in hydrogen-bond distance. In addition, the peptide backbone of E5^{-1} further coordinates Ca^{2+}; this interaction has no equivalent in the setmelanotide-receptor structure. NDP-α-MSH's unique contacts are with $I185^{4.61}$ at TM4 and $V193^{5.39}$ at TM5. The NDP-α-MSH sidechain of R8^{2} connects EL2, TM4, and TM3 and forms a tight agonist-receptor interface. The sidechain of R8^{2} is in hydrogen bond distance with $S188^{EL2}$, $I185^{4.61}$, and $D126^{3.29}$. Setmelanotide shares the HD-FRW core interaction motif with NDP-α-MSH. Upstream of this motif are three residues a D-R1^{-3}, C2^{-2}, and D-A^{-1}. The C2^{-2} forms a disulfide bond with the second cysteine, C8^{4}, one residue downstream of the core motif. Because of this disulfide bond, setmelanotide is slightly more compact than NDP-α-MSH. D-F5^{1} points in accordance with NDP-α-MSH's binding mode into the receptor core and is in the surround-

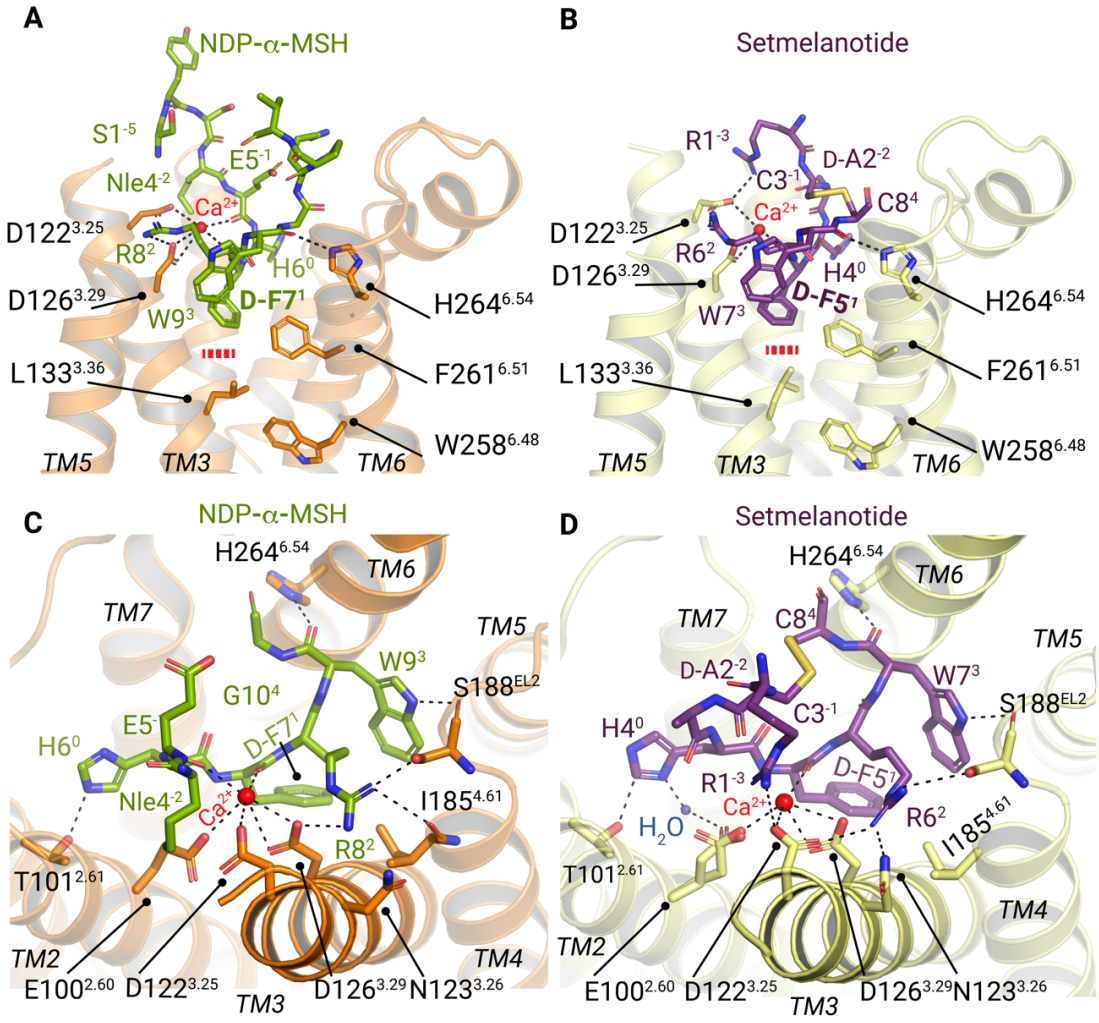

Figure 3.3.2 Peptidic agonists bind to MC4R in the presence of a Ca^{2+} ion. (**A-D**) NDP-α-MSH (stick representation of carbon atoms; green) and setmelanotide (stick representation of carbon atoms, magenta) are bound to MC4R (orange and light-yellow, respectively). (**A-B**) Membrane plane view on the orthosteric binding site of MC4R. TM3 and TM6 receptor residues stabilize the peptides and Ca^{2+}. $L133^{3.36}$ and agonist residue $D\text{-}F^{1}$ do not interact (indicated by the red dashed lines). Top view of (**A-B**), highlighting all ligand-receptor hydrophilic interactions, assessed by a minimum distance of 3.5 Å (black dashed lines). (**C**) NDP-α-MSH residues 4-11 are shown for simplicity.

ing of E100$^{2.60}$, D126$^{3.29}$, I129$^{3.32}$, L133$^{3.36}$, F261$^{6.51}$, and Ca^{2+}. Unique contacts for setmelanotide are N123$^{3.26}$, L133$^{3.36}$ at TM3, F261$^{6.51}$ at TM6, and N285$^{7.36}$ at TM7. Here, the additional hydrogen bond of R6^{2} to N123$^{3.26}$ and no interaction to I185$^{4.61}$ at TM4 indicates a subtle clockwise rotation compared to NDP-α-MSH (Figure 3.3.3).

The cryo-EM density allowed the modeling of two E100$^{2.60}$ rotamers. One directs towards Ca^{2+} and the other to a water molecule in hydrogen bond distance to the imidazole ring of H^{0} that is as well a double rotamer conformation. Here, one rotamer is in hydrogen bond distance to T101$^{2.61}$, and the second rotamer coordinates the water bridge to E100$^{2.60}$. In addition, setmelanotide forms with D122$^{3.25}$ two potential hydrogen bonds via R1^{-3} and R6^{-2}. This arginine clamp is unique to setmelanotide. On the other hand, NDP-α-MSH residues are not in hydrogen bond distance with D122$^{3.25}$. The acylated S1^{-5} of NDP-α-MSH potentially directs here as well, but D122$^{3.25}$ has shifted away from S1^{-5}, and they are at a minimum distance of 3.5 Å.

Overall, the ligand-binding mode of both agonists shows subtle differences in the interaction with the Ca^{2+} binding interface at TM3.

3.3.3 Both agonists adjust TM3 in concert with Ca^{2+}

The subtle differences of the agonist binding modes result in adjustments of the MC4R-LBP. Of note, the superposition of both agonist-MC4R complex structures (Figure 3.3.3) display a subtle clockwise rotation of setmelanotide towards NDP-α-MSH inside the orthosteric pocket (perspective according to a top-view on the LBP). As described in the section above, the main differences in the agonist binding modes are the deviating interaction pattern with TM3 via N123$^{3.26,}$ TM4 via I185$^{4.61}$, and the additional hydrogen bond of R1^{-3} in setmelanotide with D122$^{3.25}$ (Figures 3.3.2C-D).

On the receptor side, TM2 and TM3 are shifted on the membrane or horizontal plane. The co-factor Ca^{2+} accompanies the ligand-receptor interface and connects the ligand, TM2 (E100$^{2.60}$) and TM3 (D122$^{3.25}$ and D126$^{3.29}$). TM3 is shifted relatively by more than 1 Å inside the LBP for setmelanotide-MC4R, and the residues D122$^{3.25}$, D126$^{3.29}$, and N123$^{3.26}$ at TM3 follows the overall positioning. D122$^{3.25}$ in the NDP-α-MSH structure faces only one oxygen towards Ca^{2+}, in contrast to the setmelanotide bound structure, in which the

total carboxy group faces the ion. In addition, N123$^{3.26}$ displays the most extensive side-chain rearrangement among both structures besides the additional hydrogen bond in the setmelanotide MC4R structure. TM2 accompanies the positioning of TM3 and Ca^{2+}, here E100$^{2.60}$ displays varying conformations. In summary, the presented two agonist-bound structures display a horizontal adjustment of TM3 accompanied by Ca^{2+}, potentially driven by ligand interactions with TM3 and TM4.

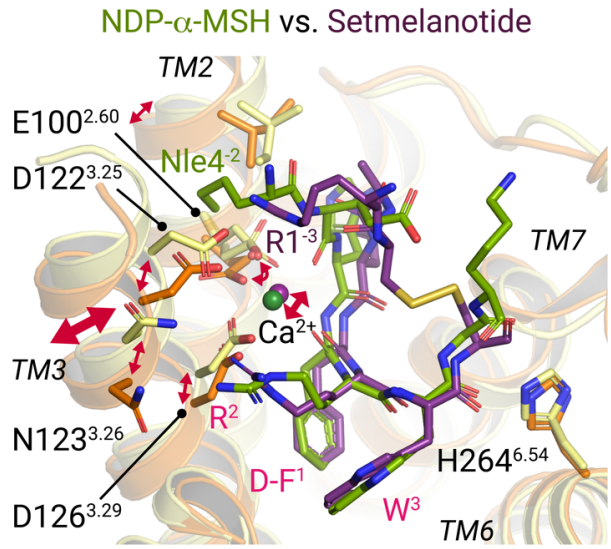

Figure 3.3.3 ***Agonist-bound structures display a relative horizontal shift of TM3 accompanied by Ca^{2+}.*** Superpositions of NDP-α-MSH (residues 4-11) and setmelanotide bound at MC4R (orange and yellow backbone cartoons). Superposition of both agonists. Residues shared by both agonists are labeled in pink. Red arrows indicate relative transpositions of residues and TM3 among both agonist-bound structures.

3.3.4 Ligand binding mode of antagonist SHU9119 bound to MC4R

The SHU9119-MC4R structure will be briefly shown here to introduce the high similarity with the agonist binding mode. The structure will be relevant for the discussion since the antagonized state is needed to understand the activation mechanism. In contrast, the active-state structures allow unraveling the molecular mechanism of SHU9119 antagonism (Section 4.1.2).

In 2020, the antagonistic peptide SHU9119 was described in the crystal structure of inactivated SHU9119-bound MC4R [91]. The peptide is a circularized (lactam bond: D2^{-1} and K7^{4}) variant of NDP-α-MSH [133] (Figure 3.2.1A). SHU9119's binding mode is like NDP-α-MSH's. The hydrophobic contacts in the central four-finger H^{0}x^{1}R^{2}W^{3} motif are in accordance with NDP-α-MSH's contacts (Table 3.5.1). W6^{3} is counter-anchored by H264$^{6.54}$ (TM6) and S188^{EL2} (EL2). R5^{2} connects TM3, TM4, and EL2 through hydrogen bonds to S188^{EL2} and I185$^{4.61}$ (Figure 3.3.4). However, R5^{2} does not interact with D126$^{3.29}$

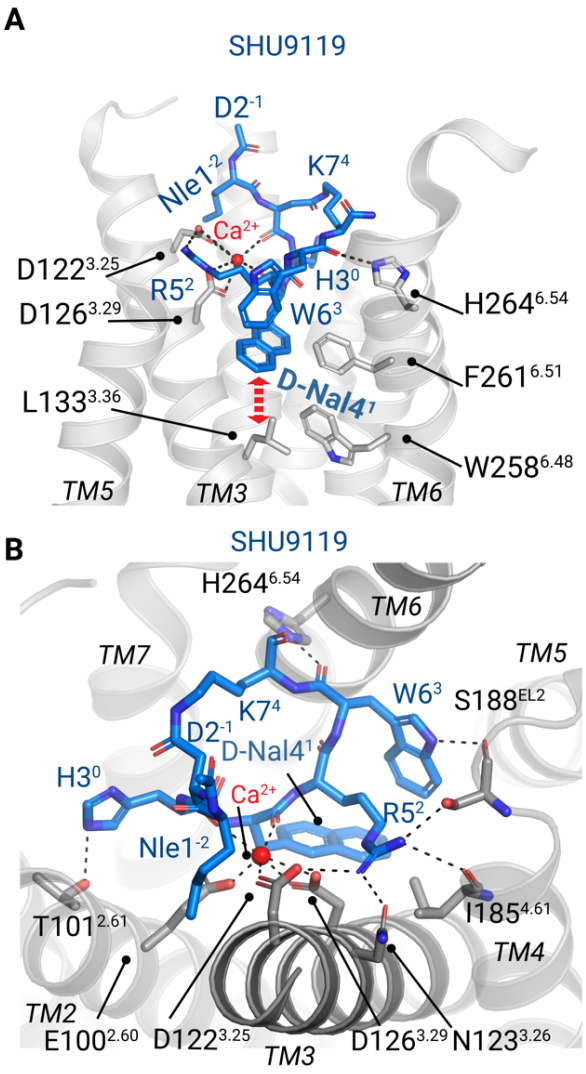

Figure 3.3.4 ***Antagonist SHU9119 and Ca^{2+} ion binding to MC4R.*** (**A-B**) Stick representation of SHU9119 (carbon atoms; blue) bound into MC4R binding site (grey). (**A**) Side view (from the membrane plane) on the orthosteric binding site of MC4R. TM3 and TM6 receptor residues involved in peptide and Ca^{2+} interactions are highlighted. The bold red dashed red arrow indicates an interaction of $L133^{3.36}$ with antagonist residue D-$Nal4^{1}$. (**B**) Top view of (**A**), highlighting all intermolecular hydrophilic inter-actions, assessed by a minimum distance of 3.5 Å (black dashed lines).

in NDP-α-MSH but with $D122^{3.25}$ and $N123^{3.26}$ as in setmelanotide (Fig. 3.3.4B). First, this is because SHU9119 has a more compact orientation to TM3 due to its circularization; secondly, $D122^{3.25}$ is oriented towards calcium than the NDP-α-MSH complex, although SHU9111 has a Nle in the vicinity like NDP-α-MSH. As a result, cofactor Ca^{2+} is fivefold coordinated with identical coordination partners as the NDP-α-MSH-MC4R complex. Nevertheless, none of the LBP contacts in the upper part are unique to SHU9119 and are shared at least with one agonist bound structure. The larger aromatic D-naphthyalalanine (D-Nal) than the D-F in the agonist sequence at the x^{1} position points into the receptor core. D-$Nal4^{1}$ is hydrophobic contact with E100, D126, I129, C130, L133, I185, L197, F261, and L288 (Table 3.5.1). In contrast to both agonists, the antagonist directly interacts with the residues in the vicinity of $W258^{6.48}$ via $L133^{3.36}$ (Figure 3.3.4A).

Table 3.5.1 Comparison of the receptor-ligand contacts in complexes of MC4R with the agonists NDP-α-MSH and setmelanotide and the antagonist SHU9119. Interactions of all relevant amino acids of the MC4R, the calcium ion (Ca^{2+}), and water molecules within the ligand-binding pocket towards the binding peptides NDP-α-MSH, setmelanotide, and SHU9119 were analyzed using CONTACT, a program of the CCP4 software suite. Interactions with a maximum distance of 3.9 Å. Amino acids were abbreviated as one-letter code, capital numbers display the receptor's or peptide's position, and superscripted numbering refers to the conserved peptide *HxRW*. D-Nal corresponds to (2R)-2-amino-3-(naphthalene-2-YL) propanoic acid abbreviated as 4J2 in the protein database (PDB ID: 6w25). D-phenylalanine, D-alanine, and norleucine are abbreviated as three-letter codes Dpn, Dal, and Nle, respectively.

MC4R	**NDP-α-MSH** (agonist)	**Setmelanotide** (agonist)	**SHU9119** (antagonist)
F51	$H6^{0}$	$H4^{0}$	-
N97	-	-	$Dnal4^{1}$ (N97L - antagonized structure)
E100	$Nle4^{-2}$, $E5^{-1}$	A: $Dpn5^{1}$ B: $Dal3^{-1}$	$H3^{0}$, $Dnal4^{1}$, $D2^{-1}$
T101	$H6^{0}$	$H4^{0}$	$H3^{0}$
V103	-	-	$Nle1^{-2}$
I104	$Nle4^{-2}$	$Dal3^{-1}$	$Nle1^{-2}$, $D2^{-1}$, $H3^{0}$
T118	-	-	$Nle1^{-2}$
D122	$Nle4^{-2}$	$R1^{-3}$, $R6^{2}$	$Nle1^{-2}$, $D2^{-1}$, $R5^{2}$
N123	-	$R6^{2}$	$R5^{2}$
D126	$Dpn7^{1}$, $R8^{2}$	$Dpn5^{1}$, $Arg6^{2}$	$Dnal4^{1}$, $Arg5^{2}$
I129	$Dpn7^{1}$	$Dpn5^{1}$	$Dnal4^{1}$
C130	-	-	$Dnal4^{1}$
L133*	-	$Dpn5^{1}$	$Dnal4^{1}$
I185*	$R8^{2}$	-	$Dnal4^{1}$, $R5^{2}$
S188	$R8^{2}$, $W9^{3}$	$R6^{2}$, $W7^{3}$	$R5^{2}$, $W6^{3}$
V193	$W9^{3}$	-	-
I194	$W9^{3}$	$W7^{3}$	$W6^{3}$
L197	$W9^{3}$	$W7^{3}$	$Dnal4^{1}$, $W6^{3}$
F261	-	$Dpn5^{1}$	$Dnal4^{1}$
H264	$W9^{3}$, $G10^{4}$	$W7^{3}$, $C8^{4}$	$W6^{3}$, $K7^{4}$, $NH_2 8^{5}$
L265	-	-	$W6^{3}$
Y268	$W9^{3}$, $K11^{5}$, $P12^{6}$	$W7^{3}$	$W6^{3}$, $K7^{4}$
M281	-	-	$NH_2 8^{5}$
F284	$H6^{0}$, $R8^{2}$, $G10^{4}$	$H4^{0}$, $R6^{2}$, $C8^{4}$	$H3^{0}$, $R5^{2}$, $K7^{4}$
N285	-	$H4^{0}$	-
L288	$H6^{0}$, $Dpn7^{1}$	$H4^{0}$	$H3^{0}$, $Dnal4^{1}$
Ca^{2+}	MC4R: E100, D122, D126; Ligand: $E5^{-1}$, $Dpn7^{1}$	MC4R: E100, D122, D126; Ligand: $Dpn5^{1}$	MC4R: E100, D122, D126; Ligand: $D2^{-1}$, $Dnal4^{1}$
water 30	-	$Dal3^{-1}$, $H4^{0}$, $C8^{4}$, Q43, F284, N285	-
water 43	$Nle4^{-2}$	-	-
water 63	-	$R1^{-3}$, $R6^{2}$	-
water 101	-	-	$D2^{-1}$, $H3^{0}$, $K7^{4}$,
water 102	-	-	$R5^{2}$, $W6^{3}$, Y268, water 103, water 104
water 103	-	-	$W6^{3}$, $K7^{4}$, Y268, water102, water 105
water 104	-	-	$Dnal4^{-1}$, $R5^{2}$, $W6^{3}$, $K7^{4}$, water 102, water 105
water 105	-	-	$D2^{-1}$, $K7^{4}$, water 103, water 104

3.4 Description of the transmembrane core and the G_s-protein interface of active MC4R

The nucleotide-free complex displays a stable low energy state of high similarity in the transmembrane core and G_s-protein interface among different GPCR G_s-protein complex structures. Here, the agonist and G_s-protein binding stabilize the receptor in one defined state. Subtle differences among both agonist-bound complex structures could indicate areas of increased mobility upon receptor activation. The agonist-bound MC4R-G_s structures show subtle changes at TM3, TM7, and IL2 and could indicate areas of relevance for the activation of MC4R and are therefore described in detail.

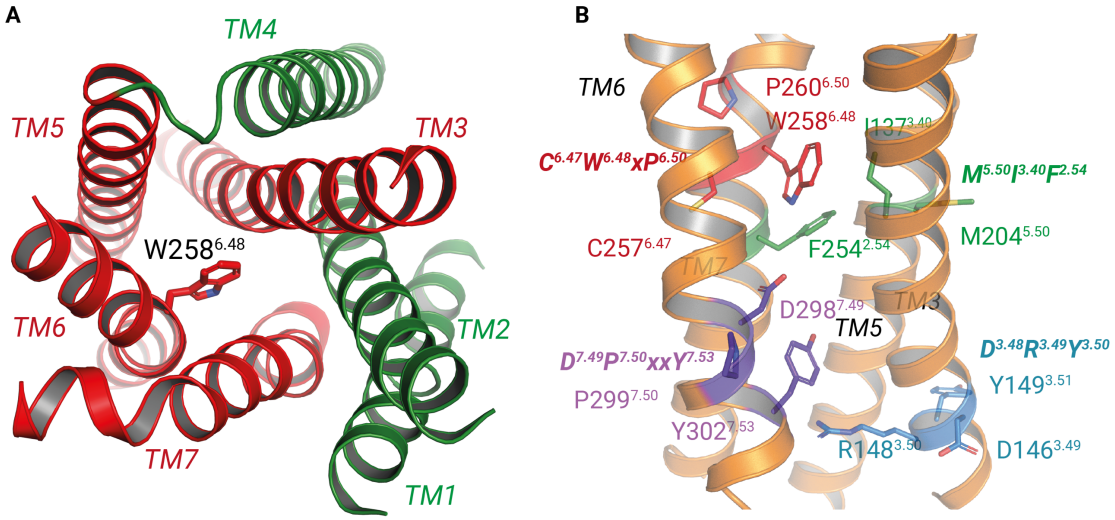

Figure 3.4.1.1 Transmembrane helices packing and conserved GPCR motifs. (**A**) TMs of active MC4R are shown from a view on top of the ligand-binding pocket. TM3, TM5, TM6, and TM7 form a tight bundle in the transmembrane core (shown in red). TM1, TM2, and TM4 (shown in green) are not involved in the tight packing and flank the core bundle. W258$^{6.48}$ is at the center of the TM bundle. (**B**) Conserved residues of *CWxP* (red), *NPxxY* (magenta), *MIF* (green), and *DRY* (blue) motifs are shown as sticks on the structure of setmelanotide-MC4R-G_s. In **B** TM1, TM2 and TM4 are not displayed.

3.4.1 A hydrophobic box surrounds W258$^{6.48}$ in the transmembrane core

The hydrophobic core of the receptor consists of TM3, TM5, TM6, and TM7 (Figure 3.4.1.1). TM1, 2, and TM4 flank the core but are not involved in the tight packing. At the core of this packing is W258$^{6.48}$, which is part of the conserved *CWxP*$^{6.50}$ motif,

surrounded by hydrophobic residues of TM3 (L133$^{3.36}$, I137$^{3.40}$), TM6 (F254$^{6.44}$, F261$^{6.54}$), and TM7 (I291$^{7.41}$) (Figure 3.4.1.1). F254$^{6.44}$ (TM6) together with I137$^{3.40}$ (TM3) and M204$^{5.50}$ (TM5) constitute the *M$^{5.50}$IF* motif in MC4R, reminiscent of the class A GPCR-typical *P$^{5.50}$IF* motif (P$^{5.50}$-I$^{3.40}$-F$^{6.44}$).

As discussed above, the D-F^{1} inside the agonist HFRW motif for NDP-α-MSH (Figure 3.3.2A) and setmelanotide (Figure 3.3.2B) point towards the receptor core but is not in contact (below 4 Å) with residues surrounding W258$^{6.48}$ (Figures 3.4.1A-C). Of note, D-F^{1} in setmelanotide is in contact with F621$^{6.54}$, but NDP-α-MSH is not (Table 3.5.1). The superposition of both agonist-bound structures displays at TM3 a subtle horizontal shift. Here, TM3 in the setmelanotide activated receptor is closer to the inside of the TM bundle (Figure 3.4.1C). This TM3 shift is observable from the extracellular end of TM3 until L133$^{3.36}$ in the transmembrane core. Residues below (closer to the intracellular side) W258$^{6.48}$ at TM3 and TM6 superimpose well (Figure 3.4.1C). In contrast, the TM7 superposition does not show any variations in the extracellular half of TM7 but displays subtle differences in the helix positioning and side chains I291$^{7.41}$ and Y3027.53 (Figure 3.4.1F).

W258$^{6.48}$ is flanked by two residues at TM7, from above (extracellular site) by I291$^{7.41}$ and below (intracellular site) by N294$^{7.45}$. The side chains of I291$^{7.41}$ at TM7 in both structures display different side-chain conformations (Figure 3.4.1C). The side chain of N294$^{7.45}$ is in hydrogen bond distance to the indoleamine of W258$^{6.48}$, and D298$^{7.49}$. Hence, N294$^{7.45}$ connects the *CWxP* motif with the *DP$^{7.50}$xxY* motif. Y302$^{7.53}$, one turn below D298$^{7.49}$ and the Y in *DPxxY* directs towards Y212$^{5.58}$ at TM5. Both tyrosines are 3.8 Å apart and therefore not in hydrogen bond distance, but TM7 displays at this exact position a bulge towards the interface of TM3 and TM5. Y212$^{5.58}$ hydrogen bonds R147$^{3.50}$ of the conserved *DR$^{3.50}$Y* motif are essential for the stable formation of the G-protein binding cavity.

In summary, W258$^{6.48}$, at the center of the TM bundle of TM3, TM5, TM6, and TM7, is surrounded by a box of hydrophobic residues. The active state structures display a connection of the *CWxP* motif with the *DPxxY* motif by hydrogen bonds via N294$^{7.45}$, and a weak interaction connects the *DPxxY* motif via Y212$^{5.58}$ with the *DRY* motif.

