Integrative Modeling to Determine the Activity, Molecular Recognition, and Membrane Trafficking of the G-Protein Coupled Bile Acid Receptor TGR5

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Christoph Gerhard Wilhelm Gertzen

aus Oberhausen

Düsseldorf, den 01.08.2016

Aus dem Institut für Medizinische Chemie und Pharmazie

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Holger Gohlke

Korreferentin: Prof. Dr. Verena Keitel-Anselmino

Tag der mündlichen Prüfung: 15.12.2016

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Diese Dissertation wurde in der vorgelegten oder einer ähnlichen Form noch bei keiner anderen Institution eingereicht, und es wurden bisher keine erfolglosen Promotionsversuche von mir unternommen.

Düsseldorf, im August 2016

To Corinna and my family

“Build a man a fire, and he’ll be warm for a day.

Set a man on fire, and he’ll be warm for the rest of his life.”

-Terry Pratchett

TABLE OF CONTENT

LIST OF PUBLICATIONS

This thesis is based on the following publications:

Peer-reviewed publications:

[§]Spomer, L., **[§]Gertzen, C.G.W.**, Schmitz, B., Häussinger, D., Gohlke, H., Keitel, V; *A membrane-proximal, C-terminal alpha-helix is required for TGR5 plasma membrane localization and function*. **J Biol Chem (2014)**, 289, 3689-3702.
Impact factor reported in 2014: 4.573

[§]Gertzen, C.G.W., [§]Spomer, L., Smits, S.H.J., Häussinger, D., Keitel, V., Gohlke, H.; *Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists*. **Eur J Med Chem (2015)**, 104, 57-72.
Impact factor reported in 2014: 3.447

Greife A., Felekyan S., Ma Q., **Gertzen C.G.W.**, Spomer L., Dimura M., Peulen T.O., Wöhler C., Häussinger D., Gohlke H., Keitel V., Seidel C.A.M.; *Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study TGR5 in live cells*. **Sci Rep (2016)**, 6, 36792.
Impact factor reported in 2015: 5.228

[§] Both authors contributed equally to this work.

ABBREVIATIONS

A_{2A}-AR	A_{2A}-adenosine receptor
ADAM17	A disintegrin and metalloproteinase 17
AFMoC	Adaptation of fields for molecular comparison
AV	Accessible volume
β_2-AR	β2-adrenergic receptor
CA	Cholic acid
cAMP	Cyclic adenosine monophosphate
CRE	cAMP responsive element
CREB	cAMP responsive element binding factor
CDCA	Chenodeoxycholic acid
DCA	Deoxycholic acid
EGFR	Endothelial growth factor receptor
EL	Extracellular loop
ER	Endoplasmic reticulum
FoxO1	Forkhead box transcription factor O1
GABA	γ-Aminobutyric acid
G-CSF	Granulocyte-colony stimulation factor
GERD	Gastric-esophageal reflux disease
GLP-1	Glucagon-like peptide-1
GPBAR-1	G-protein coupled bile acid receptor 1
GPCR	G-protein coupled receptor
G-protein	Guanylyl cyclase binding protein
GRK	GPCR kinases
IL	Intracellular loop
JNK	c-Jun N-terminal kinases
κ-OR	κ-Opioid receptor

LCA	Lithocholic acid
LDL	Low-density lipoprotein
M-BAR1	Membrane-type bile acid receptor 1
MDR3	Multidrug resistance protein 3
MFIS	Multiparameter Image Fluorescence Spectroscopy
MM-PBSA	Molecular Mechanics-Poisson Boltzmann Surface Area
MD	Molecular Dynamics
μ-OR	μ-Opioid receptor
NASH	Non-alcoholic steatohepatitis
NOX5-S	NADPH oxidase 5-S
PM	Plasma membrane
Rab-GTPases	Ras-related in brain-GTPases
RMSD	Root mean square deviation
T1DM	Type-1 diabetes mellitus
T2DM	Type-2 diabetes mellitus
T_4	Thyroxine
TGR5	Tanaka G-protein coupled receptor 5
TLC	Taurolithocholic acid
TM	Transmembrane helix
TNFα	Tumor necrosis factor α
TRPA1	Transient receptor potential ankyrin 1
UDCA	Ursodeoxycholic acid