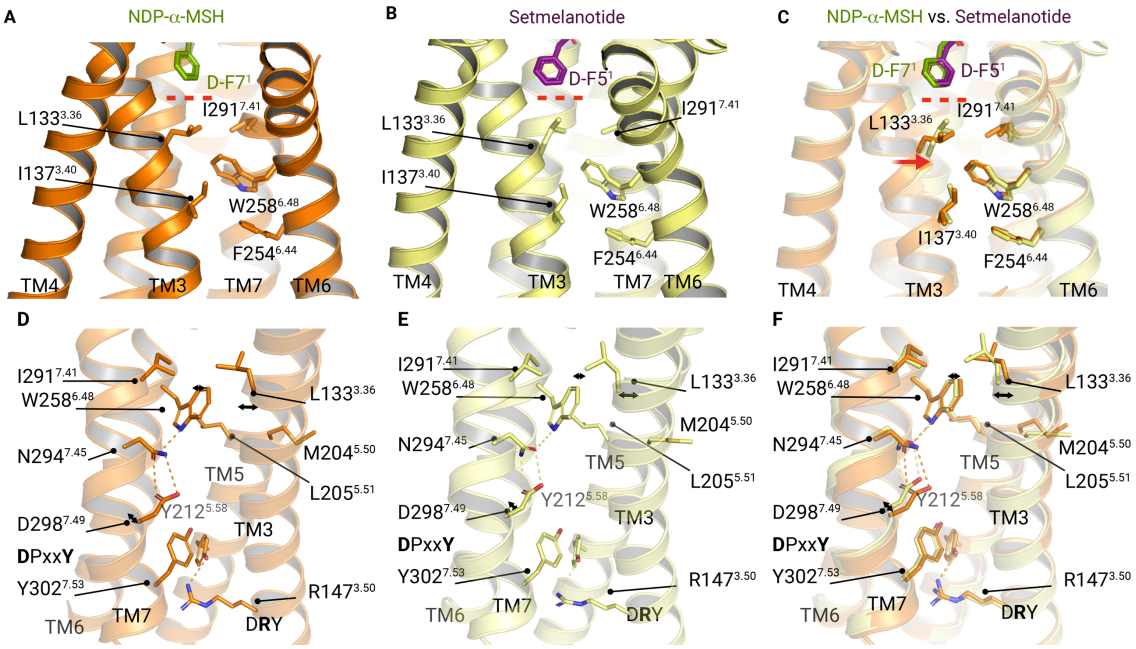

Figure 3.4.1 ***Transmembrane core of NDP-α-MSH (A) and setmelanotide-MC4R-G$_s$ structures (B).*** Hydrophobic residues surround W258$^{6.48}$ at TM3 (L133$^{3.36}$, I137$^{3.40}$), at TM6 (F254$^{6.54}$), and TM7 (I2917$^{7.41}$). The superposition of both agonist-bound structures (**C**) displays subtle changes in the side-chain residues of L133$^{3.36}$, and I291$^{7.41}$. The horizontal shift of TM3 that is apparent in the ligand-binding pocket prolongs down until L133$^{3.36}$, indicated by a red arrow in **C**. The residues below an imaginary horizontal axis at the level of W258$^{6.48}$ do not display differences. D-F^{1} points towards the transmembrane core (NDP-α-MSH green, setmelanotide in purple) but does not interact with residues at the level of L133$^{3.36}$, I291$^{7.41}$, or W258$^{6.48}$), indicated by a red dashed horizontal line. (**D-F**) Display of conserved residues of the *CWxP, NPxxY* , *MIF*, *DRY* motifs and residues that play a potential role in receptor activation. TM3, TM5, TM6, and TM7 are displayed.

3.4.2 MC4R-G_s-protein interface

The Gs-protein binding mode of both agonist-bound structures to MC4R varies only in subtle differences outlined in the next section, and the overall binding mode is in high accordance with previously published receptor-G_s structures such as the β2AR-Gs complex [20] (PDB ID: 3SN6). In brief, the Gα subunit consists of the catalytic active $G\alpha_s$Ras domain and the α-helical domain (AHD). The flexible AHD is not covered by high-frequency data in the cryo-EM datasets and was therefore not modeled. The absence of the AHD is the case for all published GPCR-G-protein complex structures acquired by cryo-EM. The $G\alpha_s$Ras domain is responsible for all interactions with MC4R; the Gβ and Gγ domains do not engage the receptor. Nb35 stabilizes the Gβ and Gαs interface, but it does not interact with the empty nucleotide-binding pocket (indicated as a beige circle in Figure 3.4.3) based on $G\alpha_s$–GTPγS (PDB ID: 1azt). The addition of Nb35 drastically increased the complexed G_s-protein (data not shown). Preincubation of G_s and Nb35 increased the complex yield further (data not shown). The yield optimization of the preincubation underlines that Nb35 does not impair GPCR binding, the GDP release, and the mobility of the AHD.

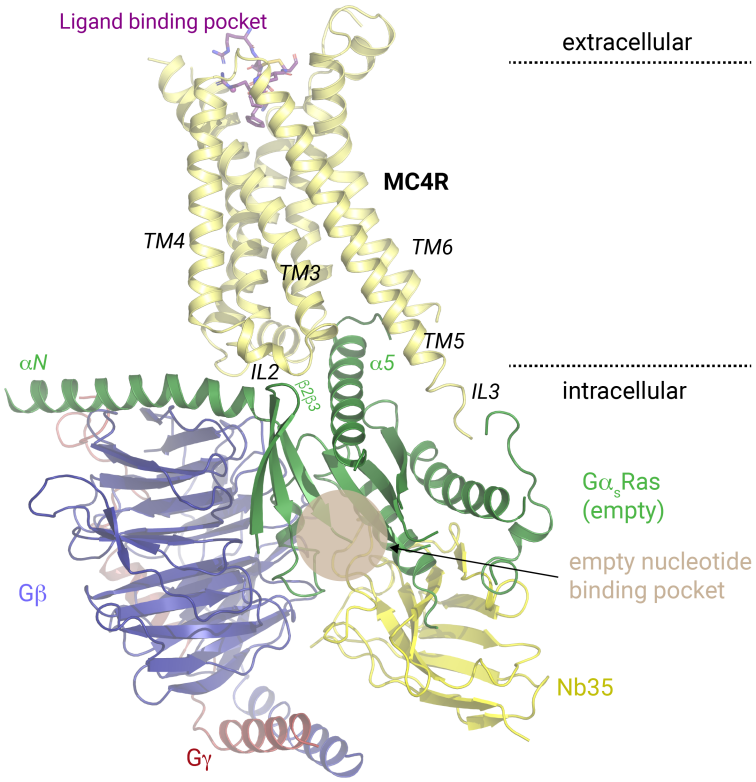

Figure 3.4.3 *G_s binding to MC4R.* Display of the cartoon representation of Nb35 stabilized G_s bound to MC4R. The area of the nucleotide-binding pocket is shown as a beige circle. Dashed lines indicate the intra- and extracellular environment.Setmelanotide-MC4R-Gs model is shown

The conformational link between MC4R and the nucleotide-binding pocket involves the amino-terminal αN-helix, the carboxy-terminal α5-helix of Gαs, and the β2β3 junction (Figure 3.4.3). Other regions are not in contact below 4 Å to the receptor. The carboxy-terminal end of the α5-helix is deeply buried in the cytoplasmic crevice of MC4R and is responsible for most interactions. The α5-helix engages predominantly with IL2, TM3, and TM5 of MC4R. IL3 is not fully resolved in both structures.

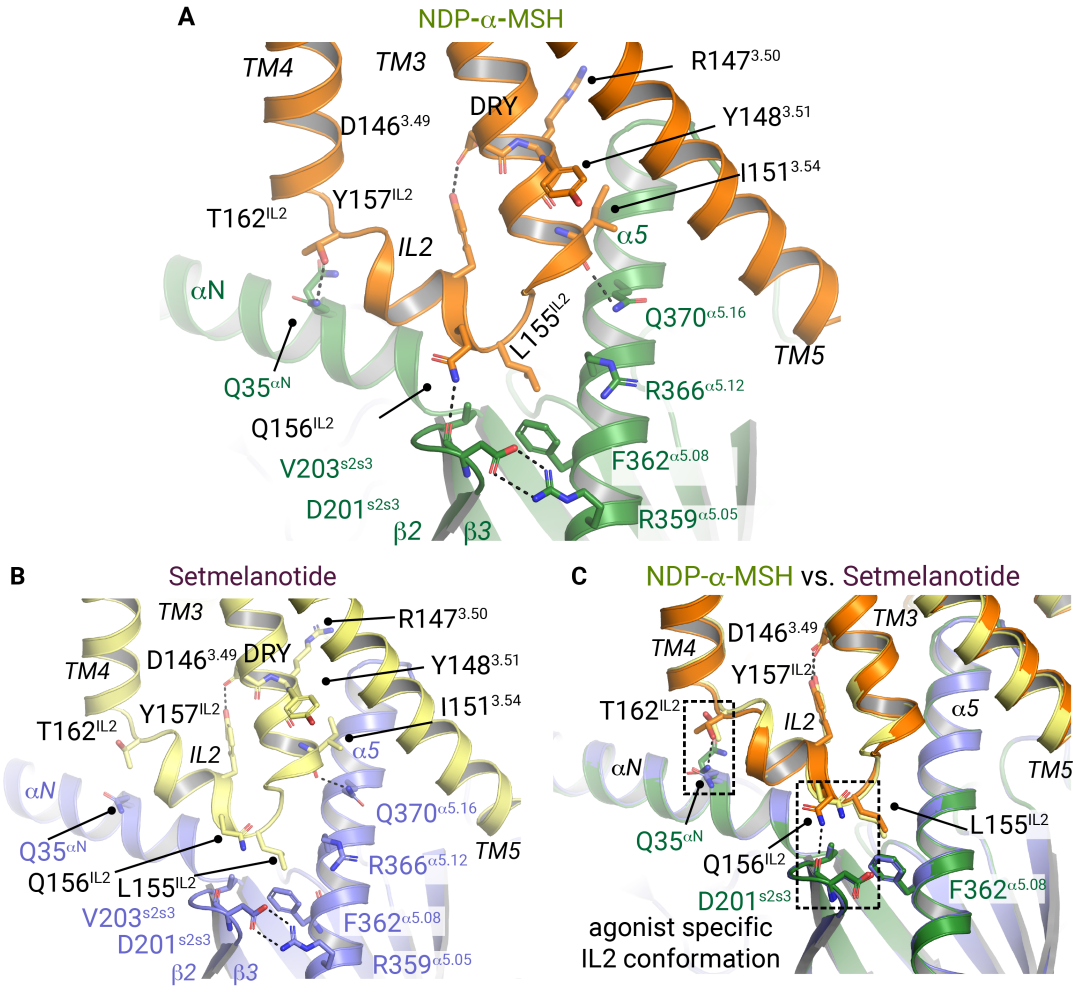

Figure 3.4.3 ***MC4R interacts predominantly with the α5-helix and αN-helix of Gαs.*** The IL2-TM3-TM6 site stabilizes the active receptor-G_s complex structures of MC4R bound to (**A**) NDP-α-MSH, (**B**) setmelanotide, and (**C**) the superposition of **A-B**. Black dashed lines indicate hydrophilic interactions.

3.4.3 IL2 variability at the G_s interface

The MC4R-G_s-protein binding interface is formed by IL2, TM3, TM5 of MC4R and the α5-helix and αN-helix of $G\alpha_s$.

As described above, the cytoplasmic half of TM3 is adjusted on the horizontal plane relatively for both agonist-bound structures (Section 3.3.3). The same helix forms the primary $G\alpha_s$ interaction site with the *DRY* motif and its junction IL2. The IL2 is stabilized as a $_{3/10}$-helix by a hydrogen bond of $Y157^{IL2}$ with $D146^{3.49}$ of the *DRY* motif. The α5-helix is stabilized turn by turn via hydrogen bonds with the TM5 (not shown), IL2, and TM3 (Figure 3.4.3). $L155^{IL2}$ is at the center of the IL2 interaction network with the α5-helix and the β2β3 loop. $L155^{IL2}$ displays a head-to-face π–π interaction with $F362^{\alpha5.08}$ (G-protein numbering according to [49]). $V203^{s2s3}$ further stabilizes $F362^{\alpha5.08}$. This site has been described as the central anchor for G-protein insertion at β2AR [212].

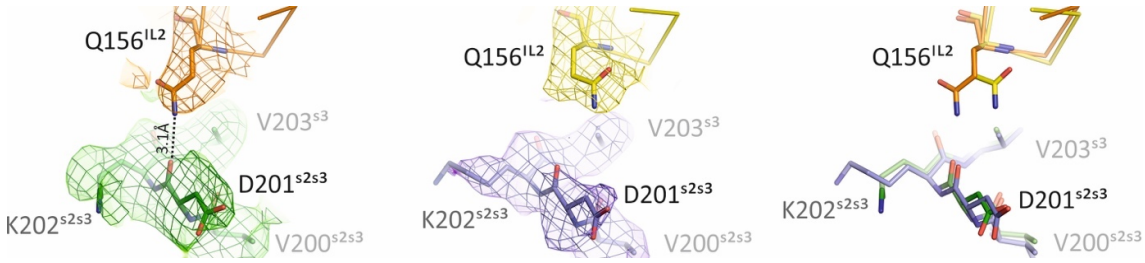

Figure 3.4.3.1 ***Cryo-EM map differences at $Q156^{IL2}$ in the IL2-G_s interface***. Close-up view on the cryo-EM maps for $Q156^{IL2}$ in the NDP-α-MSH-MC4R-G_s (left) and setmelanotide-MC4R-G_s complex (middle) and the superposition of both complexes (right). NDP-α-MSH-MC4R, the corresponding $G\alpha_s$-protein, setmelanotide-MC4R, and its $G\alpha_s$-protein are colored orange, dark green, yellow, and dark purple. Amino acids are shown as sticks. The protein is depicted as a ribbon. Cryo-EM maps are displayed as mesh and volume, contoured at 4 σ level and colored corresponding to the displayed proteins.

NDP-α-MSH stabilizes an IL2 conformation that varies subtly to the one stabilized by setmelanotide. Consequently, the hydrogen bond network differs. Two potential additional hydrogen bonds can be formed in the NDP-α-MSH complex. First, one of the two rotamers of Q35 at the αN-helix and $T162^{IL2}$ are in hydrogen-bond distance. Secondly, the backbone of V203 at the β2β3 loop is in hydrogen bond distance with $Q156^{IL2}$ (Figure 3.4.3.1). In contrast, the cryo-EM densities of the setmelanotide complex do not indicate hydrogen

bonds here. The Q156^{IL} sidechain conformation is rotated away from the β2β3 loop (Figure 3.4.3.1), and T162^{Il2} is rotated away from Q35 (Figure 3.4.3.2). The cryo-EM density suggests that the IL2 adjustment results in varying hydrogen bond interaction patterns among both agonist-bound structures.

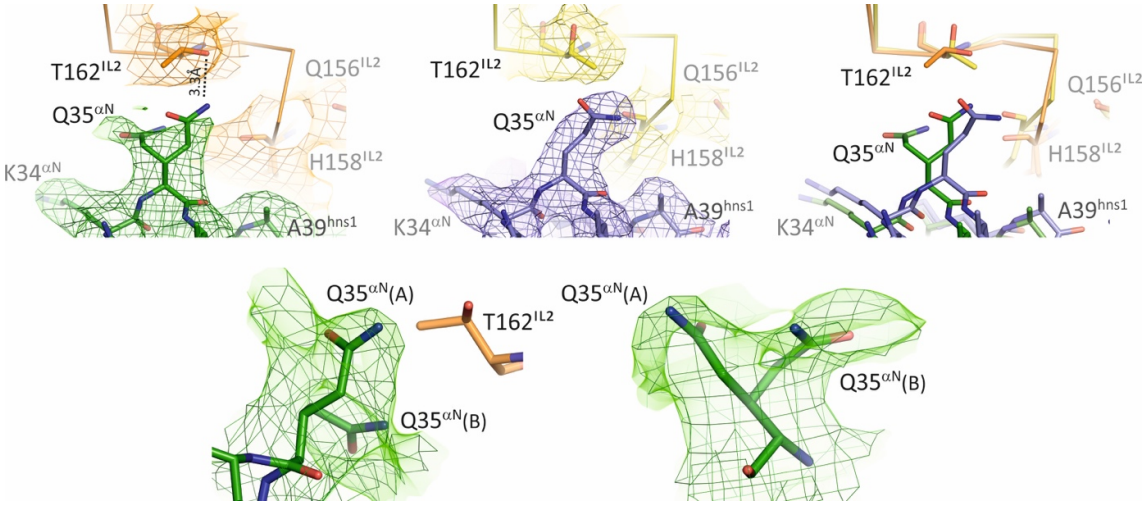

Figure 3.4.3.2 ***Cryo-EM map differences at T162^{IL2} in the IL2-G$_s$ interface.*** Close-up view on the cryo-EM maps for the amino acid T162^{IL2} in the NDP-α-MSH-MC4R-G$_s$ (left) and setmelanotide-MC4R-G$_s$ complex (middle) and the superposition of both complexes (right). NDP-α-MSH-MC4R, the corresponding Gα$_s$-protein, setmelanotide-MC4R, and its Gα$_s$-protein are colored orange, dark green, yellow, and dark purple. Amino acids are depicted in stick representation. The protein is visualized as a ribbon. Cryo-EM maps are displayed as mesh and volume. In the upper row, the cryo-EM maps are contoured at 4 σ level. To verify the double conformation of Q35^{aN}, the contour level in the bottom row is 2 σ level. All cryo-EM maps are colored corresponding to the displayed proteins.

3.4.4 MC4R specific interactions at the α5-helix of Gα$_s$

The carboxy-terminus of the α5-helix of Gα$_s$ forms extensive interactions at the MC4R G-protein binding cavity (Figure 3.4.4.1). For example, R147$^{3.50}$ of the DRY motif at TM3 interacts with Y212$^{5.58}$ at TM5. Intriguingly, Y377$^{α5.23}$ forms a hydrogen bond interaction with T150$^{3.53}$ at TM3. This novel interaction for a receptor-G$_s$ complex is apart by 3.0 Å for the setmelanotide-MC4R-G$_s$ complex and by 2.7 Å for the NDP-α-MSH-MC4R-G$_s$ complex.

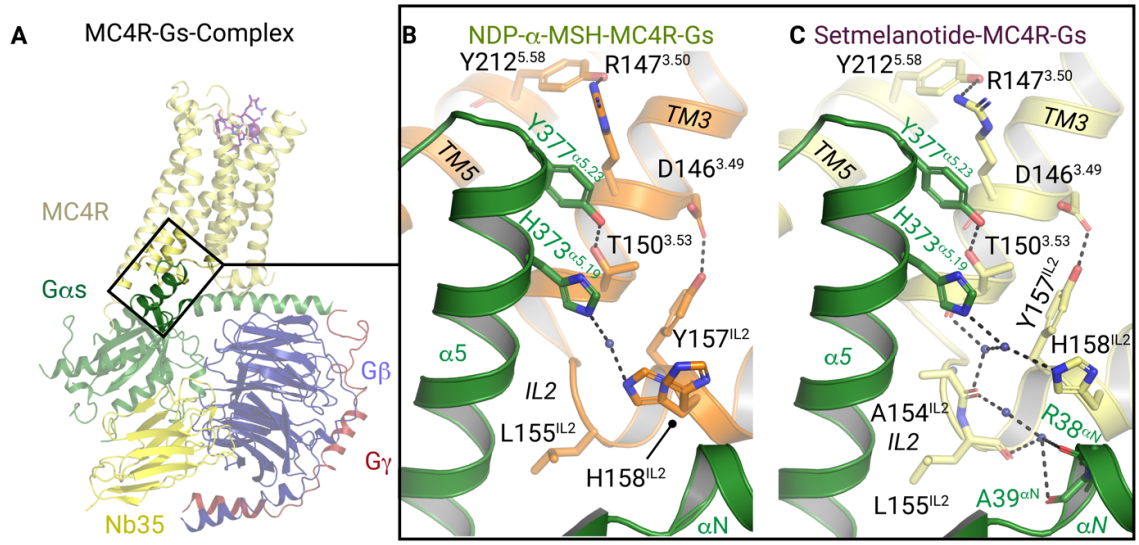

Figure 3.4.4.1 ***MC4R and G_s interaction at the intracellular receptor crevice.*** (**A**) Full-view on MC4R-G_s complex structure. (**B-C**)**.** Interface of the $G\alpha_s$ αN-, α5-helices with MC4R TM3 and-IL ((**B**) NDP-α-MSH-MC4R-Gs and (**D**) setmelanotide MC4R-Gs). Black dashed lines indicate hydrophilic interactions.

The different IL2 conformations show their most significant cryo-EM map deviations at $H158^{IL2}$ (Figure 3.4.4.2). The NDP-α-MSH stabilized structure displays two rotamers, of which one forms a stable hydrogen bridge with $H373^{\alpha5.19}$. The setmelanotide stabilized complex has one rotamer at that site but displays an extensive hydrogen bond network. This network connects IL2 with the α5-helix via $H373^{\alpha5.19}$and leads $A154^{IL2}$ to $R38^{\alpha N}$ at the αN-helix (Figure 3.4.4.1C).

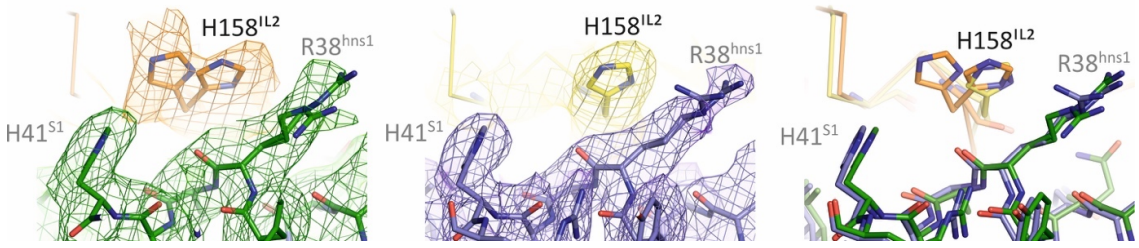

Figure 3.4.4.2 ***Cryo-EM map differences at $H158^{IL2}$ in the IL2-G_s interface.*** The close-up view reveals for $H158^{IL2}$ a double conformation in the NDP-α-MSH-MC4R-G_s complex (left) but not for the setmelanotide-MC4R-G_s complex (middle). NDP-α-MSH-MC4R, the corresponding $G\alpha_s$-protein, setmelanotide-MC4R and its $G\alpha_s$-protein are colored orange, dark green, yellow, and dark purple. Amino acids are visualized as sticks and the protein as a ribbon. Cryo-EM maps are displayed as mesh and volume, contoured at 4 σ level and colored corresponding to the displayed proteins.

3.5 MC4R related signaling data

Based on both agonist-MC4R-G_s structures, point mutations were suggested. Heike Biebermann's lab investigated the activation of the G_s pathway via cAMP accumulation assays, the PLC activation via the NFAT reporter gene assay, and the expression of the respective mutants in *HEK293* cells. Single-side chain substitutions can impact cell surface expression levels of the mutant. Receptor inactivation by single-side chain substitutions can be differentiated from reduced receptor expression by comparing signaling efficacy and expression levels (Table 3.5.1). The signaling data is from Heyder *et al.* 2021[118].

The signaling information guides the discussion, especially for interpreting the MC4R activation mechanism. GPCR structures of the inactive and active state display energy minima and do not necessarily reflect the functional relevant state for the activation process and G-protein binding. Here, signaling data can illuminate the relevance of structural observations.

The initial set-up of the signaling experiments did not include setmelanotide, and comprehensive data was generated for NDP-α-MSH and the endogenous ligand α-MSH. During the revision of Heyder *et al.* 2021 [118] a small subset of MC4R mutants was carefully selected, and G_s signaling upon setmelanotide stimulation was investigated.

cAMP signaling data

The primarily investigated signaling pathway of MC4R is the G_s signaling pathway. Here, the receptor induces G_s-protein dissociation into the $G\alpha_s$ and $G\beta\gamma$ subunits. The $G\alpha_s$ subunit activates adenylate cyclase, which converts ATP into cAMP. The second messenger cAMP is measured via the AlphaScreen cAMP detection kit (Perkin Elmer). The determination of endpoint measurements is facilitated by supplementing the assay buffer with the phosphodiesterase inhibitor 3-isobutyl-1-methylxanthine (IBMX). Variant expression levels and the data of cAMP maximum stimulation experiments are given in Table 3.5.1. Concentration-response experiments on G_s signaling are summarized in Table 3.5.2.

The G_s signaling data is separated into mutations lining the ligand-binding pocket (LBP), the receptor transmembrane core (TM core), and the G-protein binding cavity (GBC).

Table 3.5.1 ***Determination of total and cell surface expression with Nano-Glo®HiBiT Lytic/Extracellular detection system, and maximal cAMP response E_{max}.*** Data are given as the result of four to eight independent experiments performed in triplicates ± SEM. Wild-type MC4R (WT) stimulation as fold-over MC4R basal is 20 ± 1.06-fold for α-MSH, 22 ± 1.15-fold for NDP-α-MSH, and 14.53 ± 1.53-fold for setmelanotide, all set as 100%. Expression data were cleaned by performing a ROUT test with Q = 1%. Statistics were done by one-way ANOVA with Kruskal-Wallis test. WT was tested against all mutants stimulated with the indicated ligand: a: $p < 0.05$; b: $p < 0.01$; c: $p < 0.001$; d: $p < 0.0001$.