ZUSAMMENFASSUNG

TGR5 ist der zuerst entdeckte gallensalzaktivierte G-Protein gekoppelte Rezeptor. Er wird von Gallensalzen und Neurosteroiden aktiviert. Seine physiologischen Funktionen bestehen sowohl in der Regulation des Blutzuckerspiegels als auch dem Schutz der Zellen des Gallengangs durch antiapoptotische und proliferative Signale. Bei einer Überexpression des Rezeptors in Gallengangzellen kommt es jedoch zur Bildung des Cholangiozytenkarzinoms. Derzeit ist keine spezielle Therapie dieser Krebskrankheit möglich, da keine TGR5-Antagonisten bekannt sind, die das Übermaß an antiapoptotischen und proliferativen Signale blockieren könnten. Die Entwicklung von TGR5-Antagonisten ist nicht trivial, jedoch könnte die Kenntnis des Bindemodus von TGR5-Agonisten den Entwurf von Antagonisten lenken und so deren Entwicklung vereinfachen. In **Publikation III** habe ich in Kooperation mit der Arbeitsgruppe von Prof. Dr. V. Keitel ein Bindemodenmodell von TGR5-Agonisten vorhergesagt. Dieser Bindemodus ist in Einklang mit neun Mutationen, die sowohl die Aktivität von TGR5 als auch die Affinität der Liganden beeinflussen.

Damit TGR5 seine Effekte ausüben kann, muss dieser nach Synthese im endoplasmatischen Retikulum zur Plasmamembran transportiert werden. Allerdings ist nicht bekannt, welche Faktoren diesen Transport bedingen. Die meisten GPCRs enthalten hierfür gewöhnlich ein Sortiermotiv in ihrem C-Terminus, welches TGR5 nicht besitzt. In **Publikation I** konnte ich mit der Arbeitsgruppe von Prof. Dr. V. Keitel zeigen, dass die Bildung einer α-Helix im C-Terminus verantwortlich für den Membrantransport ist. TGR5 Varianten, die keine α-Helix im C-Terminus aufwiesen, verblieben im endoplasmatischen Retikulum. Wir haben Andeutungen gefunden, dass dies in allen GPCRs der Grund für den Membrantransport sein könnte und Sortiermotive die Aufgaben haben, die Bildung der α-Helix im C-Terminus zu forcieren.

Homodimerisierung von TGR5 könnte ein Grund für die Bedeutung dieser Sekundärstruktur für die Membranlokalisierung sein. Für GPCRs wurde gezeigt, dass Dimerisierung im endoplasmatischen Retikulum der auslösende Faktor für deren Membrantransport ist. Mit den Arbeitsgruppen von Prof. Dr. V. Keitel und Prof. Dr. C. Seidel konnte ich in **Publikation II** zeigen, dass TGR5 in der Tat den C-Terminus als Dimerisierungsschnittstelle benutzt. Darüber hinaus konnten wir zeigen, dass TGR5 Oligomere bildet und hierfür zwei weitere potentielle Interaktionsflächen identifizieren.

ABSTRACT

TGR5 is the first known bile acid-sensing G-protein coupled receptor. TGR5 is activated by bile acids and neurosteroids. The physiological roles of TGR5 include the regulation of blood glucose levels and the protection of bile duct cells via anti-apoptotic and proliferative effects. An overexpression of TGR5 in bile duct cells, however, leads to the formation of the cholangiocarcinoma. Currently, no specialized therapy of the cholangiocarcinoma is available as no TGR5 antagonists, inhibiting the abundancy of proliferative and anti-apoptotic effects, are known. The design of TGR5 antagonists is not trivial, but knowledge of the binding mode of TGR5 agonists could guide the design of antagonists, which could simplify their development. In **publication III** I discovered a binding mode model of TGR5 agonists in cooperation with the working group of Prof. Dr. V. Keitel using integrative modeling. This binding mode is in agreement with nine mutations, including negative controls influencing the activation of TGR5 as well as agonist affinity.