Substi-tution	Total expression [fold-WT basal]	Cell surface expression [fold-WT basal]	Basal [fold over WT basal]	α-MSH E_{max} [% of WT at 1 µM]	NDP-α-MSH E_{max} [% of WT at 1 µM]	Setmelanotide E_{max} [% of WT at 1 µM]
WT	1	1	1	100	100	100
E100N	1.30 ± 0.10	2.42 ± 0.15^{d}	0.98 ± 0.08	11.35 ± 1.45^{a}	56.10 ± 4.13^{a}	103 ± 13
T101A	1.41 ± 0.07	0.89 ± 0.11	1.16 ± 0.16^{a}	189 ± 10^{d}	155 ± 12^{b}	-
D122S	1.63 ± 0.06^{c}	1.41 ± 0.09	1.16 ± 0.06	40.10 ± 3.46	129 ± 8.56	138 ± 11
N123A	0.53 ± 0.03^{d}	0.60 ± 0.05^{c}	1.05 ± 0.05	163 ± 13^{b}	119 ± 7.89	154 ± 11^{b}
D126S	1.28 ± 0.07	1.87 ± 0.14^{c}	0.90 ± 0.05^{a}	13.92 ± 1.67^{a}	91.23 ± 7.74	81.69 ± 7.77
C130A	0.99 ± 0.04	1.13 ± 0.09	1.03 ± 0.09	147 ± 9.24^{a}	114 ± 9.25	-
L133A	1.11 ± 0.10	0.78 ± 0.15	1.03 ± 0.05	130 ± 6	152 ± 8^{a}	-
L133F	0.97 ± 0.09	1.34 ± 0.11	0.85 ± 0.04^{a}	112 ± 6	137 ± 7	-
I137A	0.83 ± 0.04	0.79 ± 0.03	1.30 ± 0.08	92.78 ± 4.96	99.40 ± 4.89	-
I137F	0.81 ± 0.07	0.57 ± 0.06^{b}	1.02 ± 0.06	76.65 ± 3.93	94.21 ± 5.53	-
T150A	0.60 ± 0.06^{c}	0.58 ± 0.04^{b}	0.76 ± 0.06^{d}	26.11. ± 1.64^{d}	18.07 ± 2.17^{d}	32.41 ± 1.87^{d}
T150I	0.69 ± 0.06^{a}	0.55 ± 0.02^{b}	0.88 ± 0.06^{c}	36.46 ± 2.24^{d}	40.74 ± 2.08^{d}	-
T150D	0.41 ± 0.03^{d}	0.47 ± 0.03^{c}	0.85 ± 0.05	9.88 ± 0.91^{d}	10.85 ± 0.86^{c}	-
T150F	1.09 ± 0.05	0.88 ± 0.03	0.83 ± 0.02^{a}	42.60 ± 2.41^{d}	55.68 ± 4.64^{a}	-
T150S	1.44 ± 0.09	1.23 ± 0.05	0.77 ± 0.03^{c}	77.81 ± 5.11^{a}	86.18 ± 5.14	-
H158A	1.35 ± 0.17	1.37 ± 0.16	2.49 ± 0.19^{d}	129 ± 15	129 ± 15	148 ± 17^{b}
F184V	1.14 ± 0.07	1.82 ± 0.14^{c}	1.24 ± 0.11	149 ± 14	91.65 ± 8.29	-
S188A	1.10 ± 0.13	1.15 ± 0.06	0.94 ± 0.07	148 ± 11^{a}	142 ± 15	-
D189S	0.98 ± 0.06	1.51 ± 0.13	0.80 ± 0.06	151 ± 16	166 ± 28	-
S191A	1.00 ± 0.06	0.89 ± 0.04	1.00 ± 0.04	199 ± 18^{c}	197 ± 22^{c}	-
L197A	0.78 ± 0.06	1.12 ± 0.16	0.84 ± 0.05	152 ± 11	171 ± 24^{a}	-
M204A	0.60 ± 0.04^{c}	0.69 ± 0.04	1.81 ± 0.13^{d}	113 ± 5	103 ± 5	-
L205A	0.92 ± 0.06	0.82 ± 0.08	1.21 ± 0.06	101 ± 7	99.85 ± 5.61	-
L205F	1.14 ± 0.05	1.02 ± 0.04	2.32 ± 0.20^{d}	112 ± 6	112 ± 6	-
F254A	1.73 ± 0.15^{a}	0.94 ± 0.04	1.02 ± 0.03	88.67 ± 5.41	85.09 ± 5.23	-
F254M	1.99 ± 0.17^{b}	1.57 ± 0.10^{a}	0.99 ± 0.04	77.33 ± 8.14	97.12 ± 10.32	-
W258A	0.73 ± 0.07	0.48 ± 0.04^{c}	0.94 ± 0.07	61.40 ± 4.13^{d}	65.04 ± 5.30^{b}	-
W258F	1.30 ± 0.09	0.96 ± 0.06	2.48 ± 0.22^{d}	78.13 ± 4.39^{a}	72.66 ± 4.83	-
F261V	0.81 ± 0.04	0.77 ± 0.10	0.93 ± 0.08	63.54 ± 4.19	61.61 ± 5.96	112 ± 8
H264A	0.47 ± 0.04^{d}	1.35 ± 0.05	1.10 ± 0.07	10.21 ± 0.94^{a}	133 ± 9.10	111 ± 10
L265A	0.62 ± 0.04^{d}	0.78 ± 0.05	0.89 ± 0.08	83.68 ± 6.16	140 ± 12	-
Y268F	0.98 ± 0.07	1.06 ± 0.06	0.87 ± 0.03	140 ± 16	133 ± 14	-
H283A	0.95 ± 0.07	1.17 ± 0.08	1.40 ± 0.18	164 ± 15^{a}	152 ± 21	-
F284V	0.67 ± 0.09^{b}	0.90 ± 0.05	1.08 ± 0.03	145 ± 14	165 ± 27	-
N285S	0.86 ± 0.07	0.99 ± 0.07	1.14 ± 0.05	173 ± 16^{a}	167 ± 22	-
Y287V	0.74 ± 0.02	1.75 ± 0.10^{b}	0.99 ± 0.04	148 ± 12	142 ± 18	-
L288A	0.51 ± 0.04^{d}	0.47 ± 0.05^{d}	0.94 ± 0.04	76.41 ± 5.51	129 ± 12	-
I291A	0.49 ± 0.03^{d}	0.54 ± 0.05^{b}	0.94 ± 0.03	11.09 ± 0.94^{d}	8.98 ± 0.69^{d}	-
I291F	0.81 ± 0.08	0.54 ± 0.04^{c}	0.92 ± 0.05	79.83 ± 6.47	71.34 ± 6.02	-

Changes in potency (EC_{50}) indicate that residues are involved in ligand binding, and differences in efficacy (E_{max}) indicate that residues are involved in receptor activation and signal transmission [213].

The concentration-response measurements and the maximum stimulation experiments have been performed independently. Therefore, the absolute numbers upon stimulation with 1 µM agonist can vary for both set-ups. Thus, data from concentration-response measurements are prioritized in the following, and the maximal cAMP response will only be discussed in the absence of dose-response measures.

Table 3.5.2 ***Determination of half-maximal effective concentration EC_{50} via cAMP accumulation using the AlphaScreen® assay after adding agonists α-MSH, NDP-α-MSH, and setmelanotide.*** Data are given as the result of four to eight independent experiments performed in triplicates ± SEM. Statistics were done by one-way ANOVA with Kruskal-Wallis test. Wild-type MC4R (WT) was tested against all mutants stimulated with the indicated ligand. a: $p < 0.05$; b: $p < 0.01$; c: $p < 0.001$; d: $p < 0.0001$; n.d. = not determined due to either too low E_{max} or due to severely shifted concentration-response curve which do not allow proper EC_{50} calculation.

Substitution	α-MSH EC_{50} [nM]	NDP-α-MSH EC_{50} [nM]	Setmelanotide EC_{50} [nM]
MC4R WT	17.6 ± 4.21	1.17 ± 0.26	1.24 ± 0.39
E100N	n.d.	383 ± 139[c]	n.d
T101A	49.4 ± 6.12	0.9 ± 0.05	-
D122S	n.d.	9.74 ± 1.61[b]	n.d.
N123A	33.3 ± 5.97	0.93 ± 0.23	6.40 ± 2.82
D126S	n.d.	166 ± 34.2[c]	n.d
C130A	136 ± 21.9[c]	2.20 ± 0.71	-
L133A	94.7 ± 1.58	2.07 ± 0.26	-
L133F	76.1 ± 17.0	3.79 ± 0.77	-
I137A	19.7 ± 5.97	1.15 ± 0.08	-
T150A	n.d.	n.d.	1.09 ± 0.08
T150I	86.2 ± 34.9	1.41 ± 0.35	-
T150F	91.9 ± 16.9	2.06 ± 0.58	-
T150S	36.2 ± 9.10	0.76 ± 0.16	-
H158A	7.59 ± 3.83	0.58 ± 0.14	2.4 ± 0.82
S188A	101 ± 30.2	0.97 ± 0.19	-
M204A	6.34 ± 1.17	1.06 ± 0.44	-
L205F	7.18 ± 1.84	1.38 ± 0.45	-
W258A	70.7 ± 17.7	2.53 ± 0.34	-
W258F	2.60 ± 0.53	0.59 ± 0.2	-
F261V	118 ± 14.5[c]	2.96 ± 1.44	28.80 ± 7.70
H264A	n.d.	179 ± 32.5[d]	n.d.

PLC signaling assay data

The MC4R-G_s-protein interfaces deviate upon setmelanotide compared to NDP-α-MSH binding. These subtle differences indicate a ligand-related allosteric connection of the ligand-binding pocket with the G_s-protein interface. To assess the potential biological relevance of these subtle changes, PLC signaling was investigated of receptor variants of critical residues involved in ligand binding and facilitating the G_s-protein interface (Table 3.5.3. and Figure 3.5.1).

Recent studies have associated $G_{q/11}$ signaling of MC4R with satiety induction in humans [102] and mice [114]. MC4R induces the activation of PLC via the $G_{q/11}$ family, measured by the NFAT responsive element located in the promoter region of the gene fire-

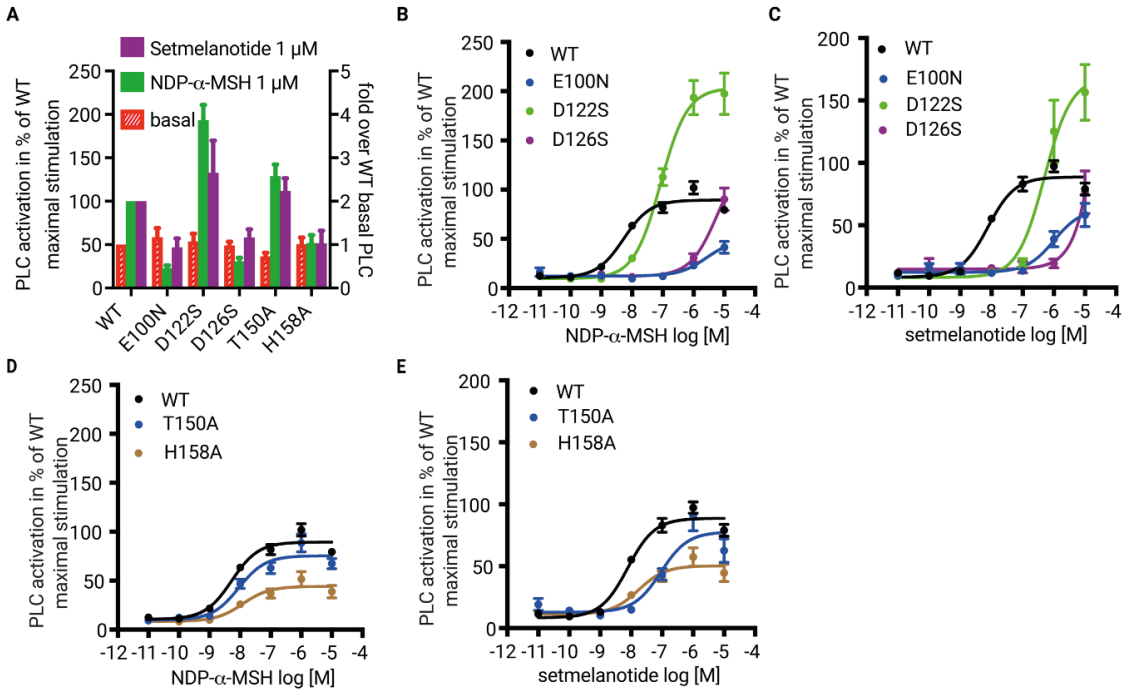

Figure 3.5.1 ***PLC signaling of MC4R variants with substitution in the ligand- and G-protein binding regions***. (**A**) Maximal $G_{q/11}$-protein signaling was determined as PLC activation of MC4R WT and mutants after adding 1 µM NDP-α-MSH and setmelanotide indicated as fold of WT max. signaling [%]. Basal PLC accumulation of WT and mutants are depicted as fold-over WT basal. (**B-C**) Concentration-response curves of $G_{q/11}$-protein signaling determined as PLC activation of WT and indicated mutants in the LBP after addition of (**B**) NDP-α-MSH and (**C**)setmelanotide shown as fold of MC4R WT max. signaling [%]. (**D-E**) Concentration-response curves of $G_{q/11}$-protein signaling determined as PLC activation of WT and indicated mutants in the LBP shown as fold of WT max. signaling [%].

fly luciferase. The luciferase gene expression is regulated by the second messenger, inositol 1,4,5-trisphosphate (IP_3). The resulting chemiluminescence is measured as a direct function of IP_3 concentration. $G_{q/11}$ signaling is discussed along with the G_s signaling data.

Table 3.5.3 ***Determination of $G_{q/11}$ activation by determination of phospholipase C activation via NFAT reporter gene assay***. Data are given as the result of four to six independent experiments performed in triplicated ± SEM. The stimulation of wild-type MC4R (WT) as fold-over WT basal is 7.56 ± 0.87 for NDP-α-MSH challenge and 10.91 ± 0.88 for stimulation with setmelanotide at a concentration of 1 µM. Statistics were done by one-way ANOVA with Kruskal-Wallis test. MC4R-WT was tested against all mutants stimulated with the indicated ligand: a: $p < 0.05$; b; $p < 0.01$; c: $p < 0.001$; d: $p < 0.0001$; n.d. = not determined due to either too low E_{max} or due to severely shifted concentration-response curve which do not allow proper EC_{50} calculation.

	Basal	**NDP-α-MSH**		**Setmelanotide (RM-493)**	
Variant	E_{max} [fold-WT basal]	E_{max} at 1µM [fold-WT at 1µM]	EC50 [nM]	E_{max} at 1µM [fold-WT at 1µM]	EC50 [nM]
WT	1	100	5.81 ± 0.88	100	6.29 ± 0.76
E100N	1.17 ± 0.21	22.94 ± 3.20[c]	n.d.	46.72 ± 10.23[a]	n.d.
D122S	1.07 ± 0.18	194 ± 17	85.5 ± 12.4	132 ± 37	401 ± 59.6[a]
D126S	0.97 ± 0.09	30.53 ± 4.33[b]	n.d.	25.72 ± 6.94[d]	n.d.
T150A	0.73 ± 0.08[b]	129 ± 14	11.7 ± 1.98	126.75 ± 13.19	92.1 ± 24.7
H158A	1.01 ± 0.15	51.66 ± 9.35	10.00 ± 4.47	57.45 ± 7.56[a]	20.6 ± 6.58

Agonist signaling potency and efficacy for G_s and PLC induced signaling

NDP-α-MSH and setmelanotide's G-protein signaling via G_s and $G_{q/11}$ have not been compared in one experimental set-up. Our data suggest that both agonists have a similar potency for the induction of G_s and $G_{q/11}$ family signaling but deviate in their efficacy. NDP-α-MSH is more efficacious in G_s and setmelanotide in $G_{q/11}$ activation (Figure 3.5.2).

3.5.1 MC4R variations lining the ligand-binding pocket

The largest dataset of signaling mutants was measured from MC4R variants with mutations in the ligand-binding pocket. Here, MC4R variants of residues E100N, T101A, D122S, D126S, S188A, H264A were selected since these are in hydrophilic contact with NDP-α-MSH. In addition, acidic receptor residues interacting with the cofactor Ca^{2+} were mutated.

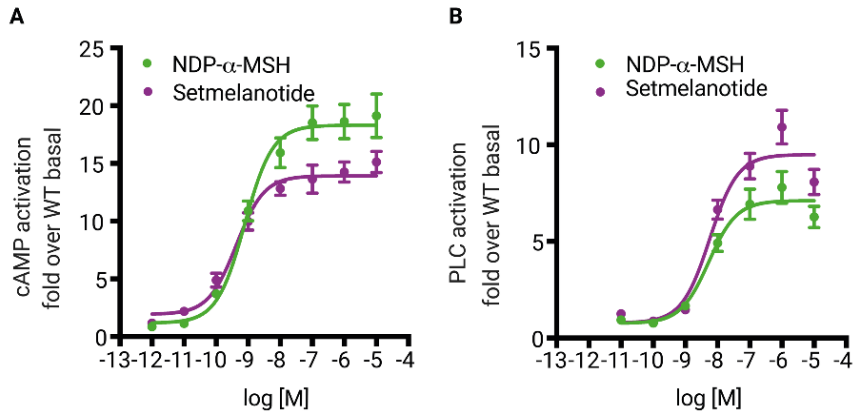

Figure 3.5.2 ***Comparison of G_s and $G_{q/11}$ induced signaling at wild-type MC4R upon stimulation with NDP-α-MSH and setmelanotide.*** (**A**) Concentration-response curves of G_s-protein signaling determined as cAMP accumulation of MC4R and indicated mutants after stimulation by NDP-α-MSH and setmelanotide shown as fold change of basal stimulation. (**B**) Concentration-response curves of $G_{q/11}$-protein signaling determined as PLC activation of MC4R and indicated mutants after the challenge of α-MSH shown as fold change of basal stimulation.

to serine or asparagine to maintain the hydrophilic side chain properties (E100N, D122S, D126S). Furthermore, we investigated residues lining the ligand-binding pocket below 4 Å to NDP-α-MSH or SHU9119: N123A, C130A, F184A, S188A, D189A, S191A, L197, L265, F261V, Y268A, F284V, N285A, Y287A, Y287V, and L288A

The MC4R mutants of variants lining the LBP did not significantly reduce cell surface expression except N123A and L288A (Figure 3.5.1.1A). However, a significant increase in cell surface expression is observed for MC4R mutants E100N, D122S, D126S, F184V, and Y287V (Figure 3.5.1.1A). Of note, all three variants of the acidic residues coordinating Ca^{2+} (E100N, D122S D126S), the *EDD* motif, display increased expression levels of up to 2.4 times (E100N) of WT MC4R (Figure 3.5.1.1A).

The increased cell surface expression of these three mutants is contradicted by a total loss of α-MSH-induced stimulation of G_s signaling (Figure 3.5.1.1B). The mutants E100N and D126S have impaired G_s signaling for all tested agonists (Figures 3.5.1D, F, and G). A ligand-dependent response was observed for MC4R-D122S. Here, NDP-α-MSH-induced signaling depends less on $D122^{3.25}$ since the D122S mutant significantly drops in potency than WT (Figure 3.5.1.1D), in contrast to the stimulation with setmelanotide (Figure 3.5.1.1G).

The variant H264A in TM6 leads to a total loss of cAMP accumulation upon α-MSH stimulation (Figure 3.5.1.1B). On the other hand, the high-affinity stimulation by setmelanotide and NDP-α-MSH do not result in a lack of cAMP accumulation but in a significant shift in potency (EC_{50}) (Figures 3.5.1D and G).

MC4R mutants T101A, N123A, C130A, S188A, D189S, S191A, L197A, N285A, and Y287S/V display significant hyperstimulation upon stimulation with NDP-α-MSH and α-MSH (Figure 3.5.1.1B). Hence, the signaling is either hampered or enhanced by manipulating the LBP but does not remain like WT.

F261V, a TM6 residue close to the transmembrane core, has a deviating response in G_s signaling for setmelanotide in contrast to α-MSH and NDP-α-MSH. A potency reduction in the cAMP accumulation was observed for setmelanotide with EC_{50} = 29 nM compared to wild-type EC_{50} = 1.2 nM in contrast to a nearly unchanged potency for NDP-α-MSH and α-MSH (Tables 3.5.2). However, the efficacy of setmelanotide was comparable to wild-type MC4R but significantly reduced for NDP-α-MSH and α-MSH (Tables 3.5.1-3.5.2).

Furthermore, the N123A mutant displays a subtle difference upon setmelanotide stimulation compared to NDP-α-MSH and α-MSH stimulation. Setmelanotide induces at mutant N123A a six-fold decreased potency compared to WT (Table 3.5.2).

In summary, three residues, $E100^{2.60}$, $D126^{3.29}$, and $H264^{6.54}$, are essential for G_s signaling upon stimulation with setmelanotide, NDP-α-MSH, and α-MSH. α-MSH-induced signaling results in a total loss of efficacy by MC4R mutants E100N, D122S, D126S, and H264A. A total loss of signaling is not the case for the high-affinity agonists and is probably driven by α-MSH's lower affinity. Intriguingly, NDP-α-MSH G_s signaling is less dependent on $D122^{3.25}$. Ultimately, MC4R induced G_s signaling stimulated by setmelanotide compared to NDP-α-MSH depends on residues $D122^{3.25}$ and $F261^{6.51}$ (indicated by a significant change in potency).

The $G_{q/11}$ signaling data of MC4R mutants E100N, D122S, D126S, and H264A upon stimulation with setmelanotide and NDP-α-MSH displays the same tendencies as for G_s signaling (Figure 3.5.1 and Tables 3.5.2-3.5.3).

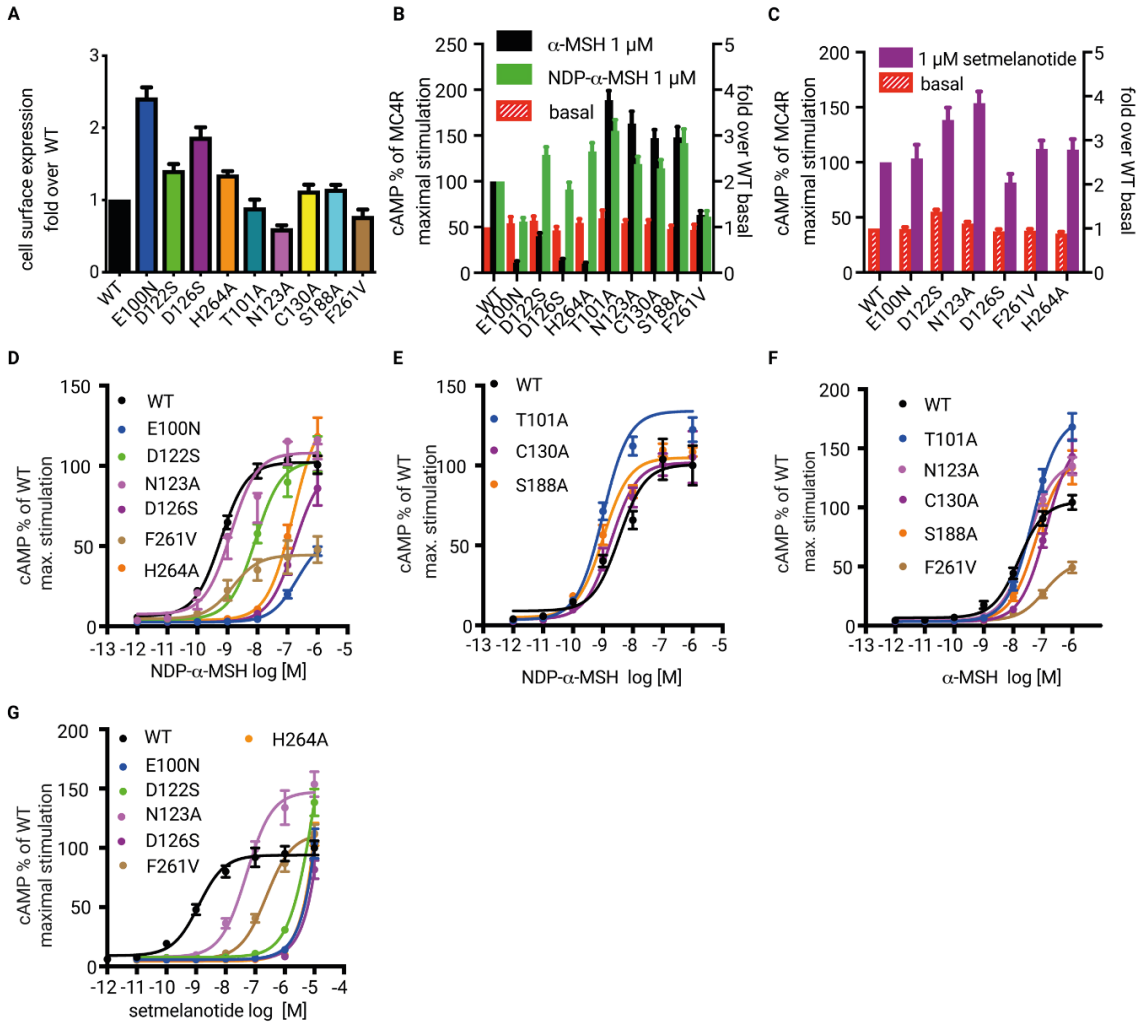

Figure 3.5.1.1 ***G_s signaling data of MC4R mutants with variants in the ligand-binding region of MC4R.*** (**A**) Cell surface expression of wild-type (WT) MC4R and mutants are shown as fold change over WT. (**B-C**) Maximal G_s-protein signaling determined as the cAMP accumulation of WT and mutants after addition of 1 μM (**C**) α-MSH, NDP-α-MSH, or (**D**) setmelanotide indicated as fold change of WT maximal signaling [%]. (**E-G**) Concentration-response curves of G_s-protein signaling determined as the cAMP accumulation of WT and mutants upon stimulation with NDP-α-MSH (**D-E**), α-MSH (**F**), and setmelanotide (**G**) shown as fold change of WT max. signaling [%].

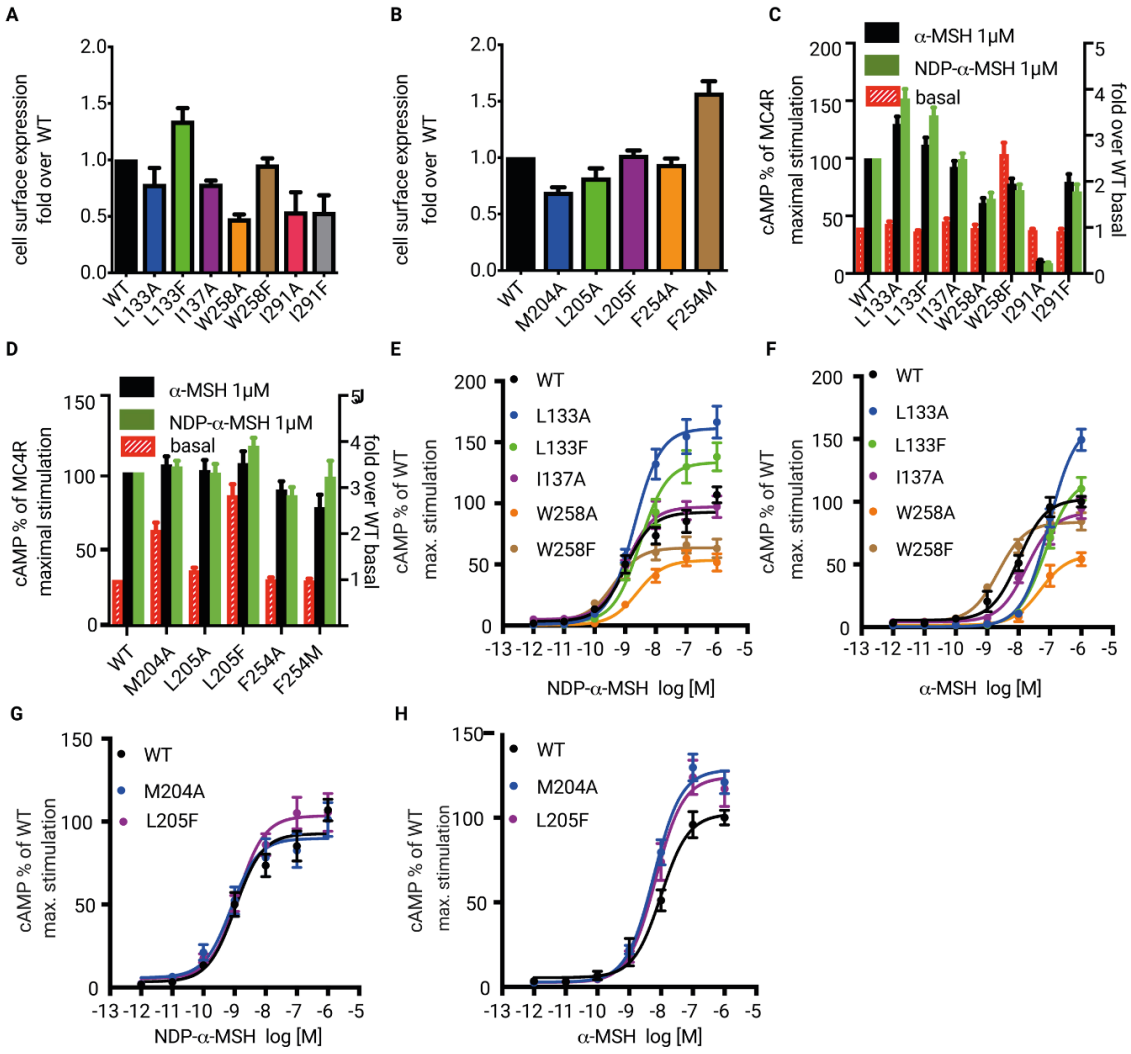

Figure 3.5.2.1 ***G_s signaling data of MC4R mutants with variants in the transmembrane region of MC4R.*** (**A-B**) Cell surface expression of wild-type MC4R (WT) and mutants with variants in (**A**) the hydrophobic box surrounding W258 and (**B**) the *MIF* motif and L205A are shown as fold-over WT. (**C-D**) Maximal G_s-protein signaling determined as the cAMP accumulation of MC4R WT and mutants with variants in (**C**) the hydrophobic box surrounding W258 and (**D**) the *MIF* motif and L205A after addition of 1 µM α-MSH or NDP-α-MSH indicated as fold of MC4R WT max. signaling [%] (left y-axis). Basal cAMP accumulation of MC4R WT and mutants are depicted as fold-over WT basal (right y-axis). (**E-H**) Concentration-response curves of G_s-protein signaling determined as the cAMP accumulation of MC4R WT and indicated mutants with variants in (**E-F**) the hydrophobic box surrounding W258 and (**G-H**) the *MIF* motif and L205A upon stimulation with NDP-α-MSH (**C** and **G)** and α-MSH (**F** and **H**) shown as fold of MC4R WT max. signaling [%].

3.5.2 MC4R variations in the transmembrane core

MC4R variants with mutations in the transmembrane core were investigated for their ability to induce G_s signaling upon stimulation with NDP-α-MSH and α-MSH. The central residue of the transmembrane core of MC4R is $W258^{6.48}$. Therefore, we focused on residues forming a hydrophobic box surrounding $W258^{6.48}$ in TM3 (L133A, L133F, I137A, I137F), in TM5 (M204A, L205A), in TM6 (F254A, F254M, W258A, W258F), and TM7 (L288A

I291A, and I291F (Table 3.5.1). Variations in the transmembrane core do not increase the cell surface expression and have no significant impact on G_s signaling potency (EC_{50}) (Tables 3.5.2). MC4R variations W258A, W258F, and I291A, induce a drop in efficacy compared to WT upon α-MSH stimulation (Figure 3.5.2.1C). L133A and L133D display a modest increase in efficacy upon stimulation with α-MSH than WT (Figure 3.5.2.1C). This effect is weakened for mutant L133A and gone for mutant L133F upon stimulation with NDP-α-MSH.

The MC4R variants M204A, I137A, I137F, F254A, and F254M, are part of the *MIF* motif. These have no significant impact on potency (Tables 3.5.2) and efficacy (Table 3.5.1). However, M204A and the neighboring residue L205A display a substantial increase in basal activity (Table 3.5.1).

In summary, in the MC4R core, we identified variants at two residues, $W258^{6.48}$ and $I291^{7.42}$, that can disturb MC4R signaling. Furthermore, residues of the *MIF* motif have no relevance for ligand-dependent G_s stimulation, but mutants M204A and L205A display an increase in basal activity compared to WT.

3.5.3 Mutations in the MC4R G-protein interface

Two characteristics at the G-protein interface were examined via signaling experiments. On the one hand, we studied $H158^{IL2}$ at IL2 that displays a double rotamer in the NDP-α-MSH-MC4R-G_s complex and only one rotamer in the setmelanotide-MC4R-G_s complex. On the other hand, we examined the novel receptor-G_s interaction of MC4R

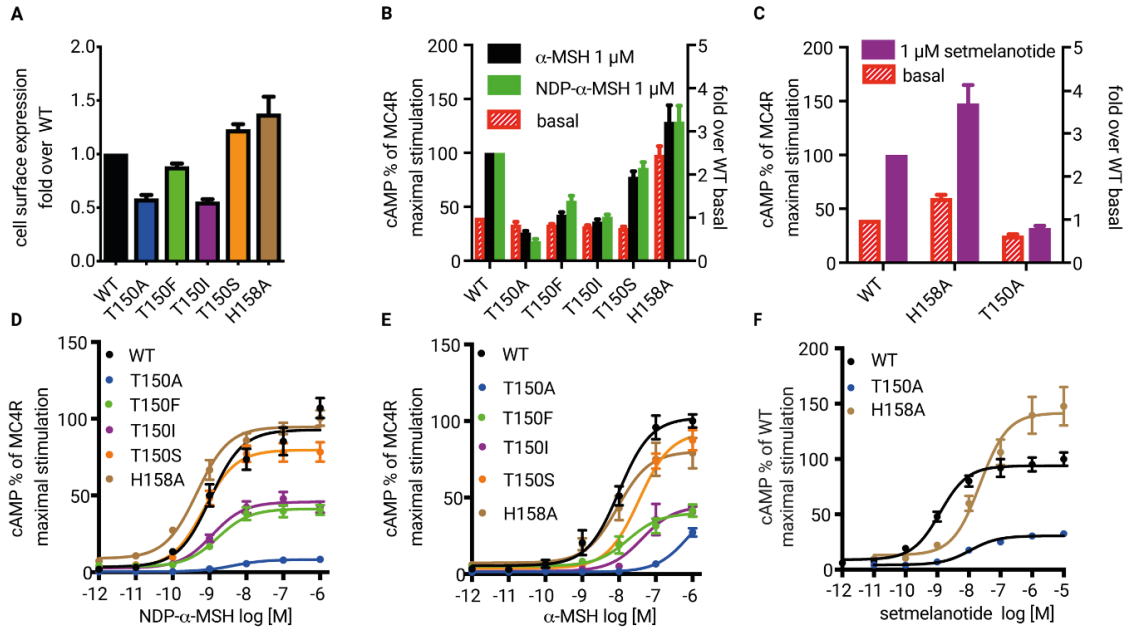

Figure 3.5.3.1 ***G_s signaling data of MC4R mutants with variants in the G_s-protein binding region of MC4R.*** (**A**) Cell surface expression of wild-type MC4R (WT) and mutants are shown as fold-over WT. (**B-C**) Maximal G_s -protein signaling is determined as the cAMP accumulation of MC4R WT and mutants after adding 1 μM α-MSH, NDP-α-MSH (**B**), and setmelanotide (**C**) indicated as fold of WT max. signaling (left y-axis) [%]. Basal cAMP accumulation of MC4R WT and mutants are depicted as fold-over WT basal (right y-axis). (**D-F**) Concentration-response curves of G_s -protein signaling determined as the cAMP accumulation of MC4R WT and indicated mutants after the challenge of NDP-α-MSH (**D**)**,** α-MSH (**E**)**,** and α-MSH (**F**) shown as fold of WT max. signaling [%].

$T150^{3.53}$ with $G\alpha_s$ $Y377^{\alpha5.23}$. Cell surface expression of MC4R variant T150A and T150I dropped to ~50% and increased for T150S (Table 3.5.1 and Figure 3.5.3.1A). The potential loss of hydrogen bond due to the substitutions T150A, T150D, T150F, and the pathogenic mutation T150I result in a substantial drop in efficacy for G_s signaling upon stimulation with NDP-α-MSH and α-MSH (Figure 3.5.3.1D-E and Table 3.5.1). Upon setmelanotide stimulation displays the T150A mutant a drop in efficacy in G_s signaling and no change in potency compared to WT (Figure 3.5.3.1.C and Tables 3.5.1-2). MC4R variant T150S displays the same WT efficacy upon stimulation with NDP-α-MSH and α-MSH.

MC4R variant H158A does not reduce the efficacy for G_s signaling upon stimulation with NDP-α-MSH and α-MSH and causes mild hyperstimulation upon stimulation with setmelanotide. In addition, the basal stimulation is increased by 2.5-fold (Figure 3.5.3.1B).

The G_s signaling data of the two side-chain variants at the G-protein interface indicate that a hydroxyl group at the 150 position is mandatory for G_s signaling, and $H158^{IL2}$ plays a role in the basal state equilibrium.

The $G_{q/11}$ signaling data are notable in the G-protein interface. Remarkably, the T150A mutant has upon stimulation with NDP-α-MSH no reduction in potency for $G_{q/11}$ activation, but with setmelanotide a 24-fold reduction; and both ligands have no significant effect on (Figures 3.5.1D-E). This observation is in strong contrast to the essential role of $T150^{3.53}$ for G_s signaling in the T150A variant (Figure 3.5.3.1D-E).
The IL2 mutation, H158A, does not affect G_s-mediated signaling but reduces its efficacy for $G_{q/11}$ signaling to ~50% for both ligands (Figures 3.5.1.D-E, Table 3.5.1, and Table 3.5.3).

In conclusion, while $T150^{3.53}$ is essential for G_s binding, the T150A mutant does not impair $G_{q/11}$ signaling efficacy but displays a ligand-dependent shift in potency. At IL2, $H158^{IL2}$ is not relevant for G_s but $G_{q/11}$ coupling.

4 Discussion

In the following discussion, the reader will be guided from the ligand-binding pocket via the receptor transmembrane core to the G-protein interface. A potential trajectory of receptor activation will be indicated. The binding of the agonists and the G-protein stabilize a conformation that is characterized by local readjustments in the microswitch regions *CWxP*$^{6.50}$ (C257$^{6.47}$- W258$^{6.48}$-P260$^{6.50}$, in TM6), *N(D)P*$^{7.50}$*xxY* (D298$^{7.49}$-P299$^{7.50}$-Y302$^{7.53}$ in TM7), and *DR*$^{3.50}$*Y* (D146$^{3.49}$-R147$^{3.50}$-Y148$^{3.51}$, in TM3) compared to the antagonized MC4R structure. These local rearrangements are associated with helical movements characteristic of an active state receptor structure, enabling G-protein binding. These helical movements are mainly a ~13 Å outward motion of TM6 and an inward twisting of TM5 (Figure 4.1).

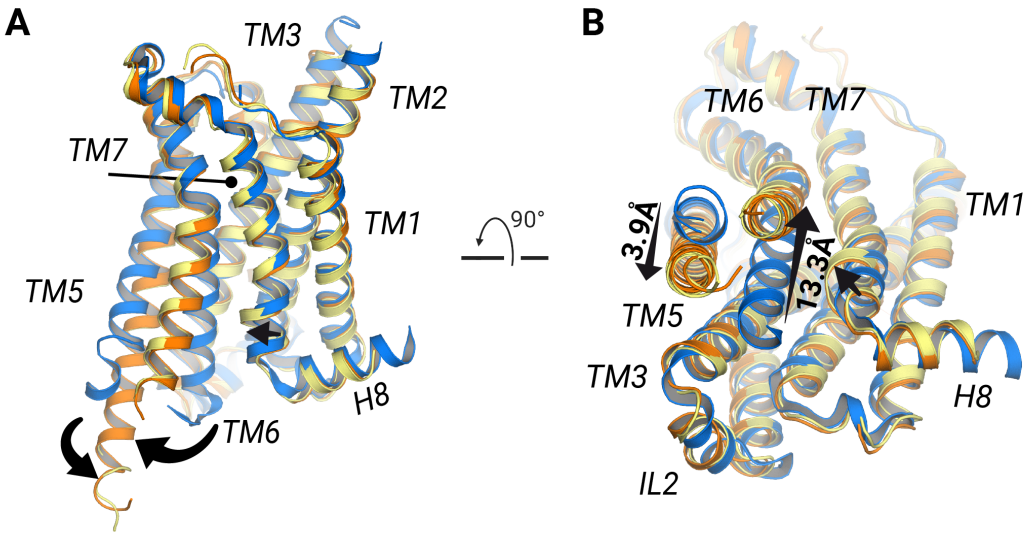

Figure 4.1 ***The TM6 outward movement and TM5 inward movement define the active state MC4R.*** (**A-B**) Superposition of the SHU9119-antagonized (blue) MC4R and the active state structures (NDP-α-MSH bound MC4R - orange, setmelanotide bound MC4R - yellow) (**A**) from inside the membrane plane, (**B**) from the cytoplasmic view. Arrows indicate significant transmembrane helix rearrangements (TM5, TM6, and TM7). The TM6 outward movement is measured at the Cα atom of M241 and the TM5 inward movement at Cα of V213.

4.1 Ligand action at the MC4R

4.1.1 The pharmacological profile of setmelanotide and NDP-α-MSH

Setmelanotide, the first FDA-approved anti-obesity drug targeting MC4R, potentially facilitates its appetite-reducing action via a more pronounced activation of the $G_{q/11}$ pathway in comparison to α-MSH [102]. NDP-α-MSH and setmelanotide share the same potency for inducing G_s and $G_{q/11}$ signaling but deviate in their efficacy. NDP-α-MSH is more efficacious in G_s signaling and setmelanotide in $G_{q/11}$ signaling (Figure 3.5.2). Deviating pharmacological ligand profiles indicate that the compared ligands have a pronounced stabilization of receptor conformations that favor either G_s or $G_{q/11}$ protein coupling. Therefore, the ligand-binding modes of setmelanotide and NDP-α-MSH were compared (Section 3.3.2) and analyzed for potential clues indicating a molecular origin of this G-protein signaling bias. Since both structures are G_s-coupled complexes and an MC4R-$G_{q/11}$ has not been determined, signaling studies need to assess the relevance of the observed deviations.

Comparison of the two structures bound to the agonist shows a slight shift of TM3 in the horizontal plane and two significant differences in the interaction pattern with the receptor.

Firstly, NDP-α-MSH is in direct contact with TM4 via $I185^{4.61}$, and setmelanotide lacks that interaction but interacts via $R6^{2}$ with $N123^{3.26\ at}$ TM3 (Figures 3.3.2C-D). Mutagenesis experiments at $I185^{4.61}$ are not reasonable since the $R8^{2}$ (NDP-α-MSH) interaction is via the backbone of $I185^{4.61}$. Therefore, we cannot assess if the $I185^{5.61}$ interaction is relevant for NDP-α-MSH's mode of action. The MC4R variant N123A compared to WT does not show a drop in cAMP accumulation upon NDP-α-MSH and α-MSH stimulation (Figure 3.5.1.1D). However, setmelanotide displays a 6-fold reduced potency for mutant N123A compared to WT (Figure 3.5.1.1G and Table 3.5.2). Hence, the G_s signaling data confirm the relevance of setmelanotide's interaction with $N123^{3.26}$ and both agonists' different ligand binding modes.

The second difference in the binding modes is the direct interaction of setmelanotide $R1^{-3}$ with $D122^{3.25}$ (Figure 3.3.2B), which has no counterpart in the NDP-α-MSH structure (Figure 3.3.2A).

Mutagenesis studies at the three acidic residues coordinating Ca^{2+} are problematic because changes in the Ca^{2+} binding could mainly drive functional changes; hence the hydrophilic side chain properties are maintained by choosing the D122S instead of a D122A mutant. The G_s and $G_{q/11}$ signaling data display a reduced dependency of NDP-α-MSH-induced signaling on $D122^{3.25}$. In contrast, α-MSH and setmelanotide display no potency for inducing G_s signaling on the D122S mutant (Table 3.5.2). The deviating responses on the D122S mutant confirm the relevance of $R1^{-3}$ for the ligand-binding mode of setmelanotide to MC4R and imply that α-MSH binding relies as well on direct interaction with $D122^{3.25}$.

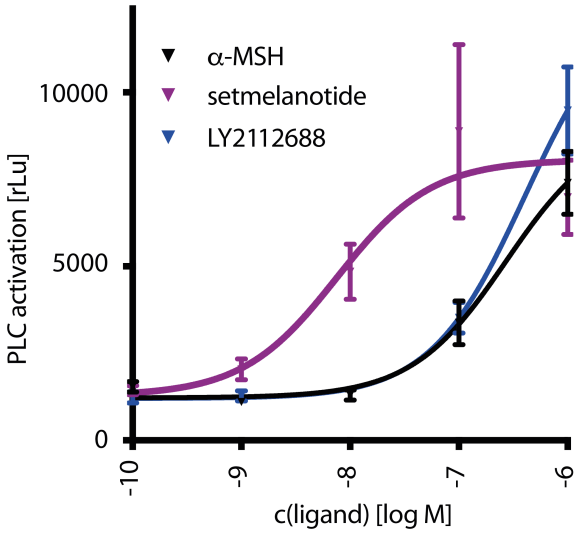

Figure 4.1.1 ***$G_{q/11}$ signaling data displays a 50-fold reduced G_q activation upon stimulation with LY2122688 compared to setmelanotide.*** HEK293 cells transfected with MC4R were stimulated with varying concentrations of α-MSH, setmelanotide, and LY2112688. Data published in Clement *et al.* 2018 [102].

To understand more about the biased signaling profile of setmelanotide, we searched the literature for ligands with varying primary sequences in the X^{-3} position. The MC4R agonist LY2112688 deviates for two positions $E3^{-1}$ instead of D-$A3^{-1}$, and D-$R1^{-3}$ instead of L-$R1^{-3}$ (sequence shown in Figure 1.3.1). The stimulation with LY2112688, in contrast to setmelanotide, induces a 50-fold reduced $G_{q/11}$ stimulation but only a 4-fold decrease in G_s signaling [102]. Hence, both substitutions cause setmelanotide's $G_{q/11}$ bias in relation to α-MSH (Figure 4.1.1). The setmelanotide-MC4R-G_s complex structure strongly suggests that the stereoisomer L-$R1^{-3}$D-R substitution is the primary factor for this bias. This stereoisomer substitution could distort the $R1^{-3}$-$D122^{3.25}$ hydrogen bond in LY2112688. In addition, the direct ligand interaction with $D122^{3.25}$ could be the origin of the deviating positioning of TM3 in the upper half of the receptor in both agonist-bound structures.

Nevertheless, a better understanding of the relevance of TM3 and a potential connection between the two TM3 conformational states and G-protein specificity should be investigated by signaling and structural studies of peptide variants targeting the interaction with TM3.

In addition, structures of the endogenous agonists α-MSH and β-MSH - the precursor peptide for the design of NDP-α-MSH and setmelanotide, respectively - might indicate which TM3 state resembles an endogenous active MC4R state. Although NDP-α-MSH is a high-affinity variant of α-MSH, the signaling data of setmelanotide indicate that α-MSH and setmelanotide share comparable ligand-receptor interactions in the EDD motif, in contrast to NDP-α-MSH

4.1.2 Origin of SHU9119 antagonism at the MC4R

The main difference between the antagonist SHU9119 and both agonists is the substantial unnatural amino acid D-naphthylalanine (D-Nal4^{1}) in SHU9119 at position x^{1} in the agonist *H^{0}F^{1}R^{2}W^{3}* motif (which is termed *HxRW* motif in the context of the antagonist) that corresponds to D-F^{1} in NDP-α-MSH and setmelanotide. The superposition of the antagonized MC4R structure bound to SHU9119 with both agonist-bound structures displays a high similarity of the binding mode in the upper part of the LBP (Figure 4.1.2.1). To our surprise, the binding position of the peptide core is only slightly shifted, and the lactam circularization of SHU9119 (Figure 4.1.2.1A) does not show a significant change at the *HxRW* motif in comparison to NDP-α-MSH. Setmelanotide, as well circularized, displays only subtle deviations but is slightly rotated in a clockwise direction (top view perspective) in comparison to SHU911 (Figure 4.1.2.1).

The most apparent difference in the LBP compared to the agonist bound structures is the transposition of TM3 by 1 Å in the SHU9119-MC4R structure (Figure 4.1.2.1). The side chains of D122$^{3.25}$, N123$^{3.26}$, and D126$^{3.29}$ at TM3 accompany this movement. The structures indicate that two factors enable this adjustment at the ligand-receptor interface due to the TM3 transposition. Firstly, Ca^{2+} remains centered according to the EDD motif's coordination geometry and is also located approximately 1 Å deeper into the LBP (Figure 4.1.2.1). Secondly, the sidechain of R^{3} allows maintaining the hydrogen bond network between the ligands and the receptors.

Ultimately, the superposition of the antagonized with both active state structures in the upper part of the LBP does not indicate the origin of the downward transposition of TM3. However, the comparison of the transmembrane interactions of D-Nal4[1] in SHU9119 with D-F[1] in both agonists displays impressively how SHU9119 impedes receptor activation (Figure 4.1.2.2).

The bulky aromatic side chain of D-Nal pushes the side chain of L133[3.36] down in comparison to the corresponding D-F of both agonists that are not in contact with L133[3.36] (Figure 4.1.2.2). This transposition is accompanied by a vertical downward shift in the upper half of TM3 by approximately 1 Å (Figure 4.1.2.1). Therefore, the side chain of L133[3.36] in the antagonized receptor is in front of W258[6.48], eliminates the TM6 outward movement, and prevents G-protein binding (Figures 4.1.2.2C-D). A single mutation functionally supports this structural observation at L133[3.36] to a more flexible methionine side chain that turns SHU9119 to an MC4R agonist [214].

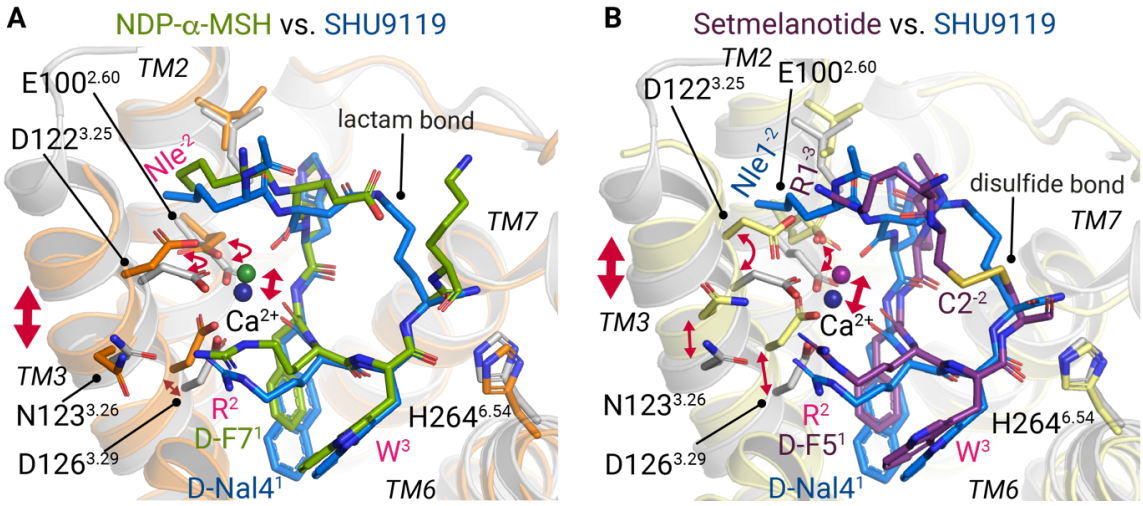

Figure 4.1.2.1 ***SHU9119 induces a vertical transposition of TM3 accompanied by Ca^{2+}.*** Comparison of ligand binding modes of NDP-α-MSH (residues 4-11), setmelanotide, and SHU9119 at MC4R (orange, light-yellow, and gray backbone cartoons). SHU9119 binding to MC4R is compared with (**A**) NDP-α-MSH, and in (**B**) with setmelanotide. Bi-directional red arrows indicate differences in the relative positioning of residues, the TM3, or the Ca^{2+} ion. Intermolecular interactions are summarized in Table 3.5.1.