In order to exert its effects, TGR5 needs to be transported to the plasma membrane after its synthesis in the endoplasmic reticulum. However, the determinants for its membrane trafficking are unknown. For this trafficking, most of the GPCRs bear sorting motifs in their C-termini, which TGR5 does not. In **publication I,** I could show in collaboration with the working group of Prof. Dr. V. Keitel that α-helix formation in the TGR5 C-terminus is responsible for its membrane trafficking. TGR5 variants, which did not show α-helix formation, remained in the endoplasmic reticulum. We discovered hints, that this α-helix formation could be the determining factor for membrane localization in all GPCRs, while the sorting motifs facilitate α-helix formation.

Homodimerization of TGR5 could be the reason for the importance of this secondary structure for its membrane localization. For other GPCRs it could be shown that dimerization in the endoplasmic reticulum triggers their membrane trafficking. In cooperation with the working groups of Prof. Dr. V. Keitel and Prof. Dr. C. Seidel, I could show in **publication II** that TGR5 utilizes its C-terminus in a dimerization interface. Furthermore, we could show that TGR5 forms oligomers, for which we identified two possible interfaces.

1 INTRODUCTION

G-protein coupled receptors (GPCRs) are cell-surface receptors which constitute a very important gene family of receptors as they are present in virtually every type of tissue in mammals, and thus the human body[1, 2]. They can be found in brain tissues[3-6], the retina[7, 8], the lung[9, 10], the heart[11, 12], the kidney[13, 14], and the intestine[15-17], only to name a few. Their prominence throughout the body allows them to take part in a variety of regulatory mechanisms. These mechanisms include neurological, cardiovascular, endocrine, and reproductive functions and encompass the senses of taste, smell and vision[18, 19]. Thus, GPCRs offer a wide range of possibilities to alter processes in the body, which makes them valuable drug targets[18-20]. In fact, GPCRs are among the most targeted gene families on the current drug market[21]. Accordingly, 27% of all present drugs influence the activity of GPCRs[21], although biologicals, especially monoclonal antibodies not targeting GPCRs, are on a steady rise[22-24]. Yet, with 40% of all prescriptions being GPCR modulating substances, the importance of GPCRs among drug targets is undisputed[1, 25, 26]. Hence, drugs targeting more than 50 different GPCRs, among them novel targets, are currently in the pipeline[18, 27].

In order to participate in such diverse mechanisms, multiple different GPCRs are necessary. To date, about 900 different human GPCRs are known of which 400 are nonolfactory receptors[1, 27, 28]. These 900 receptors can be divided into five classes depending on their phylogeny: The family of rhodopsin-like receptors (class A) including catecholamine, olfactory, and many peptide receptors; the family of glucagon-like receptors (class B) including peptide and secretin receptors; the family of metabotropic glutamate receptors (class C) including GABA, pheromone, and taste receptors; and the families of frizzled and smoothened receptors (class D, and E, respectively)[29-33]. In drug discovery, the group of nonolfactory receptors, especially class A receptors, have been subject to extensive investigation while the olfactory and taste receptors, only being involved in smell and taste, have been widely neglected[1, 27, 28]. Recent studies however, show an occurrence of olfactory and taste GPCRs in tissues, which are not associated with those senses, effectively increasing the number of possibly targetable GPCRs by about 500. Such olfactory and taste receptors have been hypothesized to be responsible for nutrient sensing in the heart[11], found to influence the chemotaxis of sperms[34] or inhibit the proliferation of prostate cancer cells[35]. Moreover, some of them regulate glucose absorption in the intestine and may thus be future targets for the treatment of metabolic diseases[15]. An intestinal target for metabolic diseases,

though a nonolfactory GPCR regulating blood glucose homeostasis, is the Tanaka G-protein coupled receptor 5 (TGR5)[36, 37] (**chapter 2.1.1**).