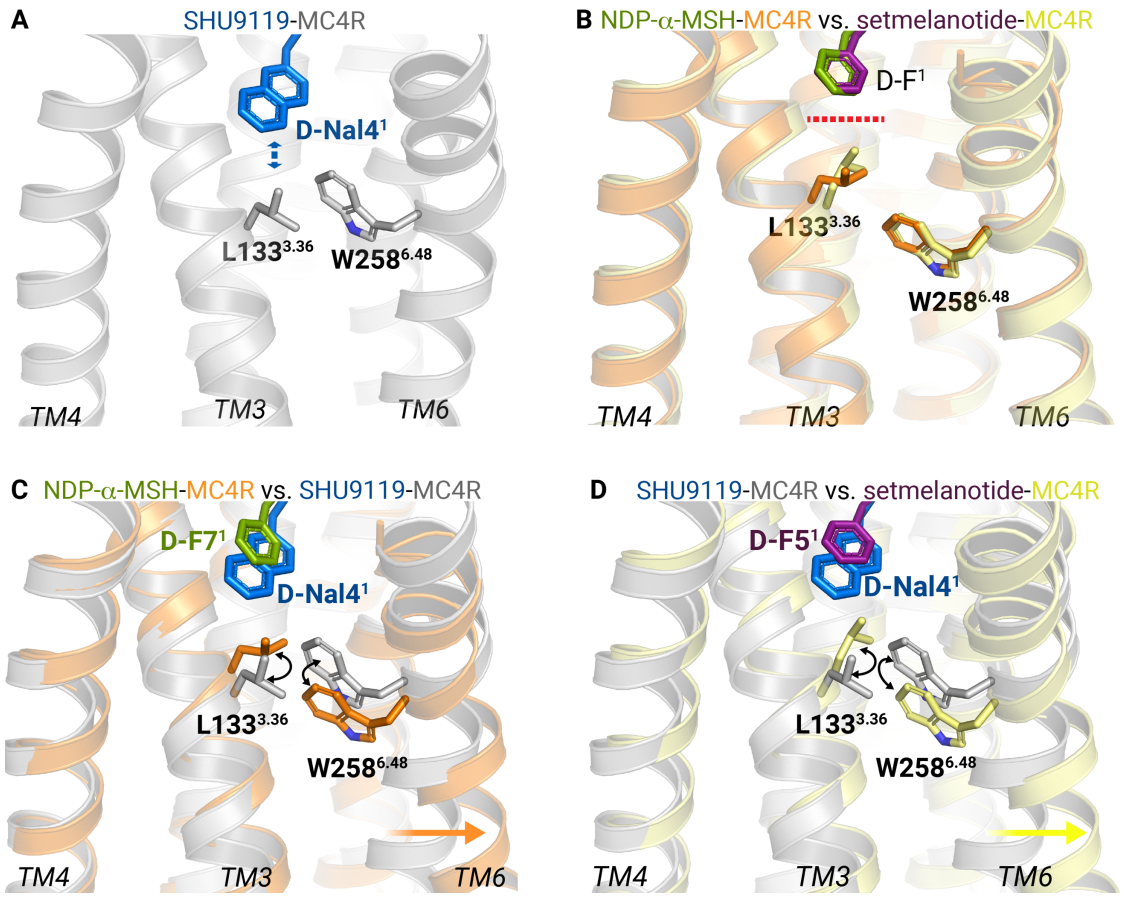

Figure 4.1.2.2 SHU9119s *D-Nal4¹ pushes down L133³.³⁶ in front of W258⁶.⁴⁸, thereby antagonizing MC4R.* (**A**) A blue bidirectional arrow indicates the interaction D-Nal4¹ of SHU9119 (blue) with L133³.³⁶ of MC4R (grey). (**B**) Superposition of NDP-α-MSH-MC4R with a red dashed horizontal line indicating that DF¹ does not interact with L133³.³⁶ in both agonist-bound structures. (**C-D**) The superposition of the (C) NDP-α-MSH-MC4R and (**D**) setmelanotide-MC4R complex structures with SHU9119 is shown. Black bidirectional images highlight the transposition of L133³.³⁶ and W258⁶.⁴⁸.

This experiment was based on the observation that SHU9119 is an agonist for the MC1R and MC5R with a natural M128³.³⁶ or V126³.³⁶ at this position. The recently preprinted cryo-EM structure of the active MC1R-G_s complex with SHU9119 confirms the here described mechanism of L133³.³⁵ mediated MC4R antagonism [215]. Comparing the active state structures of MC1R and MC4R is the prerequisite for the rational design of subtype-selective MC4R or MC1R ligands.

Complementarily, Melanotan II [216] an MC4R agonist, shares the identical sequence with SHU9119 except that it contains a D-F¹ instead of the D-Nal4¹ (see the sequence in Figure 1.3.1). Our functional characterization supports the observation that the smaller

D-F[1] does not affect the position of L133[3.36]. Since the L133[3.36] substitutions to either alanine or phenylalanine reveal no impact on NDP-α-MSH signaling (Figure 3.5.2.1).
In summary, SHU9119 acts as an antagonist due to the D-Nal4[1]-L133[3.36] interaction. Even though the antagonist has a similar binding pattern in the upper part of the MC4R LBP compared to the agonists. The D-Nal4[1]-L133[3.36] interaction also strengthens a hypothesis that W258[6.48] is crucial for MC4R activation (see also Section 4.2.1) since its stabilization by L133[3.36] stabilizes a conformation in which W258[6.48] cannot induce a TM6 kink.

4.1.3 A calcium ion acts as a cofactor for ligand binding

Thus far, no class A GPCR except MC4R has been shown to be modulated by Ca^{2+} and in which manner Ca^{2+} is involved in MC4R signaling regulation is still under investigation. Distinguishing a Ca^{2+} ion of Mg^{2+} ion is not possible at a 2.9 Å or 2.6 Å resolution. The assignment of the vacant electron density in both agonist-bound MC4R structures at the peptide-TM2-TM3 interface to a Ca^{2+} ion was reasoned by three factors. Firstly, the coordination of the surrounding residues indicates a divalent cation. Secondly, during the complex purification, Mg^{2+} ions are solely present until the extensive washing steps of 100 column volumes over two hours at the M1-FLAG antibody column, which are performed in the presence of 2.5 mM $CaCl_2$.

Furthermore, the elution of the M1-FLAG antibody column is driven by chelating divalent cations by adding 5 mM EDTA. Both complexes were incubated overnight in this elution buffer containing 5 mM EDTA during H3C protease cleavage (Section 2.1.3). Consequently, all divalent cations have been removed from the equilibrium, and during the subsequent gel filtration step, 1 mM Ca^{2+} was added but no other divalent cations. Hence, the presence of another divalent cation except Ca^{2+} is not probable. Thirdly, numerous studies have shown the relevance of Ca^{2+} for melanocortin binding (Section 1.3.3), NDP-α-MSH [91], and setmelanotide [92]. In addition, the structures of SHU9119-MC4R [91] and the setmelanotide-MC4R- G_siN structure [92] have a Ca^{2+} ion assigned at the correlating position.

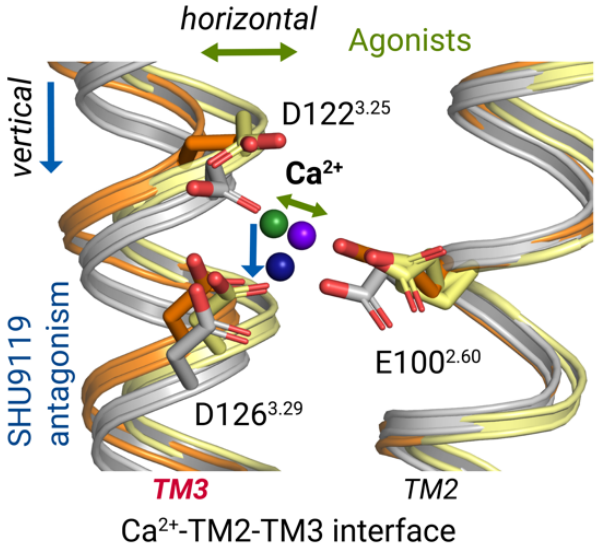

Figure 4.1.3 ***Ca^{2+} connects ligands, TM2 and TM3.*** MC4R helices TM2 and TM3 are displayed as cartoons (NDP-α-MSH-MC4R-G_s in orange, setmelanotide-MC4R-G_s in yellow, and SHU9119-MC4R in gray). The residues of the *EDD* motif are shown as sticks and Ca^{2+} as spheres (NDP-α-MSH-MC4R-G_s in green, setmelanotide-MC4R-G_s in purple, and SHU9119-MC4R in blue).

The relative positioning of TM3 observed in different ligand-MC4R complexes correlates with the positioning of the Ca^{2+} ion (Figure 4.1.3) and indicates that MC4R's orthosteric site is adjustable here. In contrast to other GPCRs, this can be explained by D122$^{3.25}$ in TM3, which is in most other GPCR's a cysteine involved in a disulfide bond to EL2. Structures of ligands in varying sizes like small molecules MCL1029 (MW = 547 Da) [217] or elongated peptide ligand AgRP(87-132) (MW = 5.2 kDa) could shed light on a potential size adaptive characteristic of the MC4R orthosteric site.

Biochemical experiments not part of this thesis indicate that forming the agonist-MC4R-G_s complex requires Ca^{2+}. The removal of Ca^{2+} by adding EDTA results in complex dissociation for low-affinity agonists like α-MSH (unpublished data, manuscript in preparation). Furthermore, the *in vitro* MC4R-G_s complex formation depends on Ca^{2+} but not on the presence of an agonist. This apo complex was not sufficiently stable to sustain vitrification, but the complex was formed *in vitro* (unpublished data). According to these observations, MC4R is stabilized by Ca^{2+} in an active conformation to a certain degree. Of note, the binding affinity of AgRP to MC4R, the known inverse agonist/antagonist of MC4R [90], is reduced in the presence of Ca^{2+} [91]. It would be interesting to decipher if MC4R's basal activity could be diminished by removing Ca^{2+} from the equilibrium. Here, cell-based assays are not ideal systems, since the depletion of Ca^{2+} impairs various vital cell functions. A fluorescence-based *in vitro* GPCR G-protein coupling assay in lipid nanodiscs could shed light on this question.

Ultimately, the question arises if Ca^{2+} is an allosteric modulator or a permanent cofactor of MC4R ligand binding. Extracellular Ca^{2+} ion concentration is between 1.1 - 1.4 mM [218], which is significantly above the EC_{50} of Ca^{2+} in the range of 10 - 100 µM [136] in the presence of NDP-α-MSH binding. The cytoplasmic Ca^{2+} concentration is 100 nM [218], and Ca^{2+} might not be bound when receptors are internalized. One could speculate that activating phospholipase C and thus Ca^{2+} ion channels could generate a feedback loop on MC4R activation by reducing Ca^{2+} levels in the close vicinity of MC4R's LBP.

4.1.4 Ligand-dependent activation trigger in MC4R

Orthosteric ligands influence the nucleotide-binding affinity of receptor-coupled G-proteins via the allosteric connection between the receptor and the nucleotide-binding site [198]. For instance, how agonists stabilize or induce a conformation with an outward tilted TM6 could be shown for the binary switch rhodopsin [33] but not for all receptors with active state structures. In the case of MC4R, this question cannot be tackled using the structures solely since the ligand binding mode of SHU9119 is comparable to both agonists in the upper part of the orthosteric site (Section 3.3.4). In addition, the agonists are not in contact with $W258^{6.48}$ or its surrounding residues (Figure 4.1.2.2B). Therefore, the relevant interactions occur in the upper part of the orthosteric site. The signaling data highlights four MC4R variants that result in a severe reduction of cAMP accumulation upon α-MSH stimulation and a significantly reduced potency. These residues are the three acidic residues forming the Ca^{2+} interface and $H264^{6.54}$ at TM6.

Residues that are essential for ligand binding display if mutated a change in potency (EC_{50}) since potency correlates with affinity [138]. In contrast, residues that are not contributing to ligand binding, but are relevant for receptor activation display a reduced efficacy without a change in potency (Figure 1.3.4).
$F261^{6.51}$ forms a hydrophobic interaction with the D-F^{1} residue in the setmelanotide-MC4R structure but does not directly interact with NDP-α-MSH (Figures 3.3.2A-B and Table 3.5.1). G_s signaling data for the F261V mutation confirmed this structural deviation in both complex structures. Here, a potency reduction in the cAMP accumulation was observed for setmelanotide (29 nM) compared to wild-type (1.2 nM) in contrast to an unchanged potency for NDP-α-MSH (Table 3.5.2). However, the efficacy of setmelanotide was comparable to

wild-type MC4R but significantly reduced for NDP-α-MSH (Figure 3.5.1.1C and Table 3.5.1).

The signaling data implies that the H264$^{6.54}$-F261$^{6.51}$ site acts as an initial agonist-dependent trigger, with H264$^{6.54}$ directly contacting both agonists, while F261$^{6.51}$ specifically interacts with setmelanotide. This trigger region appears to be essential in relaying the signal towards the adjacent helix-tilting TM6-*CWxP*$^{6.50}$ region in the transmembrane core (Figure 4.2.3.2).

4.2 Transmembrane core of MC4R

4.2.1 Toggle switch like motion of W258$^{6.48}$

In class A GPCRs, G-protein binding to the intracellular G-protein binding crevice is primarily enabled by the TM6 outward movement (~ 13 Å in MC4R, measured from M241-Cα) towards the membrane (Figure 4.1).

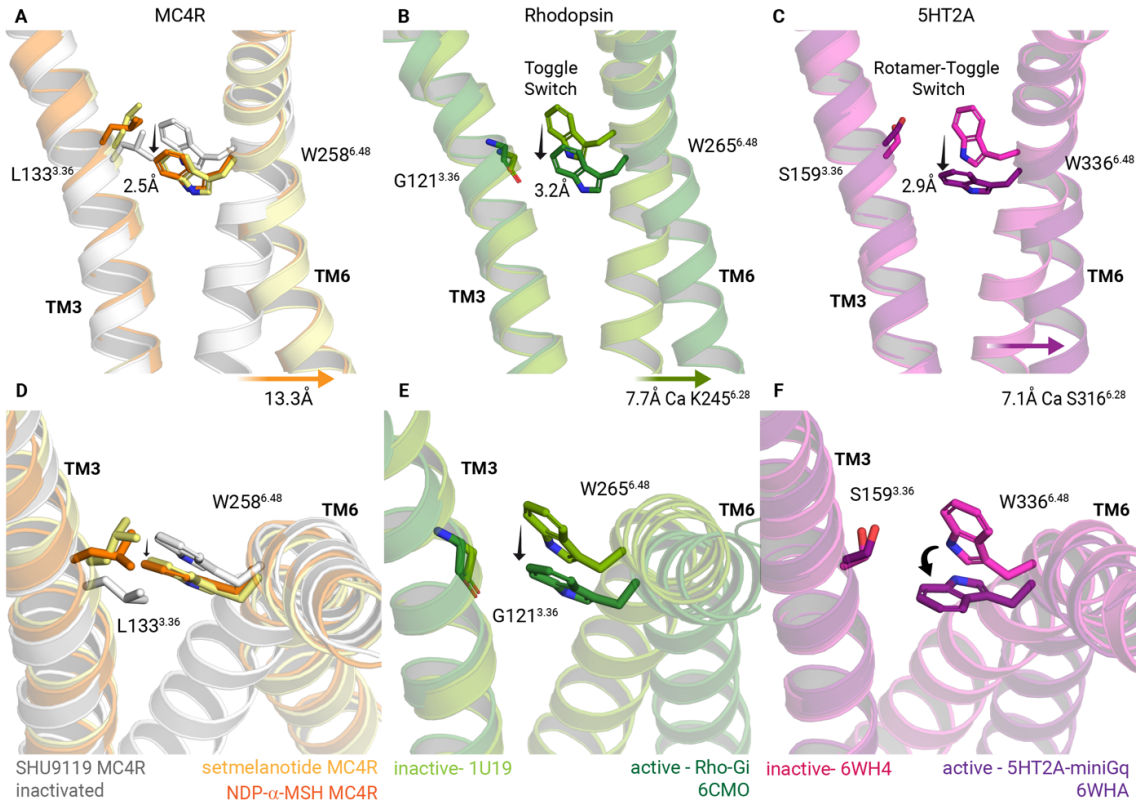

Figure 4.2.1 ***Localization of W$^{6.48}$ in different active state complex structures.*** Superposition of inactive and active state conformations of W$^{6.48}$ in (**A**) MC4R, (**B**) bovine rhodopsin (PDB IDs: inactive- 1u19, active- 6cmo), and (C) 5-HT$_{2A}$-receptor (5HT2A) (PDB IDs: 6wh4(inactive), 6wha(active)). The distances are associated with receptor activation. W$^{6.48}$ undergoes a vertical shift towards the intracellular site (measured at NE1 of W$^{6.48}$). In the 5HT2A structure, the W$^{6.48}$ side chain undergoes a rotation. The associated TM6 outward movements are measured at Cα positions of intracellularly located (**A**) M241$^{6.31}$, (**B**) K245$^{6.28}$, and (**C**) S316$^{6.28}$ and reflect the largest spatial difference of TM6 between active and inactive states. (**D-F**) Visualization of the horizontal W$^{6.48}$ shift from the top-view. The amino acid position 3.36 at TM3 opposite position 6.48 is different between all three receptors. However, in MC4R, the L133$^{3.36}$ with a large hydrophobic side chain directly impacts W$^{6.48}$. In the SHU9119 antagonized MC4R structure [91], this leucine inhibits an activation-related shift of TM6 at W$^{6.48}$.

In prototypical GPCRs such as rhodopsin [41], adenosine A1AR [58], and β2AR [20], this is facilitated by the kinking of TM6. The pivot point of the TM6 kink is the highly conserved *CWxP* motif (*CWAP* in MC4R), with W$^{6.48}$ displaying an extensive transposition towards the intracellular site. This observation has been described as the toggle switch activation mechanism in several receptors [219]. The comparison of MC4R with receptors assigned to the toggle switch theory displayed similar changes around W258$^{6.48}$ between the inactive or antagonized versus the active structures as observed, such as in rhodopsin [41] and the 5-HT2A serotonin receptor [56] upon activation (Figure 4.2.1). Hence, we observe a toggle switch-like movement of W258$^{6.48}$ during MC4R activation at the *CW$^{6.48}$xP* motif at TM6.

The signaling data of MC4R variant W258A leads not to a total loss of G_s signaling but to a significant reduction of cAMP accumulation (Figure 3.5.2.1C). A loss of structural integrity upon exchanging the aromatic moiety at the 258 position against a less hydrophobic sidechain is indicated by a reduced cell surface expression (Figure 3.5.2.1A). The W258F variant is expressed on the surface like wild-type with increased basal activity and a slightly reduced G_s signaling upon ligand stimulation than wild-type MC4R.
In summary, MC4R structures display a toggle switch-like movement of W258$^{6.48}$; however, the signaling data shows that it is not the only switch contributing to MC4R signal transduction in the transmembrane core.

4.2.2 Signal transmission inside the 7TM bundle

W258$^{6.48}$ is situated in a hydrophobic surrounding. This hydrophobic environment is constituted by L133$^{3.36}$, I137$^{3.40}$, F201$^{5.47}$, F254$^{6.44}$, F261$^{6.51}$, and I291$^{7.42}$; these display significant displacements upon receptor activation (Figure 4.2.2). G_s signaling data discriminates the MC4R activation relevant residues from non-relevant residues. The only residue with a total loss of function is I291$^{7.42}$ at TM7 (Figure 3.5.2.1C). This observation, combined with the interaction of I291$^{7.42}$ to W258$^{6.48}$ observed in the agonist-bound structures, implies that both side chains and their interplay are mandatory for MC4R activation.
The active structures further display that W258$^{6.48}$ hydrogen bonds to N294$^{7.45}$, which is coupled to D298$^{7.49}$ of the *N(D)P$^{7.50}$xxY* motif in TM7 (Figures 3.4.1.2D-F). The *N(D)P$^{7.50}$xxY* motif is slightly shifted toward TM3 (Figure 4.2.3.1B), accompanied by a

rotation of Y302$^{7.53}$ in the direction of Y212$^{5.58}$ in TM5. This orientates Y212$^{5.58}$ into a position that stabilizes the R147$^{3.50}$ in the *DR$^{3.50}$Y* motif, which is part of the G-protein interface.

The MC4R transmembrane core displays a loose allosteric connection of the orthosteric site with the G-protein interface via the conserved GPCR microswitches [3]. Furthermore, the signaling data reveals that W258$^{6.48}$ and I291$^{7.42}$ are key residues regulating the activation at the transmembrane core.

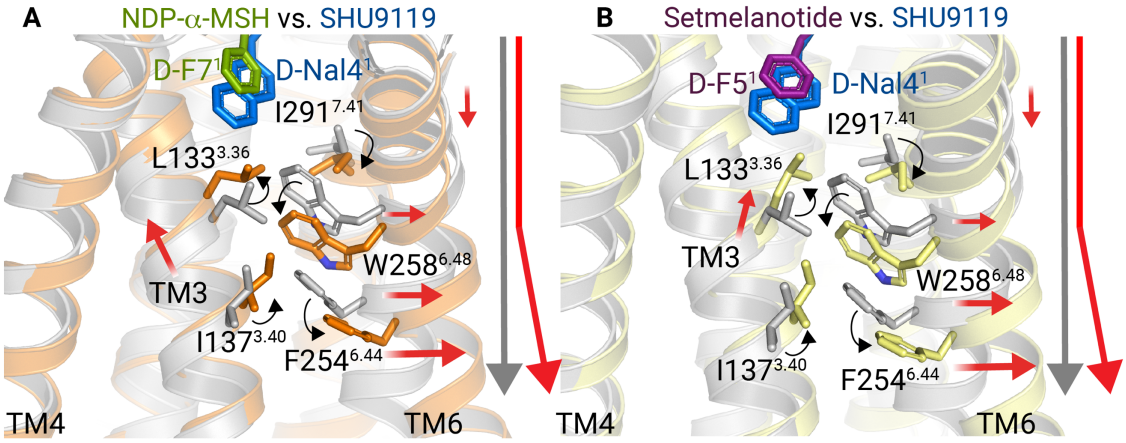

Figure 4.2.2 ***Hydrophobic residues surround W258$^{6.48}$ - this hydrophobic box translocate in concert upon activation.*** Superposition of (**A**) NDP-α-MSH-MC4R-G$_s$ or (**B**) setmelanotide-MC4R-G$_s$ with SHU9119-MC4R. Stick representation of hydrophobic residues in the vicinity of W258$^{6.48}$. Translocations upon activation are indicated by black arrows and the transpositions of TM3 and TM6 of the active states by red arrows. The grey arrow indicates the positioning of TM6 in the inactive state.

4.2.3 Microswitches govern MC4R activation

The static active-state agonist-MC4R-G$_s$ structures compared to the antagonized SHU9119 structure do not discriminate an order of events during the receptor activation pathway. Furthermore, the model of GPCR allostery (Section 1.1.3) claims that agonist binding induces changes in the G-protein interface and vice versa. The following model of MC4R activation (Figure 4.2.3.2) implies that an agonist binding event initiates the activation, leading to the outward movement of TM6 and allowing G-protein binding. This simplifies a process that is defined as a conformational ensemble, which is shifted by the presence of ligands and G-proteins. This thesis deciphers the sites of interest for receptor activation and evaluates the structural observations by signaling data.

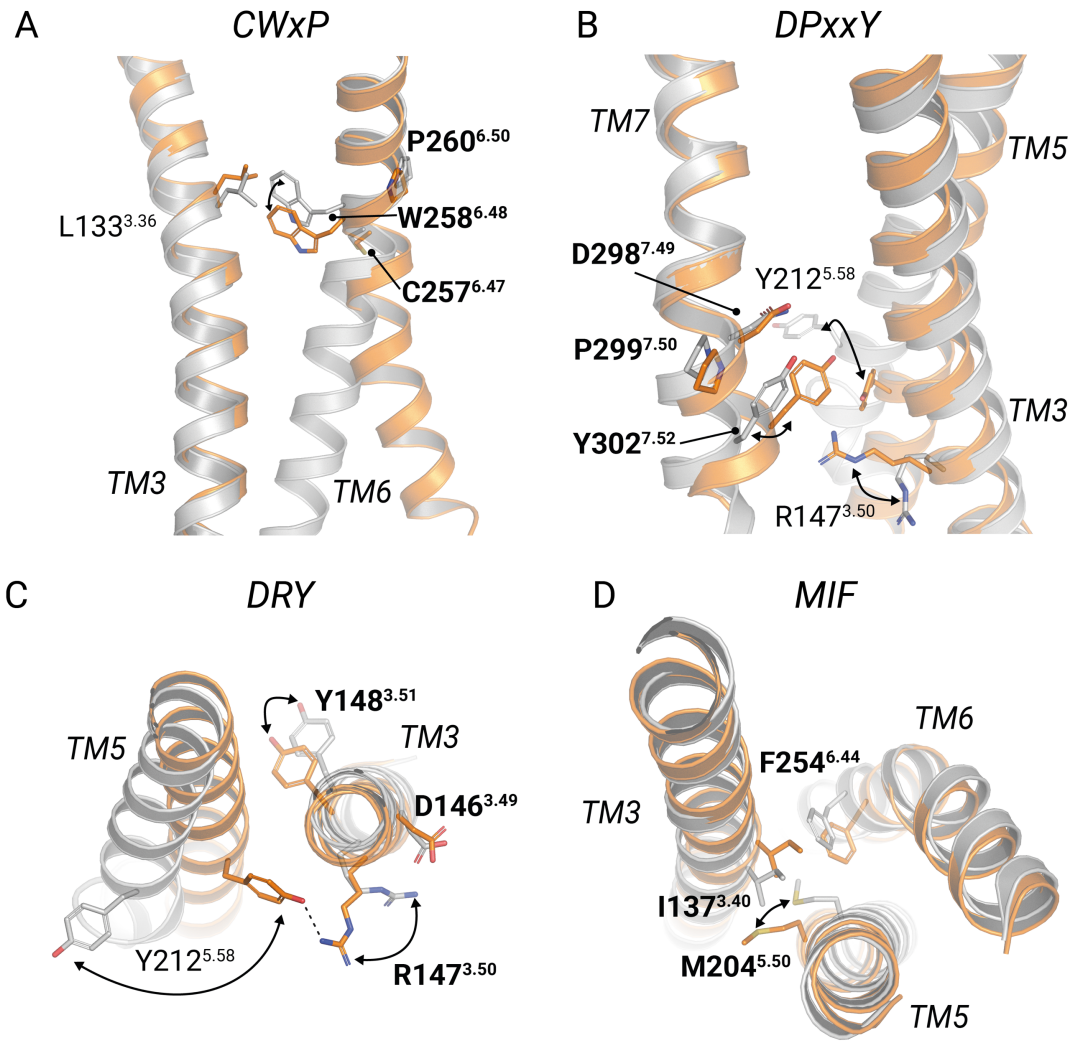

Figure 4.2.3.1 ***Comparison of highly conserved GPCR motifs between the antagonized and active MC4R structures.*** (**A**) The *CWxP* motif with $W^{6.48}$ has been described in several GPCRs as central for the "toggle switch" activation mechanism. In MC4R, $W258^{6.48}$ undergoes a toggle switch-like motion after agonist and G_s binding and is at the pivot point of the TM6 outward movement. (**B**) The conserved *NPxxY* motif of class A GPCRs is in MC4R a *DPxxY* motif. Upon activation, $Y302^{7.52}$ gets directed towards TM5 and TM7 and interacts with the $Y^{5.58}$ in TM5, which is crucial for active state stabilization. **(C)** The *DRY* motif is situated at the top of the G-protein binding cavity. $Y212^{5.58}$ stabilizes $R147^{3.50}$ in the active state. (**D**) Most class A GPCRs are characterized by proline in TM5 that induces a TM bulge. In MC4R instead, M2045.50 leads to a regular helical conformation without any king or bulge. Together with $I137^{3.40}$ in TM3 and $F254^{6.54}$ in TM6, they form the *MIF* motif, reminiscent of the *PIF* motif of class A GPCRs.

The structural data indicates that the four conserved microswitches *NPxxY*, *DRY*, *MIF,* and *CWxP* motifs are rearranged upon activation (Figure 4.2.3.1

Our comprehensive signaling data does not cover all relevant sites; thus, the functional relevance of the microswitch rearrangements cannot be answered fully by this data. Therefore, the following hypothesis is based as well on the screening of literature data. The signaling data here covered the *MIF* and W258$^{6.48}$ of the *CWxP* motif (Table 3.5.1). The mutation of the residues of the *MIF* motif did not result in a reduced cAMP accumulation but an increased basal signaling level (Table 3.5.1 and Figure 3.5.2.1). In addition, signaling studies of the D298A mutant indicate a loss of function [220], and an alanine scan of the D146$^{3.49}$, R147$^{3.50}$, and Y148$^{3.51}$ did not impair G_s coupling, but D146A induced a significant increase in basal signaling levels [221]. Combining our signaling data and literature data indicates that the *CWxP* and *NPxxY* motif are mandatory for G_s coupling. Furthermore, the *DRY* motif and the *MIF* motif may stabilize inactive conformations since their disruption increases basal signaling levels.

Model of MC4R activation

The structural comparison of the antagonized [91] and active-state structures in combination with the signaling data indicates a potential route of MC4R activation (Figure 4.2.3.2). Agonist binding depends primarily on four MC4R residues, the three acidic residues of the *EED* motif coordinating Ca^{2+} and H264$^{6.54}$ at TM6. Here at TM6, the agonist induces the activation along the H264$^{6.54}$ -F261$^{6.51}$ region connecting to the *CWxP* motif. W258$^{6.58}$ undergoes during activation a vertical downward shift towards the intracellular space. I291$^{7.41}$ at TM7 is an essential player during MC4R activation, and the complex structures imply that I291$^{7.41}$ interplays with W258$^{6.48}$. The outward movement of TM6 enables G-protein binding via the α5-helix.

Furthermore, W258$^{6.48}$ is in hydrogen-bond contact to N294$^{7.45}$ at TM7, further connected to the *DPxxY* motif and loosely to the *DRY* motif at the top of the G-protein binding cavity. Hence, MC4R's regulatory switches connect the ligand-binding pocket loosely with the G-protein interface, and conversely, G-protein binding could induce a reverse order of events. Molecular dynamics simulations done by the Hildebrand group (data in preparation)

can potentially shed light on the activation trajectory, but the here shown data display that the presented residues and sites are essential for the MC4R activation.

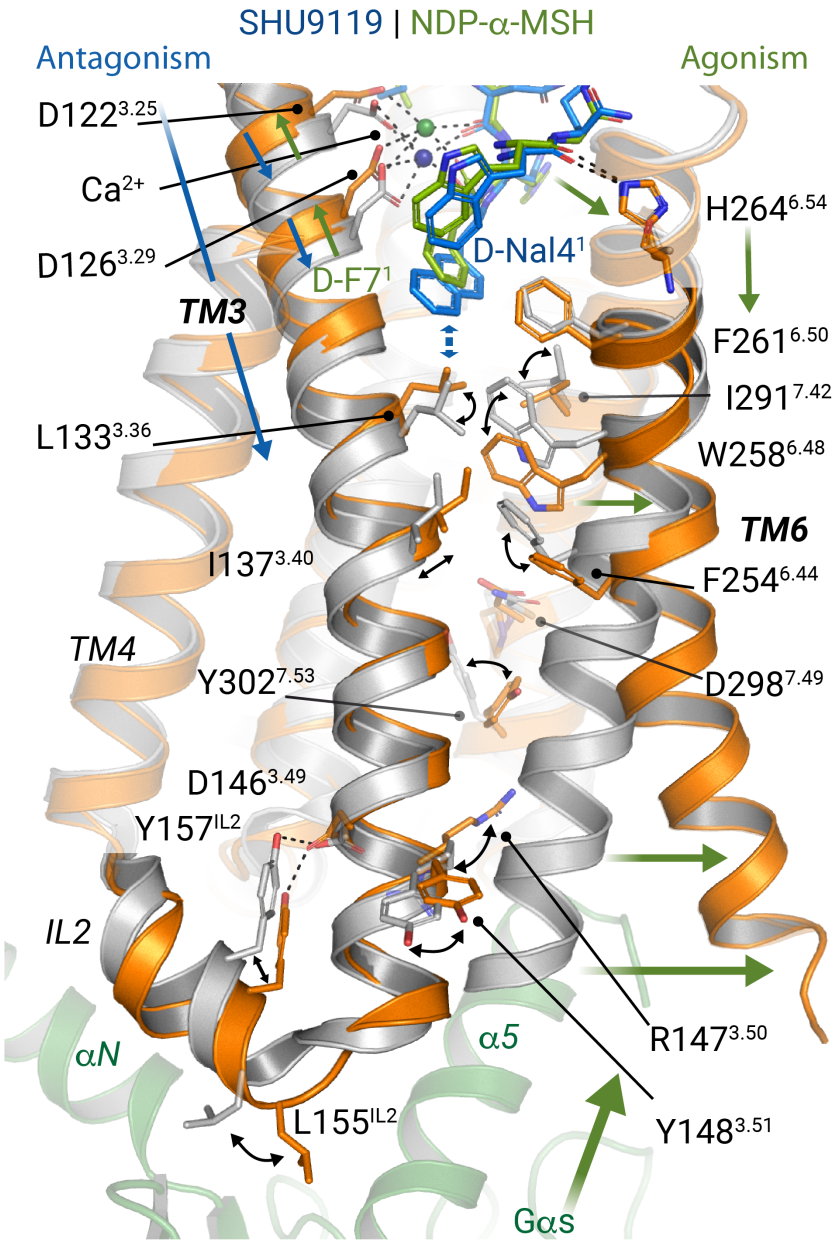

Figure 4.2.3.2 ***Model of MC4R receptor activation.*** Superposition of the agonist NDP-α-MSH (green/orange) or the antagonist SHU9119 (blue/gray) bound to MC4R. The proposed activation-pathway from the LBP to the G_s-protein interface is highlighted by green arrows and is triggered by ligand interactions at TM6, inducing the TM6 opening around W258$^{6.48}$. The interplay of D-Nal4^{1} facilitates antagonistic action with L133$^{3.36}$ (indicated by blue arrows) and a subsequently blocked "toggle switch" at W258$^{6.48}$.

4.3 MC4R-G_s protein binding

4.3.1 Agonist-specific IL2 conformation – a potential site for MC4R G-protein promiscuity

The comparison of the antagonized SHU9119-MC4R structure and the active state MC4R structures with the inactive and active β2AR displays typical class A GPCR features and MC4R specificities at the G-protein interface (Figure 4.3.1).

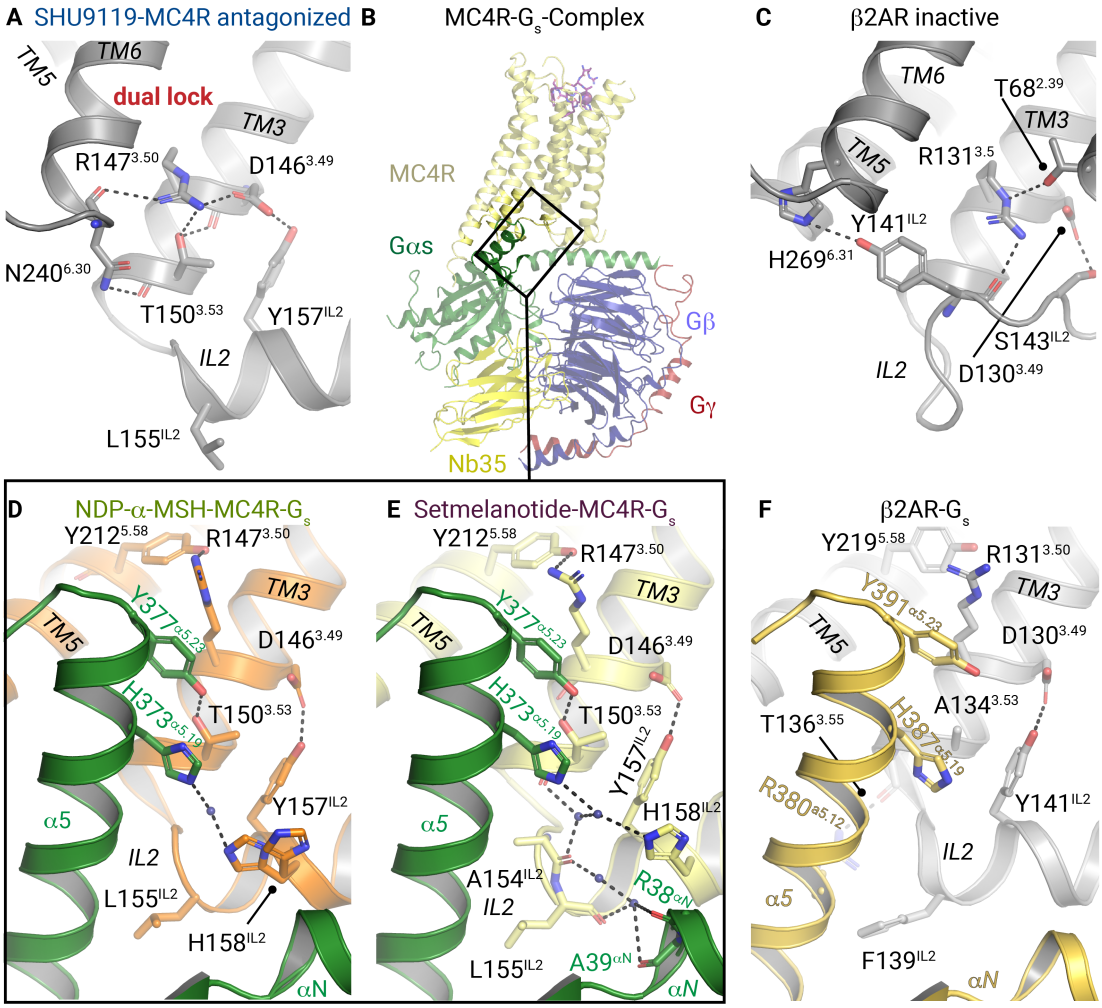

Figure 4.3.1 ***Receptor-G_s binding interface comparison between MC4R and β2AR.*** Intramolecular interactions at the IL2-TM3-TM6 site stabilizing (**A**) the antagonized MC4R (PDB ID: 6w25), and (**C**) the inactive β2AR (PDB ID: 2rh1). **(B)** Overall view on MC4R-G_s with enlarged areas in the panel shown in **(D-E).** Active receptor-G_s complex structures of MC4R bound to (**D**) NDP-α-MSH, (**E**) setmelanotide, and (**F**) β2AR (PDB ID: 3sn6). Black dashed lines indicate hydrophilic interactions.

The main interface of MC4R with G_s is at TM5, TM3, and its extension IL2 (Section 3.4.2). At TM3, we identified a previously unknown hydrogen bond interaction of MC4R with the end of the α5-helix of $G\alpha_s$ via $T150^{3.53}$. Furthermore, IL2 displayed an agonist-dependent positioning, indicating a region of increased flexibility. In which manner these observations have a functional relevance will be discussed.

4.3.1.1 $T150^{3.53}$ the critical residue in MC4R dependent G_s activation

Several inactive class A GPCRs are stabilized by a TM3-TM6 lock. In the antagonized MC4R structure, this lock is formed at the intracellular end of TM3 and TM6 via $R147^{3.50}$ and $N240^{6.30}$, respectively (Figure 4.3.1A). The corresponding interaction in dark state rhodopsin is the salt bridge between $R135^{3.50}$ with $E247^{6.30}$ (PDB ID: 1f88 [16]). In contrast, the inactive β2AR forms an interaction between $R131^{3.50}$ and $S134^{IL2}$ and $T68^{2.39}$ (Figure 4.3.1C). In β2AR, the TM6 is locked via $H269^{6.31}$ to $Y141^{IL2}$ at IL2 but not directly to TM3 [18]. Hence, MC4R forms a TM3-TM6 lock that resembles the rhodopsin ionic lock without an ionic interaction (Section 1.1.2).
In the antagonized MC4R structure, this lock is additionally supported by $T150^{3.53}$ located at the adjacent turn to $R147^{3.50}$. $T150^{3.53}$ forms hydrogen bonds to $R147^{3.50}$ and $N240^{6.30}$ (Figure 4.3.1A) and is therefore involved in stabilizing the antagonized state and the active state by forming a hydrogen bond to $H373^{\alpha5.19}$ at the $G\alpha_s$-α5-helix (Figures 4.3.1D- E). The pathogenic MC4R mutation $T150I^{222}$ is connected with severe obesity, and the corresponding MC4R-T150I variant induces 35% of wild-type cAMP signaling upon α-MSH stimulation (Table 3.5.1). The reduction of cell surface expression of T150A to 58% (Table 3.5.1) indicates that the structural integrity is reduced, but the variant is still expressed. Therefore, the total loss of function of T150A upon α-MSH stimulation (Figure 3.5.3.1E) is most likely driven by an impaired MC4R-G_s binding and not the loss of structural integrity due to the weakened TM3-TM6 lock. Furthermore, this variant has no impaired NDP-α-MSH binding in radioligand studies [221]. Hence, the presented fraction of the receptor is still functional and properly folded.

On the contrary, the T150A mutant does not display a significant drop in efficacy for $G_{q/11}$ signaling upon stimulation with NDP-α-MSH and setmelanotide (Figures 3.5.1D-E). Hence, $T150^{3.53}$ is not essential for the coupling of $G_{q/11}$. The deviating pathway responses

imply that G_s and $G_{q/11}$ have a deviating binding mode at the intracellular end of TM3, and $T150^{3.53}$ is one essential site defining G-protein coupling specificity. Intriguingly, the $G_{q/11}$ potency of setmelanotide stimulating the T150A mutant is reduced by 24-fold compared to NDP-α-MSH (Figures 3.5.1D-E and Table 3.5.3). A shift in potency correlates with changes in ligand affinity and hence a single-side chain substitution at the G-protein interface impacts the strength of the receptor-ligand interaction (Section 1.3.4). This shift in potency implies that that ligand-binding pocket and the G-protein interface are allosterically coupled, and ligands could convey their action via TM3 towards the G-protein interface at and nearby $T150^{3.53}$.

4.3.2 Relevance of IL2 for G-protein activation

MC4R's $_{3/10}$-helix at IL2 is also formed in the antagonized MC4R structure. $D146^{3.49}$ is in hydrogen bond distance to $Y157^{IL2}$ in the active and the antagonized structures. In contrast, rhodopsin and β2AR display here large conformational rearrangements. $Y141^{IL2}$ in β2AR rotates away from TM6 and towards $R131^{3.50}$ at TM3 (Figures 4.3.1C and F), and the helical IL2 conformation in β2AR is formed upon G-protein binding. If the antagonized MC4R structure represents an entirely inactive state cannot be stated since the MC4R construct was intensely modified and SHU9119 is not an inverse agonist. Hence, a feature like the preformed IL2-DRY interaction could indicate a transitory state. An inactive structure with the endogenous inverse agonist AgRP could clarify that.

In 2020 the Tao lab published extensive alanine exchange experiments at the *DRY* motif and IL2 interface, they assessed cAMP signaling upon NDP-α-MSH stimulation and direct NDP-α-MSH ligand binding [221]. R147A, L155A, Y157A, and H158A displayed increased basal signaling, but no changes in NDP-α-MSH stimulated cAMP accumulation. A significant reduction in cAMP accumulation upon NDP-α-MSH stimulation was shown solely for MC4R variant T150A (30-fold reduced potency), but not for the variants of the *DRY* motif, L155A, and Y157A. Total ligand binding was significantly reduced for Q156A, Y157A, M161A, and T162A without impairing cAMP accumulation [221]. The loss of NDP-α-MSH binding for the MC4R-Q156A and T162A variant is intriguing because the NDP-α-MSH-MC4R-G_s complex indicates hydrogen bonds between those residues and $G\alpha_s$ ($Q156^{IL2}$-$V203^{\beta2\beta3}$ and $T162^{IL2}$-$Q35^{\alpha N}$) (Figure 3.4.3.1A). These bonds are not

apparent in the setmelanotide-MC4R-G_s complex (Figure 3.4.3.1B). The structural changes are subtle differences and require further studies, such as ligand binding studies for NDP-α-MSH and setmelanotide. The TAMRA-NDP-α-MSH nanoLuc based ligand binding assay or radioligand assays using 125I-NDP-α-MSH are not suitable to dissect a potential deviating impact on ligand binding. Since the competition of the radioactive or fluorescent-labeled NDP-α-MSH binding cannot be adequately compared if NDP-α-MSH binding is impaired. A TAMRA-setmelanotide nLuc based assay could shed light on this question, but fluorescent labeling of setmelanotide resulted thus far in ligand inactivation.

The dual rotamer conformation of the central IL2 residue, $H158^{IL2}$, in the NDP-α-MSH complex structure (Figure 4.3.1D) contrasts with only one rotamer conformation in the setmelanotide complex structure. The MC4R-H158A mutant displays increased basal signaling but does not display a significant impact on NDP-α-MSH or setmelanotide induced G_s signaling, except a subtle increase in efficacy for all three tested agonists (Figure 3.5.3.1).
In contrast to the G_s signaling, the maximum of $G_{q/11}$ signaling is reduced to ~50% for the H158A variant in the case of NDP-α-MSH and setmelanotide (Figure 3.5.1D-E). Thus, this residue, along with $T150^{3.53}$, is a second identified sensitive receptor site for a shifted G-protein signaling profile suggesting differences in specific G-protein binding. Finally, our data strongly support an essential role of this IL2 receptor residue for $G_{q/11}$ activation, but barely for G_s activation.

4.3.3 A unique receptor-G-protein positioning

The relative positioning of the G_s-protein towards the receptor varies among β2AR-G_s (PDB ID: 3sn6) and both MC4R-G_s complexes (Figures 4.3.3.1A and C). The carboxy-terminus, C-cap, of the α5-helix superimposes well but turn by turn, going towards the amino-terminal end of the α5-helix, their positioning deviates more prominently. The G-protein appears to behave like a rigid body, and a deviating α5-helix positioning is translated throughout the G-protein. This translational rotation at the membrane plane is most prominently apparent by the 9 Å αN-helix shift (measured from the Cα of MC4R-A18 and β2AR-N14) (Figure 4.3.3.1A-B, and 4.3.3.2B).

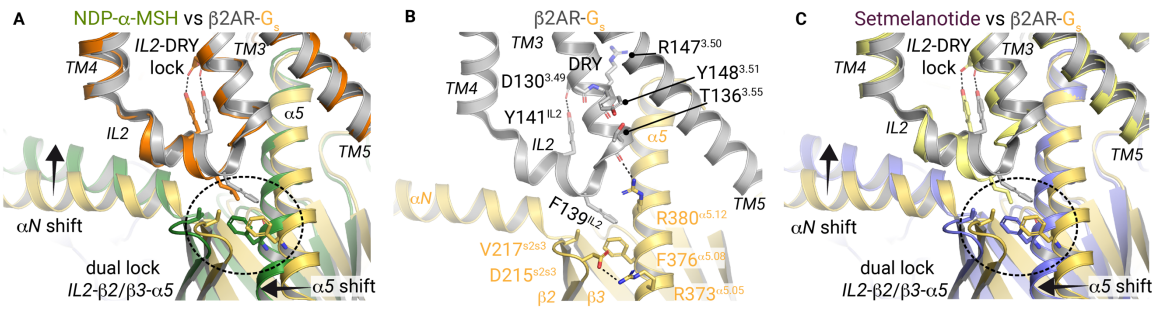

Figure 4.3.3.1 ***G_s protein binding mode comparison of MC4R and β2AR.*** (**A** and **C**) Superposition of β2AR-G_s (PDB ID: 3sn6 [20]) with NDP-α-MSH-MC4R-G_s (**A**) and setmelanotide-MC4R-G_s (**C**). Interactions at the β2AR-G_s interface are shown in detail in (**B**). Black arrows indicate the relative positioning of $G\alpha_s$ in MC4R-G_s versus β2AR-G_s structures. Figure 3.4.3 displays a detailed figure of the MC4R-G_s interface for both agonists.

A key site for G_s protein binding to β2AR is $F139^{IL2}$ [223], which is located at IL2 adjacent to TM3 and is part of a hydrophobic pocket formed by $G\alpha_s$ $H41^{\beta1.02}$, at the beginning of the β1 strand, $V217^{s2s3}$ at the β3 strand, and $F376^{\alpha5.08}$ at the α5-helix (Figure 4.3.1F and Figure 4.3.3.1B). The β2AR-F139A mutant displays strongly impaired G_s activation [224] and is unable to trigger GDP release [225]. The corresponding residue in MC4R is $L155^{IL2}$ which engages in the same hydrophobic contacts as β2AR-$F139^{IL2}$ but has a smaller hydrophilic side chain. The MC4R-L155A variant displays WT cAMP accumulation upon NDP-α-MSH stimulation [221]. MC4R does not depend on $L155^{IL2}$ for GDP release in G_s, and thus, MC4R has a different G-protein binding mode as β2AR.

The superposition of β2AR and MC4R at IL2 (Figure 4.3.3.1) suggests that the varying G-protein positioning is based on the size of the MC4R $L155^{IL2}$ side chain. Therefore, the α5-helix can pack more tightly at IL2, and the G-protein is slightly rotated compared to β2AR. The superposition of all available class A GPCR-G_s complexes stabilized by Nb35 displays two receptor G-protein complexes, dopamine 1 receptor (D1R)-G_s and GPR52-G_s that have an equal G-protein angle towards the receptor (Figures 4.3.3.2A and C). GPR52 has at the MC4R-$L155^{IL2}$ position $L147^{IL2}$, and D1R has $F129^{IL2}$ (Figure 4.3.3.2D). Hence, G-protein binding cannot be delineated to a single position at the primary sequence level. The IL2 of D1R is slightly positioned away from the α5-helix, compensating for the bigger side-chain size of $F129^{IL2}$. Ultimately, the question arises what functional state GPCR-G-protein complexes in the nucleotide-free G-protein state resemble.

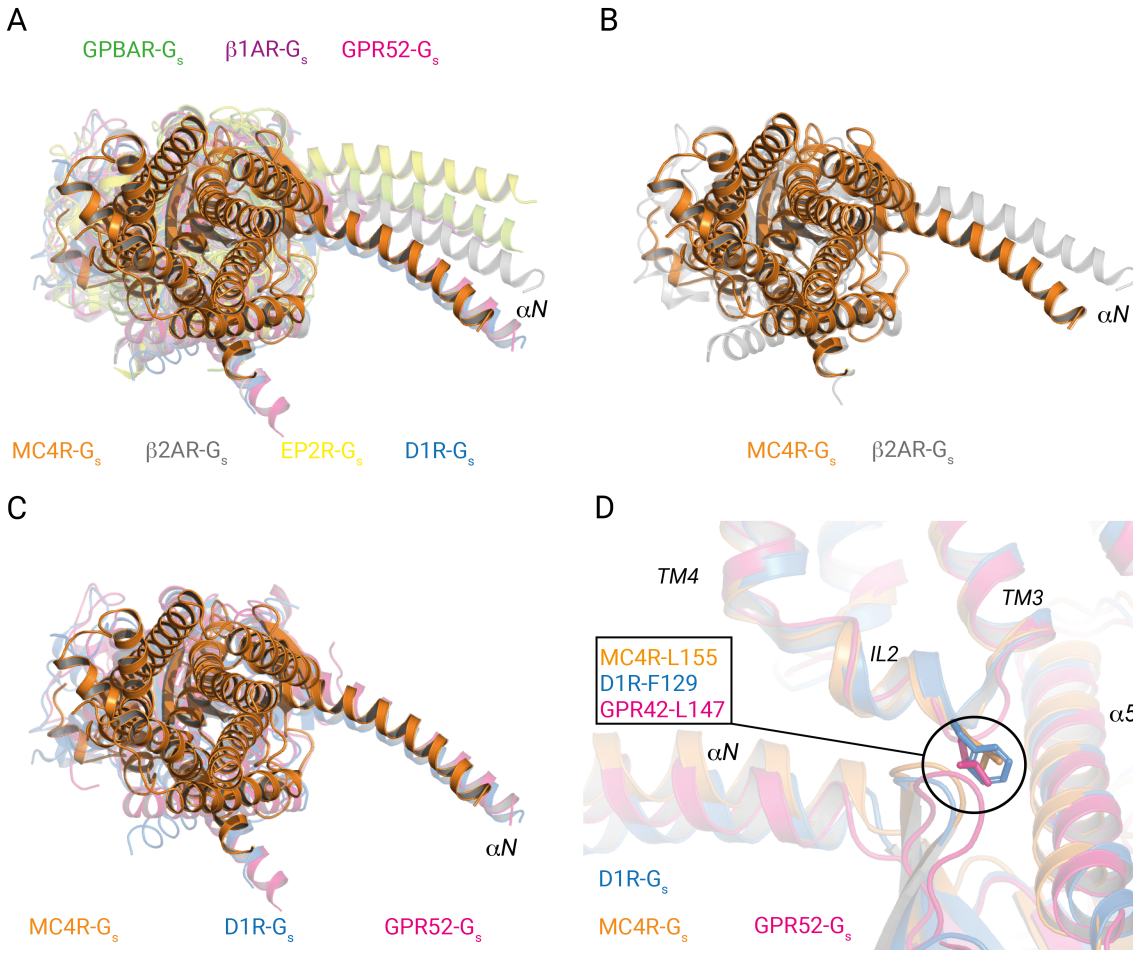

Figure 4.3.3.2 ***Superposition of all available GPCR- G_s -Nb35 complexes*** [footnote 8]. The first PDB deposited structure was selected to decrease the visual complexity of multiple structures with deviating ligands. Structures were aligned on MC4R, and solely the receptor and the Gαs subunit are displayed for clarity. Receptor-G-protein complexes are colored as follows: MC4R-G_s in orange, b2AR-G_s (PDB ID: 3sn6) in gray; EP2-G_s (PDB ID:7cx3) in yellow, D1R-G_s (PDB ID: 7Jjoz) in blue, GPBAR-G_s (PDB ID 7cfm) in light green, β1AR-G_s (PDB ID:7jjo) in purple, and GPR52-G_s (PDB ID: 6li3) in magenta. The superposition MC4R-G_s with of (**A**) GPBAR-G_s, β1AR-G_s, GPR52-G_s, β2AR-G_s, EP2R-G_s and D1R-G_s highlights that the relative positioning of the receptor to G_s deviates between complex structures. (**B**) displays the superposition of MC4R-G_s with β2AR-G_s and (**C**) of receptors that share an equal receptor-G-protein rotation angle with MC4R. (**D**) Enlarged display of the IL2-G_s interface focuses on residue at the MC4R-L155 position. Displayed are complex structures from (**C**).