TGR5, a class A receptor and the subject of this thesis, is the first known bile acid-sensing GPCR and is expressed throughout the body[38-42]. Locations with high expression levels of TGR5 are the liver, the bile duct, gallbladder, intestine, immunocompetent cells, and the brain[43-49]. In the brain, bile acids cannot be the natural agonists of TGR5 as they are actively excreted through the blood-brain barrier, and thus cannot reach the brain[50, 51]. Here, neurosteroids act as the natural ligands[46] (**chapter 2.1.2**). The effects mediated by TGR5 are highly interesting to exploit from a drug developmental perspective. For instance, upon activation TGR5 has the ability to reduce inflammation (**chapter 2.1.5**) and thus attenuate atherosclerosis, which has been shown in mice[52-55]. Additionally, the activation of TGR5 has beneficial effects on metabolism and energy expenditure (**chapter 2.1.4**) what could be utilized in the treatment of diabetes and other metabolic diseases[36, 49, 52, 56]. For the latter, the popularity of TGR5 agonists has had a setback, as they also lead to gallbladder filling which can be quite unpleasant in patients[57]. However, not only substances activating TGR5 are of high pharmacological interest, as it has been found to be overexpressed in esophagus, gastric, and gallbladder cancers[58-60] (**chapter 2.1.7**). Upon activation, TGR5 has been shown to increase proliferation and reduce apoptosis in cells, so the enhancement of these effects due to overexpression may be a cause for the development of these types of cancer[49, 56, 61]. Furthermore, TGR5 is responsible for mediating bile acid-induced pruritus[55, 62, 63]. Thus, substances inhibiting the activation of TGR5 are very promising drugs for the treatment of TGR5 mediated forms of cancer, and pruritus in cholestatic diseases. Unsurprisingly, a lot of effort has been put into the identification of new TGR5 ligands[64-75]. However, no inhibitors of TGR5 are currently known, and despite extensive research on the field of GPCRs, the design of antagonists is not trivial. The absence of an X-ray crystal structure of TGR5 and thereby a binding mode of agonists in TGR5 complicates the endeavor to identify TGR5 antagonists or more potent agonists. Here, computational structure-based methods (**chapter 2.3**) can help to predict the binding pose of ligands and, thus, guide the design of new and more potent drugs[76].

The application of such methods requires knowledge about the target structure but the absence of a crystal structure of TGR5 seems to contradict their use. However, if a crystal structure is unavailable, a prediction of the structure can be made by homology modeling. In homology modeling, the crystal structures of one or several evolutionarily related proteins are used to build a structural model of the target protein. For this, homology modeling

(**chapter 2.3.2**) utilizes the fact that protein structure is more conserved than sequence[77]. This means that two related proteins, e.g. from the same class of GPCRs, adopt a very similar structure despite their sequences showing a very low identity to each other. As a matter of fact, all GPCRs, which have been crystallized to date, exhibit seven transmembrane α-helices (TMs) and differ mainly in the structure of their extra- or intracellular parts[18, 78] (**chapter 2.2.1**). The general shape of the TMs, which also constitute the binding site, is highly conserved among GPCRs, which renders the subsequent prediction of an accurate binding mode more likely[18, 78-80].

In order to predict a binding mode, the method of molecular docking (**chapter 2.3.3**) is usually used. Here, conformations of a known ligand inside the binding pocket of the target are generated and subsequently evaluated energetically. This evaluation can be done according to various methodologies ranging from the application of force fields to the use of knowledge-based potentials[81-86]. The methods predict the energetically most favorable conformation (binding mode) of the ligand in the binding pocket and thus the most populated one. In the field of GPCRs the combination of homology modeling and molecular docking was applied with great success leading to the prediction of near native binding poses[79, 87, 88]. Moreover, these predicted binding modes were so accurate that on their basis new ligands could be identified for several GPCRs via virtual screening[89-92]. Molecular docking often does not incorporate target flexibility, which can be achieved by the use of molecular dynamics simulations (MD simulations), however[93, 94]. Here, the molecular motions e.g. of the GPCR with a ligand embedded into a cell membrane are predicted on a femtosecond to microsecond scale.