[8] www.gpcrdb.com/structures; 27.07.2021

4.3.4 The physiological relevance of GPCR-G-protein complexes in the nucleotide-free state

The surge in published receptor G-protein complexes since 2018 increased the knowledge of agonist action, receptor conformations typical for the active state, and the overall receptor-G-protein interface. The receptor-G-protein interface and the G-protein itself do not display significant variations among the G-protein subtype-specific complex structures. The α5-helix is the main conformational link between the nucleotide-binding pocket and the GPCR in the known class A and class B complex structures [212].

Therefore, the complex is stabilized in this state by two factors: the G-protein stabilizing entity (*e.g.*, scFv16 for G_i or Nb35 for G_s) and Gα in the nucleotide-free state. The recently published structures of MC1R-G_s-iN [footnote 9] complexes stabilized by Nb35 or scFv16 [215] displayed no significant changes in the receptor-G-protein interface except a transposition of the αN-helix by 5.7 Å along the membrane plane due to the scFv16 binding. This transposition is comparable with MC4R-G_s compared with β2AR-G_s (Figure 4.3.3.2B). The here described setmelanotide-MC4R-G_s complex stabilized by Nb35 compared with the recently published complex stabilized by scFv16 displays at the αN-helix a shift of more than 8 Å (PDB ID: 7aue [92]). Therefore, the choice of the G-protein stabilizing entity affects the overall positioning of the G-protein towards the receptor. The stabilization of the G_s complex with Nb35 does not impair the receptor G-protein interface at the αN-helix.

Removing GDP during the complex formation step by adding nucleotide-hydrolase apyrase is crucial for stable complex formation. Due to high intracellular GTP levels, the thereby stabilized low energy state does not exist in cells for longer than a minuscule fraction of a second. Hence, the short existence of the nucleotide-free receptor G-protein complex state implies that it is not the significant physiological state for G-protein coupling. β2AR G_s mediated activation in living cells occurs on a time span of 2 - 3s [226], and the relevant conformational changes allowing G_s to bind are highly transient on a time scale << 100 ms [227]. The observation that the rate-limiting conformational changes for receptor-G-protein binding precede GTP binding and occur in such a short time span has been shown as well for rhodopsin-G_t by FRET studies [228]. These studies imply that the

[9] Gαs-iN is a $G\alpha_s$ with the aminoterminal 15 residues exchanged for the aminoterminal 15 residues of $G\alpha_i$ and therefore having the epitope for scFv16 [45]

receptor adopts a different conformation during G-protein binding in the GDP bound state as observed in the nucleotide-free state. Such a state has been recently proposed by a crystal structure of β2AR with a tethered high-affinity peptide [212]. In addition, the structures of the M1R-G_{11} and the M2R-G_o complexes display varying receptor-G-protein interactions, but a functional relevance in G-protein specificity could not be assigned to them [45]. The recently published structures of cholecystokinin 1 (CCK1) receptor bound to heterotrimeric miniG_s and miniG_q proteins do increase the understanding of the mechanism of G-protein selectivity, but sites that are responsible for G-protein selectivity could not be identified [229].

Nucleotide-free receptor G-protein complexes have tremendous relevance in understanding the pharmacology of agonists, the activation of GPCRs, and the design of novel therapeutics. However, nucleotide-free complexes do not harbor all information necessary to answer all questions regarding G-protein selectivity, G-protein activation, and the receptor conformation that makes the initial contact to the G-protein.

The limited physiological relevance of the nucleotide-free state for the G-protein recognition challenges the observed differences at the receptor-G_s interface between the $G_{q/11}$ favoring setmelanotide and the G_s favoring NDP-α-MSH.

Nevertheless, it is striking that observed differences at TM3 and IL2 upon agonist binding are relevant for G_s and especially $G_{q/11}$ signaling. Since the G-protein binding stabilizes a low energy receptor state and the here observed differences around $H158^{IL2}$ could indicate that in transient states preceding the nucleotide release that TM3 adjusts IL2 and thereby plays a role in G-protein recognition. Additional studies are required that define MC4R's cognate $G_{q/11}$ coupling partner to support this hypothesis. The *in vitro* complex formation experiments with heterotrimeric G_q-iN and engineered miniG_q proteins did not yield in biochemical evidence of G_q binding to MC4R (data not shown)

5 Summary and outlook

The investigation of the active-state MC4R-G_s-protein complex structures required the establishment of three experimental prerequisites.

(1) This doctorate's primary challenge was establishing GPCR expression and purification methods needed for the structural investigation of peptidic GPCRs. In Mai 2018, I was invited to the lab of Brian Kobilka at Stanford University to identify several relevant variables in the production of GPCRs and their purification. Therefore, the insect cell culture was adapted to and optimized for the expression of MC4R, including novel cell lines, the in-house production of a recombinant bacmid, splitting cycles, and the definition of parameters relevant for reproducible virus generation. Those initial implementations enabled the production of a multitude of different MC4R constructs. This doctorate intended to solve an inactive MC4R structure by X-ray crystallography during the first two years. These efforts resulted in the cloning, virus generation, and expression of 55 constructs; however, the aggregation tendency of the receptor reduced the amount of purified functional protein. After a second visit to the Kobilka lab at Tsinghua University in Beijing, our strategy was pivoted to focus on the active-state MC4R structure using cryo-EM. Between January 2019 and August 2019, I established the G-protein bacmid purification, defined the best expression set-up, and adapted the purification protocol for heterotrimeric G_s, G_i, and G_q. Another detrimental step for the MC4R purification was using the M1-FLAG antibody column. The rapid and highly pure purification via this M1-FLAG antibody had an unwanted positive side effect, adding 2.5 mM $CaCl_2$ to most purification steps. This coincidence enabled the first successful complexation experiments since the initial purifications were done using NiNTA, and the used buffers did not contain $CaCl_2$ in all experimental steps.

(2) The establishment of a ligand-binding assay for MC4R was crucial for identifying potent agonists and assessing the biological activity of MC4R constructs. Here, the adaption of the nanoLuc based protocol by Leigh Stoddard *et al.* [143] was successful. The addition of the nLuc to the amino-terminal side of MC4R allowed searching for suitable high-affinity ligands for structural studies. In hindsight, this method was detrimental for shifting the strategy towards an active-state structure.

(3) The third objective was the large-scale production of a stable agonist-MC4R-G_s construct. Here, the fluorescent nature of MC4R-eGFP was a prerequisite for measuring the complex stability using F-SEC. The removal of the eGFP was guided by initial cryo-EM experiments, and the cleavage by H3C protease ultimately resulted in micrographs with homogeneous particles with random orientations. The cryo-EM data processing using cryoSPARC and RELION yielded in the cryo-EM maps of NDP-α-MSH and setmelanotide-MC4R-G_s with resolutions of 2.9 Å and 2.6 Å. The resulting models are the first GPCR-G-protein complex structures depositions coming from Germany.

The determination of the complex structures in combination with signaling studies allowed to study the three objectives of this study: To (i) characterize the peptide-binding modes of NDP-α-MSH and setmelanotide, (ii) unravel the mechanism of MC4R activation, and (iii) characterize the molecular determinants of G_s-protein coupling to MC4R.

(i) Agonist-binding modes

The agonist-MC4R-G_s complex structures reveal subtle modulations of the agonist-Ca^{2+}-receptor interfaces. The structures deviate in specific TM3 positions, and setmelanotide displays deviations in the binding mode to MC4R and Ca^{2+} compared to NDP-α-MSH's binding mode. The most notable is the direct setmelanotide interaction of D-R1^{-3} to MC4R-D122$^{3.25}$ at TM3, which has no counterpart in the NDP-α-MSH binding mode.

Intriguingly, LY2112688, an MC4R agonist that shares setmelanotide sequence with two deviations, is G_s-biased and induces significant cardiovascular adverse effects in rodents and obese rhesus macaques [230] compared to setmelanotide. One of the substitutions potentially impairs the ligand's ability to address D122$^{3.25}$. This comparison provides a first structural indication for the enhanced $G_{q/11}$ signaling of setmelanotide and links our structural observations with setmelanotide successful application in appetite regulation.

Comparing our NDP-α-MSH-MC4R structure with the recently preprinted report of an NDP-α-MSH-MC1R-G_s structure by the Eric Xu lab [215] may provide a sound basis for developing more subtype-selective drugs, as various positions of the LBP differ between the MCR subtypes. Furthermore, the structures could be used in computational docking

approaches that filter for MC4R or MC1R subtype selectivity. Since the non-subtype selective nature of MC4R drugs binding to MC1R results in adverse effects. Moreover, the active MC4R structures illuminated the subtype-specific nature of SHU9119, and novel designed MC4R antagonists could target L133$^{3.36}$.

(ii) MC4R activation

The active state MC4R-G_s complex structures compared to the antagonist bound MC4R structure allowed identifying active state properties relevant to signal transduction. Among these changes is a sizeable outward movement of TM6. This movement is pivoted at residue W258$^{6.48}$, which shifts towards the intracellular side during activation and is accompanied by I291$^{7.42}$ at TM7. Here, slight conformational alterations are related to local shifts at the *N(D)P$^{7.50}$xxY* motif. Mutagenesis signaling data of the G_s and $G_{q/11}$ pathway was used to assess their significance for G-protein coupling. The data allows assigning the agonist-dependent activation trigger at TM6 starting from the ligand-binding region of H264$^{6.54}$-F261$^{6.51}$ to the helix-tilting area of the *CWxP$^{6.50}$* motif. The conserved *DR$^{3.50}$Y* motif in TM3 and the adjacent IL2 constitute the central interfaces to G_s. Furthermore, the active-state structures clearly demonstrate that SHU9119 acts as an antagonist because it blocks the TM6 movement by locking W258$^{6.48}$ via L133$^{3.25}$.

(iii) MC4R-G_s-protein coupling

This study used two ligands with subtle deviating pharmacological profiles to identify potential sites of agonist-induced flexibility at the receptor-G_s-protein interface. On the one hand, NDP-α-MSH is characterized by a higher G_s-signaling efficacy; on the other hand, the FDA-approved anti-obesity drug setmelanotide with a higher $G_{q/11}$ signaling efficacy. Differences in the G_s-protein binding interface, particularly in IL2, in both agonist-bound MC4R structures provide initial evidence for the molecular underpinnings of the altered G-protein selectivity of NDP-α-MSH and setmelanotide. Here, T150$^{3.53}$ at TM3 is essential for G_s but not for $G_{q/11}$ signaling and displays ligand-dependent changes in potency for

$G_{q/11}$ signaling. In addition, the signaling data of $T150^{3.53}$ implies that TM3 in MC4R is a marker for G-protein specificity.

Consequently, the ligand-binding pocket and the G-protein interface are loosely allosterically connected, potentially via TM3. The second identified marker for G-protein specificity is $H158^{IL2}$ at IL2 relevant for $G_{q/11}$ signaling but not for G_s signaling. These two identified potential markers for G-protein coupling specificity suggest that the orientation of the α5-helix from $G_{q/11}$ into the G-protein binding cavity of MC4R must have a different binding interface compared with the G_s protein.

A variety of follow-up studies can be based on the active MC4R structures. For example, a setmelanotide-MC4R-$G_{q/11}$ structure has the potential to unravel sites that define $G_{q/11}$ specificity. However, such work requires extensive signaling studies that decipher which interactions are unique to the nucleotide-free state and which are involved in the transient state of G-protein recognition.

Furthermore, understanding the agonist-binding modes allows for rational design of peptides, that bind. focusing on MC4R subtype selectivity, ligands that bind specifically to MC4R and not on the other MCR subtypes, and $G_{q/11}$ bias. The two deviating residues in primary sequence of LY2112688 and setmelanotide are an ideal starting point for developing $G_{q/11}$ (setmelanotide) or G_s (LY2112688) biased drugs. Structures of MC4R bound to the endogenous agonists α-MSH and β-MSH could unravel if setmelanotide interaction via $R1^{-3}$ with $D122^{3.25}$ originates from β-MSH and if β-MSH's subtle $G_{q/11}$ bias originates here.

An apo-MC4R-G_s structure could further illuminate the role of Ca^{2+} in MC4R activation and basal activity. However, the nucleotide-free G_s' allosteric stabilization of the ligand-binding pocket could be an artificial state that is not physiologically relevant. Ultimately, the next intriguing question in the GPCR field is the investigation of transient receptor states. These could be investigated by combining spectroscopic methods, such as electron paramagnetic resonance (EPR), with next-generation structural approaches like mass spectrometry, free-electron laser serial crystallography, time-resolved single-particle EM, or cryogenic-electron tomography in living cells.

Bibliography

1. Fredriksson, R., Lagerström, M. C., Lundin, L.-G. & Schiöth, H. B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* **63**, 1256 LP – 1272 (2003).
2. Sanchez-Reyes, O. B. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).
3. Hofmann, K. P. *et al.* A G protein-coupled receptor at work : the rhodopsin model. (2009) doi:10.1016/j.tibs.2009.07.005.
4. Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B. & Gloriam, D. E. Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).
5. Manglik, A., Kruse, A. & Kruse, A. C. Structural basis for G protein-coupled receptor activation. (2017) doi:10.1021/acs.biochem.7b00747.
6. Lohse, M. J. *et al.* Receptor-specific desensitization with purified proteins. Kinase dependence and receptor specificity of beta-arrestin and arrestin in the beta 2-adrenergic receptor and rhodopsin systems. *J. Biol. Chem.* **267**, 8558–8564 (1992).
7. DeWire, S. M., Ahn, S., Lefkowitz, R. J. & Shenoy, S. K. β-Arrestins and cell signaling. *Annu. Rev. Physiol.* **69**, 483–510 (2007).
8. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. **455**, 2–8 (2008).
9. Kang, Y. *et al.* Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **523**, (2015).
10. Chen, Q. *et al.* Structures of rhodopsin in complex with G-protein-coupled receptor kinase 1. *Nat. 2021* **595**, 1–6 (2021).
11. Chambers, A. C. *et al.* Overview of the Baculovirus Expression System. *Curr. Protoc. Protein Sci.* 5.4.1-5.4.6 (2018) doi:10.1002/cpps.47.
12. Chae, P. S. *et al.* Glucose-neopentyl glycol (GNG) amphiphiles for membrane protein study. *Chem. Commun.* **49**, 2287–2289 (2013).
13. Kobilka, B. K. Amino and carboxyl terminal modifications to facilitate the production and purification of a G protein-coupled receptor. *Analytical Biochemistry* vol. 231 269–271 (1995).
14. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009).
15. Chun, E. *et al.* Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**, 967–976 (2012).
16. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science (80-.).* **289**, 739–745 (2000).

17. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. **454**, 183–188 (2008).

18. Cherezov, V. *et al.* High-Resolution Crystal Structure of an Engineered Human b 2 -Adrenergic. **318**, 1258–1266 (2007).

19. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* **469**, 175–181 (2011).

20. Rasmussen, S. G. F. *et al.* Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–557 (2011).

21. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. in *Receptor Molecular Biology* (ed. Sealfon, S. C. B. T.-M. in N.) vol. 25 366–428 (Academic Press, 1995).

22. Hurtley, S. M. Continuing the resolution revolution. *Science (80-.).* **360**, 282 (2014).

23. García-Nafría, J. & Tate, C. G. Cryo-Electron Microscopy: Moving Beyond X-Ray Crystal Structures for Drug Receptors and Drug Development. *Annu. Rev. Pharmacol. Toxicol. Annu. Rev. Pharmacol. Toxicol* (2019) doi:10.1146/annurev-pharmtox-010919.

24. Kato, H. E. *et al.* Conformational transitions of a neurotensin receptor 1–Gi1 complex. *Nature* (2019) doi:10.1038/s41586-019-1337-6.

25. Deupi, X. & Kobilka, B. K. Energy landscapes as a tool to integrate GPCR structure, dynamics, and function. *Physiology* **25**, 293–303 (2010).

26. Peleg, G., Ghanouni, P., Kobilka, B. K. & Zare, R. N. Single-molecule spectroscopy of β2 adrenergic receptor: Observation of conformational substates in a membrane protein. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8469–8474 (2001).

27. Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* (2016) doi:10.1038/nature18636.

28. Manglik, A. *et al.* Structural insights into the dynamic process of beta 2-adrenergic receptor signaling. *Cell* **161**, 1101–1111 (2015).

29. Violin, J. D., Crombie, A. L., Soergel, D. G. & Lark, M. W. Biased ligands at G-protein-coupled receptors: Promise and progress. *Trends Pharmacol. Sci.* **35**, 308–316 (2014).

30. Dixon, R. a. F. *et al.* Cloning of the gene and cDNA for mammalian β-adrenergic receptor and homology with rhodopsin. *Nature* **321**, 75–79 (1986).

31. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist-b2 adrenoceptor complex. *Nature* (2011) doi:10.1038/nature09665.

32. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist-β2 adrenoceptor complex. *Nature* **469**, 236–240 (2011).

33. Choe, H. W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471**, 651–655 (2011).

34. Che, T. *et al.* Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell* **172**, 55-67.e15 (2018).

35. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature* **469**, 175–180 (2011).

36. Carpenter, B., Nehmé, R., Warne, T., Leslie, A. G. W. & Tate, C. G. Structure of the adenosine A2A receptor bound to an engineered G protein. *Nature* **536**, 104–107 (2016).

37. Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535**, 448–452 (2016).

38. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).

39. Schlinkmann, K. M. *et al.* Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9810–9815 (2012).

40. Shi, L. *et al.* β2 adrenergic receptor activation: Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **277**, 40989–40996 (2002).

41. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).

42. Tehan, B. G., Bortolato, A., Blaney, F. E., Weir, M. P. & Mason, J. S. Unifying Family A GPCR Theories of Activation. *Pharmacol. Ther.* **143**, 51–60 (2014).

43. Saleh, N. *et al.* Binding, thermodynamics, and selectivity of a non-peptide antagonist to the melanocortin-4 receptor. *Front. Pharmacol.* **9**, 1–10 (2018).

44. Heyder, N. *et al.* Signal Transduction and Pathogenic Modifications at the Melanocortin-4 Receptor: A Structural Perspective. *Front. Endocrinol. (Lausanne).* **10**, 1–18 (2019).

45. Maeda, S., Qu, Q., Robertson, M. J., Skiniotis, G. & Kobilka, B. K. Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes. *Science (80-.).* **364**, 552–557 (2019).

46. Mirzadegan, T., Benkö, G., Filipek, S. & Palczewski, K. Sequence Analyses of G-Protein-Coupled Receptors: Similarities to Rhodopsin. *Biochemistry* **42**, 2759–2767 (2003).

47. Inoue, A. *et al.* Illuminating G-Protein-Coupling Selectivity of GPCRs Resource Illuminating G-Protein-Coupling Selectivity of GPCRs. *Cell* **177**, 1933-1947.e25 (2019).

48. Harding, S. D. *et al.* The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: Updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res.* **46**, D1091–D1106 (2018).

49. Flock, T. *et al.* Universal allosteric mechanism for Gα activation by GPCRs. *Nature* **524**, 173–179 (2015).

50. Wall, M. A. *et al.* The structure of the G protein heterotrimer Giα1β1γ2. *Cell* **83**, 1047–1058 (1995).

51. Wettschureck, N. & Offermanns, S. Mammalian G proteins and their cell type specific functions. *Physiol. Rev.* **85**, 1159–1204 (2005).

52. Sgourakis, N. G., Bagos, P. G. & Hamodrakas, S. J. Prediction of the coupling specificity of GPCRs to four families of G-proteins using hidden Markov models and artificial neural networks. *Bioinformatics* **21**, 4101–4106 (2005).

53. Yabuki, Y., Muramatsu, T., Hirokawa, T., Mukai, H. & Suwa, M. GRIFFIN: A system for predicting GPCR-G-protein coupling selectivity using a support vector machine and a hidden Markov model. *Nucleic Acids Res.* **33**, 148–153 (2005).

54. Bruce R. Conklin, Zvi Farfelt, Kevin D. Lustig, D. J. & H. R. B. Substitution of three amino acids switches receptor specificity of Gqa to that of Gia. *Nature* **255**, 243–244 (1975).

55. Rose, A. S. *et al.* Position of transmembrane helix 6 determines receptor G protein coupling specificity. *J. Am. Chem. Soc.* **136**, 11244–11247 (2014).

56. Kim, K. *et al.* Structure of a Hallucinogen-Activated Gq-Coupled 5-HT2A Serotonin Receptor. *Cell* **182**, 1574-1588.e19 (2020).

57. Maeda, S., Qu, Q., Robertson, M. J., Skiniotis, G. & Kobilka, B. K. Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes. *Science (80-.).* **364**, 552–557 (2019).

58. Draper-Joyce, C. J. *et al.* Structure of the adenosine-bound human adenosine A_1 receptor–G_i complex. *Nature* **558**, 559–63 (2018).

59. Guettier, J. *et al.* A chemical-genetic approach to study G protein regulation of ␤ cell function in vivo. **106**, 19197–19202 (2009).

60. Koehl, A. *et al.* Structure of the μ-opioid receptor-G i protein complex. *Nature* **558**, 547–552 (2018).

61. Kang, Y. *et al.* Cryo-EM structure of human rhodopsin bound to an inhibitory G protein. *Nature* **558**, 553–558 (2018).

62. Hauner, H. *Obesity and Diabetes*. *Textbook of Diabetes: Fourth Edition* (2020). doi:10.1002/9781444324808.ch14.

63. Myers A. & Rosen J.C. Obesity stigmatization and coping: Relation to mental health symptoms, body image, and self-esteem. *Int. J. Obes.* **23**, 221–230 (1999).

64. Stunkard, A. J., Harris, J. R., Pedersen, N. L. & McClearn, G. E. The body-mass index of twins who have been reared apart. *N. Engl. J. Med.* **322**, 1483–1487 (1990).

65. Stunkard, A. J., Foch, T. T. & Hrubec, Z. A twin study of human obesity. *JAMA* **256**, 51–54 (1986).

66. Yeo, G. S. H. *et al.* The melanocortin pathway and energy homeostasis: From discovery to obesity therapy. *Mol. Metab.* 101206 (2021) doi:10.1016/j.molmet.2021.101206.

67. Lu, D. *et al.* Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature* **371**, 799–802 (1994).

68. Ollmann, M. M. Antagonism of Central Melanocortin Receptors in Vitro and in Vivo by Agouti-Related Protein. *Science (80-.).* **278**, 135–138 (1997).

69. FERRARI, W. Behavioural changes in animals after intracisternal injection with adrenocorticotrophic hormone and melanocyte-stimulating hormone. *Nature* **181**, 925–926 (1958).

70. WOODBURY, D. M. Effect of adrenocortical steroids and adrenocorticotrophic hormone on electroshock seizure threshold. *J. Pharmacol. Exp. Ther.* **105**, 27–36 (1952).

71. Hadley, M. E., Hruby, V. J. & Bower, S. A. Cellular mechanisms controlling melanophore stimulating hormone (MSH) release. *Gen. Comp. Endocrinol.* **26**, 24–35 (1975).

72. Gee, C. E., Chen, C. L., Roberts, J. L., Thompson, R. & Watson, S. J. Identification of proopiomelanocortin neurones in rat hypothalamus by in situ cDNA-mRNA hybridization. *Nature* **306**, 374–376 (1983).

73. Sawyer, T. K. *et al.* 4-Norleucine, 7-D-phenylalanine-a-melanocyte-stimulating hormone: A highly potent a-melanotropin with ultralong biological activity. **77**, 5754–5758 (1980).

74. TOMI K. SAWYER, VICTOR J. HRUBY, PAUL S. DARMAN, A. M. E. H. [Half-Cystine4, Half- Cystine10]-α-melanocyte stimulating hormone: A cyclic α-melanotropin exhibiting superagonist biological activity. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 1751–1755 (1982).

75. Tatro, J. B. Melanotropin receptors in the brain are differentially distributed and recognize both corticotropin and alpha-melanocyte stimulating hormone. *Brain Res.* **536**, 124–132 (1990).

76. Mountjoy, K. G., Robbins, L. S., Mortrud, M. T. & Cone, R. D. The cloning of a family of genes that encode the melanocortin receptors. *Science (80-.).* **257**, 1248–1251 (1992).

77. Gantz, I. *et al.* Molecular cloning, expression, and gene localization of a fourth melanocortin receptor. *J. Biol. Chem.* **268**, 15174–15179 (1993).

78. Gantz, I. *et al.* Molecular cloning, expression, and characterization of a fifth melanocortin receptor. *Biochem. Biophys. Res. Commun.* **200**, 1214–1220 (1994).

79. Poggioli, R., Vergoni, A. V. & Bertolini, A. ACTH-(1-24) and α-MSH antagonize feeding behavior stimulated by kappa opiate agonists. *Peptides* **7**, 843–848 (1986).

80. Huszar, D. *et al.* Targeted Disruption of the Melanocortin-4 Receptor Results in Obesity in Mice. *Cell* **88**, 131–141 (1997).

81. Yeo, G. S. *et al.* A frameshift mutation in MC4R associated with dominantly inherited human obesity. *Nature genetics* vol. 20 111–112 (1998).

82. Vaisse, C., Clement, K., Guy-grand, B. & Froguel, P. A frameshift mutation in human MC4R is associated with a dominant form of obesity. **20**, 113–114 (1998).

83. Vaisse, C. *et al.* Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J. Clin. Invest.* **106**, 253–262 (2000).

84. Yeo, G. S. H. *et al.* Mutations in the human melanocortin-4 receptor gene associated with severe familial obesity disrupts receptor function through multiple molecular mechanisms. *Hum. Mol. Genet.* **12**, 561–574 (2003).

85. O'Rahilly, S., Sadaf Farooqi, I., Yeo, G. S. H. & Challis, B. G. Minireview: Human obesity - Lessons from monogenic disorders. *Endocrinology* **144**, 3757–3764 (2003).

86. Panaro, B. L. & Cone, R. D. Melanocortin-4 receptor mutations paradoxically reduce preference for palatable foods. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7050–7055 (2013).

87. Litt, M. J. *et al.* Loss of the melanocortin-4 receptor in mice causes dilated cardiomyopathy. *Elife* **6**, 1–21 (2017).

88. Fan, W., Boston, B. A., Kesterson, R. A., Hruby, V. J. & Cone, R. D. Role of melanocortlnergic neurons in feeding and the agouti obesity syndrome. *Nature* vol. 385 165–168 (1997).

89. Nijenhuis, W. A. J., Oosterom, J. & Adan, R. A. H. AgRP (83 – 132) Acts as an Inverse Agonist on the Human- Melanocortin-4 Receptor. 164–171 (2016).

90. Chai, B. X. *et al.* Inverse agonist activity of agouti and agouti-related protein. *Peptides* **24**, 603–609 (2003).

91. Yu, J. *et al.* Determination of the melanocortin-4 receptor structure identifies Ca 2+ as a cofactor for ligand binding. **433**, 428–433 (2020).