At the beginning of this thesis, the membrane localization determining factors of TGR5 were elusive as the C-terminus of TGR5 does not contain a known sorting motif in the C-terminus. A factor for the membrane localization of TGR5 could also be its dimer formation in the endoplasmic reticulum, leading to a membrane trafficking (**chapter 2.2.3**). Here, the C-terminus of TGR5 may play a role, as it is utilized in a known dimer interface of GPCRs. TGR5 had been identified to form dimers and higher-order oligomers[95], but the orientation of TGR5 upon dimer- and oligomerization was also unknown as was the binding mode of TGR5 agonists. Hence, I applied the aforementioned computational methods in an interdisciplinary integrative modeling approach to investigate the determinants of the membrane localization, the dimerization interfaces, and the ligand recognition of TGR5.

2 BACKGROUND

First, I will review the knowledge about the structure and ligand recognition spectrum of TGR5 and its functions in health and disease in order to demonstrate the value of TGR5 as a potential drug target. Then I will provide an overview of the structural properties and modes of activation of GPCRs in general. After this, I will highlight known factors that influence the activation and membrane localization of GPCRs. This knowledge is crucial in order to understand the decision-making process and the results of this thesis and was extensively used to bolster the modeling process with experimental data. Finally, I will provide an introduction to FRET, and structure-based methods with a focus on integrative modeling and docking, which form the basis of my thesis.

2.1 TGR5

Comprehensive information about TGR5 and its physiological roles can be found in refs.[49, 52, 54, 59, 96-99].

2.1.1 The GPCR TGR5

TGR5, also known as the G-protein coupled bile acid receptor-1 (GPBAR-1) or Membrane-type bile acid receptor 1 (M-BAR 1), is the first known bile acid-sensing G-protein coupled receptor and is activated by bile acids and neurosteroids[39, 40, 46, 65, 68, 70, 72, 73, 100]. It was first discovered by T. Maruyama *et al.* in the year 2002 and can thus be considered a relatively young GPCR[39]. Its first detection in the laboratory of K. Tanaka with T. Maruyama as a coworker lead to the propagation of the most commonly used abbreviation TGR5, the Tanaka G-protein coupled receptor 5. It is a member of the family of class A GPCRs with a length of 330 amino acids (UniProt ID: Q8TDU6)[101, 102]. Although no X-ray crystal structure of TGR5 is known, we can infer from its membership in class A GPCRs that its structure resembles seven TMs with a schematic representation shown in **Figure 1**. The intracellular loop 3 (IL3) is 42 amino acids long and unstructured, as is the C-terminus of TGR5. This was the structural information available about TGR5 at the beginning of this thesis[103-105]. While in many X-ray crystal structures of GPCRs the C-terminus adopts an α-helix, not all GPCRs show such secondary structure formation. The C-terminus of the CXCR4 receptor for example shows loop formation in the crystal structure[106]. The C-terminus of TGR5 does not contain a known sorting motif for the trafficking from the endoplasmic reticulum to the membrane[42]. While several sorting motifs have been identified

in the C-termini of GPCRs (**chapter 2.2.3**), their exact mode of action remains elusive as they are quite diverse[28, 107].