92. Israeli, H. *et al.* Structure reveals the activation mechanism of the MC4 receptor to initiate satiation signaling. **7958**, 1–14 (2021).

93. Kühnen, P., Krude, H. & Biebermann, H. Melanocortin-4 Receptor Signalling : Importance for Weight Regulation and Obesity Treatment. *Trends Mol. Med.* **25**, 136–148 (2019).

94. Montague, C. T. *et al.* Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* **387**, 903–908 (1997).

95. Clément, K. *et al.* A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* **392**, 398–401 (1998).

96. Jackson, R. S. *et al.* Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat. Genet.* **16**, 303–306 (1997).

97. Krude, H. *et al.* Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat. Genet.* **19**, 155–157 (1998).

98. Asai, M. *et al.* Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. *Science (80-.).* **341**, 275–278 (2013).

99. Baron, M. *et al.* Loss-of-function mutations in MRAP2 are pathogenic in hyperphagic obesity with hyperglycemia and hypertension. *Nat. Med.* **25**, 1733–

1738 (2019).

100. Ghamari-Langroudi, M. *et al.* Regulation of energy rheostasis by the melanocortin-3 receptor. *Sci. Adv.* **4**, eaat0866 (2018).

101. Kühnen, P. *et al.* Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. *N. Engl. J. Med.* **375**, 240–246 (2016).

102. Clément, K. *et al.* MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nat. Med.* **24**, (2018).

103. Podyma, B. *et al.* The stimulatory G protein G_s α is required in melanocortin 4 receptor-expressing cells for normal energy balance, thermogenesis and glucose metabolism. *J. Biol. Chem.* jbc.RA118.003450 (2018) doi:10.1074/jbc.RA118.003450.

104. Glas, E., Mückter, H., Gudermann, T. & Breit, A. Exchange factors directly activated by cAMP mediate melanocortin 4 receptor-induced gene expression. *Sci. Rep.* **6**, 32776 (2016).

105. Chiappini, F., Cunha, L. L., Harris, J. C. & Hollenberg, A. N. Lack of cAMP-response element-binding protein 1 in the hypothalamus causes obesity. *J. Biol. Chem.* **286**, 8094–8105 (2011).

106. Yang, L. K. & Tao, Y. X. Biased signaling at neural melanocortin receptors in regulation of energy homeostasis. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1863**, 2486–2495 (2017).

107. Chai, B. *et al.* Melanocortin-4 receptor-mediated inhibition of apoptosis in immortalized hypothalamic neurons via mitogen-activated protein kinase. *Peptides* **27**, 2846–2857 (2006).

108. Breit, A. *et al.* The natural inverse agonist agouti-related protein induces arrestin-mediated endocytosis of melanocortin-3 and -4 receptors. *J. Biol. Chem.* **281**, 37447–37456 (2006).

109. Damm, E., Buech, T. R. H., Gudermann, T. & Breit, A. Melanocortin-induced PKA activation inhibits AMPK activity via ERK-1/2 and LKB-1 in hypothalamic GT1-7 cells. *Mol. Endocrinol.* **26**, 643–654 (2012).

110. Ghamari-Langroudi, M. *et al.* G-protein-independent coupling of MC4R to Kir7.1 in hypothalamic neurons. *Nature* **520**, 94–98 (2015).

111. Anderson, E. J. P. *et al.* Late onset obesity in mice with targeted deletion of potassium inward rectifier Kir7.1 from cells expressing the melanocortin-4 receptor. *J. Neuroendocrinol.* **31**, e12670 (2019).

112. Mountjoy, K. G., Kong, P. L., Taylor, J. A., Willard, D. H. & Wilkison, W. O. Melanocortin receptor-mediated mobilization of intracellular free calcium in HEK293 cells. *Physiol. Genomics* **5**, 11–19 (2001).

113. Newman, E. A. *et al.* Activation of the melanocortin-4 receptor mobilizes intracellular free calcium in immortalized hypothalamic neurons. *J. Surg. Res.* **132**, 201–207 (2006).

114. Li, Y. *et al.* G q / 11 α and G s α mediate distinct physiological responses to central

melanocortins. **126**, 1–10 (2016).

115. Inoue, A. *et al.* Illuminating G-Protein-Coupling Selectivity of GPCRs. *Cell* **177**, 1933-1947.e25 (2019).

116. Avet, C. *et al.* Selectivity Landscape of 100 Therapeutically Relevant GPCR Profiled by an Effector Translocation-Based BRET Platform. *bioRxiv* 1–41 (2020) doi:10.2139/ssrn.3586569.

117. Yu, J. *et al.* SI:Determination of the melanocortin-4 receptor structure identifies Ca2+ as a cofactor for ligand binding. *Science (80-.).* **368**, 428–433 (2020).

118. Heyder, N. A. *et al.* Structures of active melanocortin-4 receptor − Gs-protein complexes with NDP- α -MSH and setmelanotide. *Cell Res.* 1–14 (2021) https://doi.org/10.1038/s41422-021-00569-8.

119. Hsiung, H. M. *et al.* A novel and selective β-melanocyte-stimulating hormone-derived peptide agonist for melanocortin 4 receptor potently decreased food intake and body weight gain in diet-induced obese rats. *Endocrinology* **146**, 5257–5266 (2005).

120. Yan, L. Z. *et al.* Structure-activity relationships of beta-MSH derived melanocortin-4 receptor peptide agonists. *Curr. Top. Med. Chem.* **7**, 1052–1067 (2007).

121. Kumar, K. G. *et al.* Peptides Analysis of the therapeutic functions of novel melanocortin receptor agonists in MC3R- and MC4R-deficient C57BL / 6J mice. **30**, 1892–1900 (2009).

122. Kievit, P. *et al.* Chronic treatment with a melanocortin-4 receptor agonist causes weight loss, reduces insulin resistance, and improves cardiovascular function in diet-induced obese rhesus macaques. *Diabetes* **62**, 490–497 (2013).

123. Paisdzior, S. *et al.* Differential signaling profiles of MC4R mutations with three different ligands. *Int. J. Mol. Sci.* **21**, (2020).

124. O'Donohue, T. L., Handelmann, G. E., Chaconas, T., Miller, R. L. & Jacobowitz, D. M. Evidence that N-acetylation regulates the behavioral activity of α-MSH in the rat and human central nervous system. *Peptides* **2**, 333–344 (1981).

125. Lee, Y. S. *et al.* A POMC variant implicates beta-melanocyte-stimulating hormone in the control of human energy balance. *Cell Metab.* **3**, 135–140 (2006).

126. Bregman, M. D., Sawyer, T. K., Hadley, M. E. & Hruby, V. J. Adenosine and divalent cation effects on S-91 melanoma adenylate cyclase. *Arch. Biochem. Biophys.* **201**, 1–7 (1980).

127. Langendonk, J. G. *et al.* Afamelanotide for Erythropoietic Protoporphyria. *N. Engl. J. Med.* **373**, 48–59 (2015).

128. Sawyer, T. K., Hruby, V. J., Darman, P. S. & Hadleyt, M. A. C. E. [half-Cys4 , half-Cysl0I-a-Melanocyte-stimulating hormone : A cyclic. **79**, 1751–1755 (1982).

129. Sawyer, T. K. *et al.* α-Melanocyte-stimulating hormone structure-activity studies: Comparative analysis of melanotropic and CNS bioactivities. *Synapse* **2**, 288–292 (1988).

130. Sharma, S. *et al.* Current mechanistic and pharmacodynamic understanding of melanocortin-4 receptor activation. *Molecules* **24**, (2019).

131. Li, Y. *et al.* G q / 11 a and G s a mediate distinct physiological responses to central melanocortins Find the latest version : G q / 11 α and G s α mediate distinct physiological responses to central melanocortins. **126**, 40–49 (2016).

132. Nickolls, S. A. *et al.* Molecular determinants of melanocortin 4 receptor ligand binding and MC4/MC3 receptor selectivity. *J. Pharmacol. Exp. Ther.* **304**, 1217–27 (2003).

133. Hruby, V. J. *et al.* Cyclic Lactam α-Melanotropin Analogues of Ac-Nle4-cyclo[Asp5, D-Phe7, Lys10] α-Melanocyte-Stimulating Hormone-(4-10)-NH2 with Bulky Aromatic Amino Acids at Position 7 Show High Antagonist Potency and Selectivity at Specific Melanocortin Receptors. *J. Med. Chem.* **38**, 3454–3461 (1995).

134. BIRMINGHAM, M. K., ELLIOTT, F. H. & VALERK, P. H.-L. THE NEED FOR THE PRESENCE OF CALCIUM FOR THE STIMULATION IN VITRO OF RAT ADRENAL GLANDS BY ADRENOCORTICOTROPHIC HORMONE. *Endocrinology* **53**, 687–689 (1953).

135. Hadley, M. E., Anderson, B., Heward, C. B., Sawyer, T. K. & Hruby, V. J. Calcium-Dependent Prolonged Effects on Melanophores of [4-norleucine , 7-D-phenylalanine] -α-melanotropin Published by : American Association for the Advancement of Science Stable URL : https://www.jstor.org/stable/1687070 American Association for the. *Science (80-.).* **213**, 2–5 (1981).

136. Kopanchuk, S. *et al.* Co-operative regulation of ligand binding to melanocortin receptor subtypes: Evidence for interacting binding sites. *Eur. J. Pharmacol.* **512**, 85–95 (2005).

137. Kenakin, T. Principles: Receptor theory in pharmacology. *Trends Pharmacol. Sci.* **25**, 186–192 (2004).

138. Kenakin, T. Drug Efficacy at G Protein–Coupled Receptors. *Annu. Rev. Pharmacol. Toxicol.* **42**, 349–379 (2002).

139. Xu, P. *et al.* Structural insights into the lipid and ligand regulation of serotonin receptors. *Nature* **592**, 469–473 (2021).

140. Shorr, R. L. R. and C. M. Purification of β-Adrenergic Receptor. *J. Biol. Chem.* **256**, 5820–5826 (1981).

141. Kobilka, B. K. *et al.* Delineation of the intronless nature of the genes for the human and hamster β2-adrenergic receptor and their putative promoter regions. *J. Biol. Chem.* **262**, 7321–7327 (1987).

142. Stoddart, L. A., Kilpatrick, L. E. & Hill, S. J. NanoBRET Approaches to Study Ligand Binding to GPCRs and RTKs. *Trends Pharmacol. Sci.* **39**, 136–147 (2018).

143. Stoddart, L. A. *et al.* Application of BRET to monitor ligand binding to GPCRs. *Nat. Methods* **12**, 661–663 (2015).

144. Wiseman, D. N. *et al.* Expression and purification of recombinant G protein-coupled receptors: A review. *Protein Expr. Purif.* **167**, 105524 (2020).

145. Lagane, B. *et al.* Role of sterols in modulating the human μ-opioid receptor function in Saccharomyces cerevisiae. *J. Biol. Chem.* **275**, 33197–33200 (2000).

146. Schrottke, S. *et al.* Expression, Functional Characterization, and Solid-State NMR Investigation of the G Protein-Coupled GHS Receptor in Bilayer Membranes. *Sci. Rep.* **7**, 1–11 (2017).

147. Shimamura, T. *et al.* Structure of the human histamine H 1 receptor complex with doxepin. *Nature* **475**, 65–72 (2011).

148. Hino, T. *et al.* G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* **482**, 237–240 (2012).

149. Klenk, C., Ehrenmann, J., Schütz, M. & Plückthun, A. A generic selection system for improved expression and thermostability of G protein-coupled receptors by directed evolution. *Sci. Rep.* **6**, 21294 (2016).

150. Schütz, M. *et al.* Directed evolution of G protein-coupled receptors in yeast for higher functional production in eukaryotic expression hosts. *Sci. Rep.* **6**, 21508 (2016).

151. Staus, D. P. *et al.* Sortase ligation enables homogeneous GPCR phosphorylation to reveal diversity in β-arrestin coupling. *Proc. Natl. Acad. Sci.* 201722336 (2018) doi:10.1073/pnas.1722336115.

152. Wingler, L. M. *et al.* Distinctive Activation Mechanism for Angiotensin Receptor Revealed by a Synthetic Nanobody Distinctive Activation Mechanism for Angiotensin Receptor Revealed by a Synthetic Nanobody. *Cell* 1–12 (2019) doi:10.1016/j.cell.2018.12.006.

153. Smith, G. E., Summers, M. D. & Fraser, M. J. Production of human beta interferon in insect cells infected with a baculovirus expression vector. 1983. *Biotechnology* **24**, 434–443 (1992).

154. Luckow, V. A. & Summers, M. D. Trends in the development of baculovirus expression vectors. *Bio/Technology* **6**, 47–55 (1988).

155. McKenzie, E. A. & Abbott, W. M. Expression of recombinant proteins in insect and mammalian cells. *Methods* **147**, 40–49 (2018).

156. Zhao, Y. Improving baculovirus recombination. *Nucleic Acids Res.* **31**, 6e – 6 (2003).

157. Rasmussen, S. G. F. *et al.* Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–388 (2007).

158. Maeda, S. *et al.* Development of an antibody fragment that stabilizes GPCR/G-protein complexes. *Nat. Commun.* **9**, 1–9 (2018).

159. Dawaliby, R. *et al.* Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nat. Chem. Biol.* **12**, 35–39 (2016).

160. Aloia, A. L., Glatz, R. V, McMurchie, E. J. & Leifert, W. R. GPCR Expression Using Baculovirus-Infected Sf9 Cells BT - G Protein-Coupled Receptors in Drug Discovery. in (ed. Leifert, W. R.) 115–129 (Humana Press, 2009). doi:10.1007/978-1-60327-317-6_8.

161. Chen, T. *et al.* Structural basis of ligand binding modes at the human formyl peptide receptor 2. *Nat. Commun.* **11**, 1–9 (2020).

162. García-Nafría, J., Lee, Y., Bai, X., Carpenter, B. & Tate, C. G. Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife* **7**, 1–19 (2018).

163. Sun, B. *et al.* Crystal structure of dopamine D1 receptor in complex with G protein and a non-catechol agonist. *Nat. Commun.* **12**, 1–9 (2021).

164. Zhang, Y. *et al.* Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* (2017) doi:10.1038/nature22394.

165. Pardon, E. *et al.* A general protocol for the generation of Nanobodies for structural biology. *Nat. Protoc.* **9**, 674–693 (2014).

166. Westfield, G. H. *et al.* Structural flexibility of the Gαs α-helical domain in the β 2-adrenoceptor Gs complex. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16086–16091 (2011).

167. Zhang, X. *et al.* Evolving cryo-EM structural approaches for GPCR drug discovery. *Structure* 2021.01.11.426276 (2021) doi:10.1016/j.str.2021.04.008.

168. Knoll, M. and R. E. Das Elektronenmikroskop. *Zeitschrift für Phys.* **78**, 318–319 (1932).

169. Broglie, L. De. Recherches sur la théorie des Quanta Louis De Broglie. *Diss. Migr. en cours d'affectation* (1924).

170. Rutherford, E. LXXIX. The scattering of α and β particles by matter and the structure of the atom . *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **21**, 669–688 (1911).

171. Dubochet, J., Lepault, J., Freeman, R., Berriman, J. A. & Homo, J. -C. Electron microscopy of frozen water and aqueous solutions. *J. Microsc.* **128**, 219–237 (1982).

172. Frank, J. Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state. in (Oxford University Pres, 2006).

173. Penczek, P. A., Zhu, J., Schröder, R., & Frank, J. Three dimensional reconstruction with contrast transfer compensation from defocus series. *Scanning Microsc.* **11**, 147–154 (1997).

174. Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438–449 (2015).

175. Bracewell, R. N. Strip integration in radio astronomy. *Aust. J. Phys.* **9**, 198–217 (1956).

176. Van Heel, M. Angular reconstitution: A posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* **21**, 111–123 (1987).

177. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

178. Schwefel, D. *et al.* Structural basis of lentiviral subversion of a cellular protein degradation pathway. *Nature* **505**, 234–238 (2014).

179. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

180. Zhang, K. & Biology, M. Brief Manual of Gautomatch Features : Requirement : Usage : Basic options : default values , description :

181. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

182. Scheres, S. H. W. Single-particle processing in RELION. *Manuals* 1–21 (2013).

183. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: Adaptive regularization improves single particle cryo-EM reconstruction. *bioRxiv* 1–20 (2019) doi:10.1101/2019.12.15.877092.

184. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).

185. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).

186. Vagin, A. A. *et al.* REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2184–2195 (2004).

187. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).

188. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).

189. Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D. Biol. Crystallogr.* **55**, 191–205 (1999).

190. McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793 (1994).

191. Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).

192. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

193. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

194. Paisdzior, S. *et al.* Differential Signaling Profiles of MC4R Mutations with Three Different Ligands. *Int. J. Mol. Sci.* **21**, 1224 (2020).

195. Biebermann, H. *et al.* A role for beta-melanocyte-stimulating hormone in human body-weight regulation. *Cell Metab.* **3**, 141–146 (2006).

196. Stoddart, L. A. *et al.* Application of BRET to monitor ligand binding to GPCRs Europe PMC Funders Group. *Nat Methods* **12**, 661–663 (2015).

197. Manglik, A. *et al.* Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **161**, 1101–1111 (2015).

198. Hilger, D., Masureel, M. & Kobilka, B. K. Structure and dynamics of GPCR signaling complexes. *Nat. Struct. Mol. Biol.* **25**, 1–34 (2017).

199. Yang, F., Moss, L. G. & Phillips, G. N. J. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246–1251 (1996).

200. Gewering, T., Januliene, D., Ries, A. B. & Moeller, A. Know your detergents: A case study on detergent background in negative stain electron microscopy. *J. Struct. Biol.* **203**, 242–246 (2018).

201. Hilger, D. *et al.* Structural insights into differences in G protein activation by family A and family B GPCRs. *Science (80-.).* **369**, eaba3373 (2020).

202. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).

203. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

204. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).

205. Woolley, M. J. & Conner, A. C. Understanding the common themes and diverse roles of the second extracellular loop (ECL2) of the GPCR super-family. *Mol. Cell. Endocrinol.* **449**, 3–11 (2017).

206. Chai, B.-X. *et al.* Receptor-antagonist interactions in the complexes of agouti and agouti-related protein with human melanocortin 1 and 4 receptors. *Biochemistry* **44**, 3418–3431 (2005).

207. Tarnow, P., Schöneberg, T., Krude, H., Grüters, A. & Biebermann, H. Mutationally induced disulfide bond formation within the third extracellular loop causes melanocortin 4 receptor inactivation in patients with obesity. *J. Biol. Chem.* **278**, 48666–48673 (2003).

208. Hanson, M. a. *et al.* Crystal Structure of a Lipid G Protein-Coupled Receptor. *Science (80-.).* (2012) doi:10.1126/science.1215904.

209. Chrencik, J. E. *et al.* Crystal Structure of Antagonist Bound Human Lysophosphatidic Acid Receptor 1. *Cell* **161**, 1633–1643 (2015).

210. Yin, J. *et al.* Structure of a D2 dopamine receptor–G-protein complex in a lipid membrane. *Nat. 2020* **2020**, 1–5 (2020).

211. Liu, K. *et al.* Structural basis of CXC chemokine receptor 2 activation and signalling. *Nature* **585**, 135–140 (2020).

212. Liu, X. *et al.* Structural Insights into the Process of GPCR-G Protein Complex Formation. *Cell* **177**, 1243-1251.e12 (2019).

213. Proteins, O. F. G. Drug efficacy Kenakin 2002. *N. Engl. J. Med.* **41**, 751–73 (1999).

214. Yang, Y. *et al.* Molecular Determinants of Human Melanocortin-4 Receptor Responsible for Antagonist SHU9119 Selective Activity *. *J. Biol. Chem.* **277**, 20328–20335 (2002).

215. Ma, S. *et al.* Structural mechanism of calcium-mediated hormone recognition and Gβ interaction by the human melanocortin-1 receptor. (2021).

216. Dorr, R. T. *et al.* Evaluation of melanotan-II, a superpotent cyclic melanotropic peptide in a pilot phase-I clinical study. *Life Sci.* **58**, 1777–1784 (1996).

217. Saleh, N., Kleinau, G., Heyder, N. & Clark, T. Binding , thermodynamics and selectivity of a non-peptide antagonist to the Melanocortin-4 receptor. *submitted* (2018).

218. Atchison, D. K. & Beierwaltes, W. H. The influence of extracellular and intracellular calcium on the secretion of renin. *Pflugers Arch.* **465**, 59–69 (2013).

219. Visiers, I., Ballesteros, J. A. & Weinstein, H. Three-dimensional representations of G protein-coupled receptor structures and mechanisms. *Methods Enzymol.* **343**, 329–371 (2002).

220. Patten, C. S., Daniels, D., Suzuki, A., Fluharty, S. J. & Yee, D. K. Structural and signaling requirements of the human melanocortin 4 receptor for MAP kinase activation. *Regul. Pept.* **142**, 111–122 (2007).

221. Yang, L. K. & Tao, Y. X. Alanine scanning mutagenesis of the DRYxxI motif and intracellular loop 2 of human melanocortin-4 receptor. *Int. J. Mol. Sci.* **21**, 1–18 (2020).

222. Lubrano-Berthelier, C. *et al.* Melanocortin 4 receptor mutations in a large cohort of severely obese adults: Prevalence, functional classification, genotype-phenotype relationship, and lack of association with binge eating. *J. Clin. Endocrinol. Metab.* **91**, 1811–1818 (2006).

223. Liu, X. *et al.* Structural Insights into the Process of GPCR-G Protein Complex Formation. *Cell* **177**, 1243-1251.e12 (2019).

224. Moro, O., Lameh, J., Hogger, P. & Sadee, W. Hydrophobic amino acid in the i2 loop plays a key role in receptor-G protein coupling. *J. Biol. Chem.* **268**, 22273–22276 (1993).

225. Du, Y., Duc, N. M., Lodowski, D. T. & Kobilka, B. K. Assembly of a GPCR-G Protein Complex. *Cell* **177**, 1232-1242.e11 (2019).

226. Galés, C. *et al.* Real-time monitoring of receptor and G-protein interactions in living cells. *Nat. Methods* **2**, 177–184 (2005).

227. Gregorio, G. G. *et al.* Single-molecule analysis of ligand efficacy in β2AR–G-protein activation. *Nature* (2017) doi:10.1038/nature22354.

228. Ernst, O. P., Gramse, V., Kolbe, M., Hofmann, K. P. & Heck, M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10859–10864 (2007).

229. Mobbs, J. I. *et al.* Structures of the human cholecystokinin 1 (CCK1) receptor bound to Gs and Gq mimetic proteins provide insight into mechanisms of G protein selectivity. *PLOS Biol.* **19**, e3001295 (2021).

230. Kievit, P. *et al.* Chronic treatment with a melanocortin-4 receptor agonist causes weight loss, reduces insulin resistance, and improves cardiovascular function in diet-induced obese rhesus macaques. *Diabetes* **62**, 490–497 (2013).

Supplementary data

Table S1 ***Data collection, processing, and refinement statistics.***

	NDP-α-MSH-MC4R-G_sαβγ-Nb35 complex	**Setmelanotide-MC4R-G_sαβγ-Nb35 complex**
Magnification	96000	
Voltage (kV)	300	
Electron exposure (e^-/ Å²)	40	
Defocus range (μm)	-0.8 / -2.0	
Pixel size (Å)	0.832	
Symmetry imposed	*C1*	
Micrographs used (total)	5403 (5618)	6979 (7583)
Initial particle image (no.)	2746119	4330500
Final particle images (no.)	221682	370621
Map resolution (Å)	2.86	2.58
FSC threshold	0.143	0.143
Refinement		
Initial model used (PDB ID:)	6W25 / 3SN6	NDP-MC4R-G_sαβγ-Nb35
Model resolution (Å)	2.88	2.60
Model resolution range (Å)	233 – 2.88	233 – 2.60
Map sharpening B factor (Å²)	-99	-36
Model composition		
Total atoms	7814	7788
Water	115	88
Protein atoms	7696	7696
B-factors (Å²)		
Overall	69.37	48.11
R.m.s. deviations bond length (Å)	0.013	0.013
R.m.s. deviations bond angle (°)	1.705	1.707
Validation		
Molprobity score	1.54	1.42
Clash score	3.53	3.39
Poor rotamers (%)	0.86	0.97
Ramachandran plot		
Favoured (%)	94.09	95.77
Allowed (%)	5.70	4.13
Disallowed (%)	0.21	0.11

Publications

N.A. Heyder, G. Kleinau, D. Speck, A. Schmidt, S. Paisdzior, M. Szczepek, B. Bauer, A. Koch, M. Gallandi, D.M Kwiatkowski, J. Bürger, T. Mielke, A. Beck-Sickinger, P.W. Hildebrandt, C.M.T. Spahn, D. Hilger, M. Schacherl, H. Biebermann, T. Hilal, P. Kühnen, B.K. Kobilka and P. Scheerer. Structures of active melanocortin 4 receptor Gs-protein complexes with NDP-α-MSH and setmelanotide *Cell Research*, 2021. https://www.nature.com/articles/s41422-021-00569-8

G. Kleinau, **N.A. Heyder**, Y.Tao, P. Scheerer. Structural complexity and plasticity of signaling regulation at the melanocortin-4 receptor *International Journal of Molecular Sciences*, 21(16), 1-29. https://doi.org/10.3390/ijms21165728

N.A. Heyder, G. Kleinau, M. Szczepek, D. Kwiatkowski, D. Speck, L. Soletto, J. M. Cerdá-Reverter, H. Krude, P. Kühnen, H. Biebermann, and P. Scheerer. Signal transduction and pathogenic modifications at the melanocortin 4 receptor: A structural perspective Frontiers in Endocrinology, 10(July), 1-18. https://doi.org/10.3389/fendo.2019.00515

K. Clément, H. Biebermann, I. Farooqi, L. Van Der Ploeg, B. Wolters, C. Poitou, L. Puder, F. Fiedorek, K. Gottesdiener, G. Kleinau, **N.A. Heyder**, P. Scheerer, U. Blume-Peytavi, I. Jahnke, S. Sharma, J. Mokrosinski, S. Wie- gand, A. Müller, K. Weiß, K. Mai, J. Spranger, A. Grüters, O. Blankenstein, H. Krude, and P. Kühnen. MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nature Medicine*, 24(5), 551-555. https://www.nature.com/articles/s41591-018-0015-9

N. Saleh, G. Kleinau, **N.A. Heyder**, T. Clark, P. W. Hildebrand, and P. Scheerer Binding, thermodynamics, and selectivity of a non-peptide antagonist to the melanocortin-4 receptor, *Frontiers in Pharmacology*, 10(July), 1-18. https://doi.org/10.3389/fendo.2019.00515