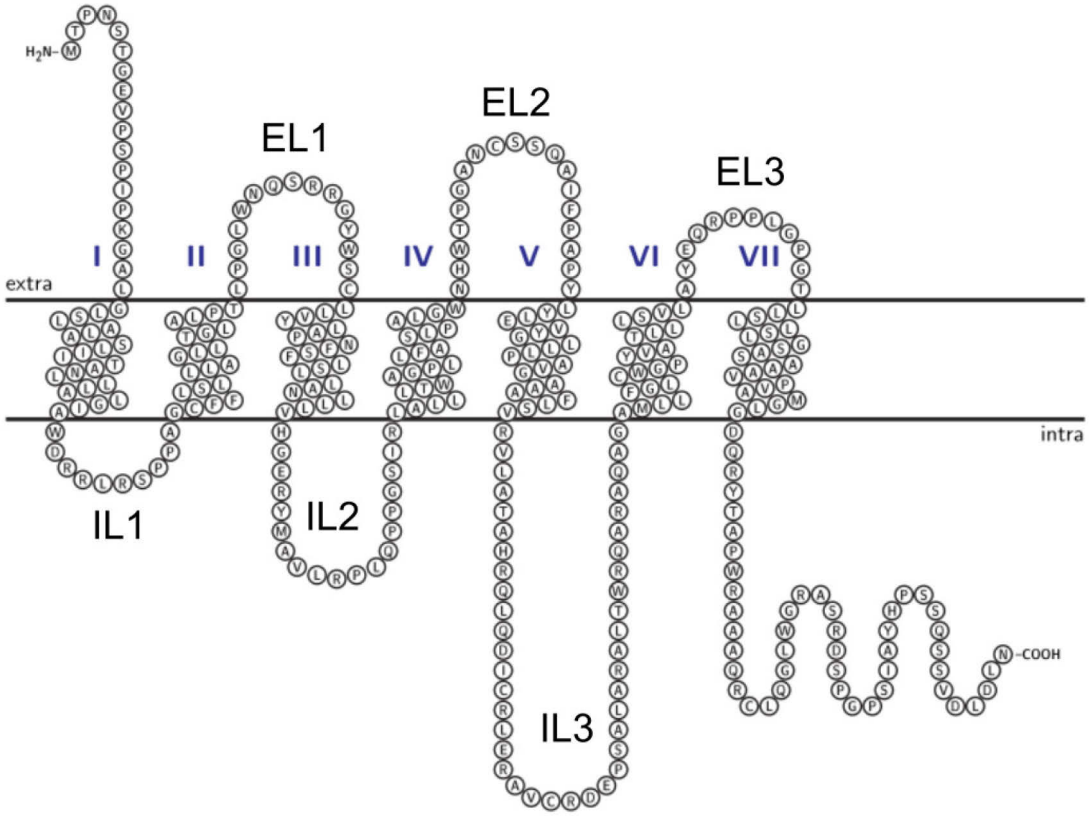


Figure 1 Snake-plot of the sequence of TGR5. This plot shows the sequence of TGR5 in an arrangement indicating the TMs. The boundaries of the TMs were predicted by information available from the GPCRDB[105] at the beginning of this thesis. Similarly, the IL3 is long while the C-terminus of TGR5 is unstructured. The roman numbers indicate the seven TMs typical for GPCRs connected by the extracellular (EL) and intracellular loops (IL). Figure adapted from **publication I**.

Generally, it is assumed that these motifs are recognized by Rab-GTPases (Ras-related in brain-GTPases), which traffic the receptors to the membrane and lead to endocytosis[108]. However, none of those sorting motifs are present in the C-terminus of TGR5, which poses the question what the determining factors of the membrane trafficking of TGR5 are. As the membrane trafficking of TGR5 is crucial for its activity, this question has been addressed in **publication I**.

2.1.2 TGR5 ligand recognition spectrum

The factor influencing the activity of TGR5 is its activation by ligands. As stated before, the natural ligands of TGR5 are bile acids and neurosteroids. The latter has only been determined recently[46, 65]. Previously, scientists were interested in the identity of the natural ligands in the brain, as bile acids cannot penetrate the blood-brain barrier and are also not synthesized in the brain in relevant amounts[50, 51, 96]. The discovery of neurosteroids such as estradiol as the natural ligands in the brain solved this issue although they are not as potent

agonists as bile acids[46, 65]. The potency of bile acids to activate TGR5 correlates with the hydrophobicity of their cholane scaffold. Essentially, the more hydrophobic the cholane scaffold of the respective bile acid, the lower the concentration of the bile acid is needed to activate TGR5. Thus, primary bile acids such as cholic acid (CA) or chenodeoxycholic acid (CDCA) bearing more hydroxyl groups are less potent than secondary bile acids such as deoxycholic acid (DCA) or lithocholic acid (LCA), which are more hydrophobic (**Figure 2**). Not only the number of hydroxyl groups on the cholane scaffold is important for the effectiveness of the respective bile acids, but also their stereochemical configuration is relevant. The configuration of the hydroxyl group in position seven of the cholane scaffold has the highest impact on the activity of bile acids if present. This is most prominent when comparing CDCA to ursodeoxycholic acid (UDCA).

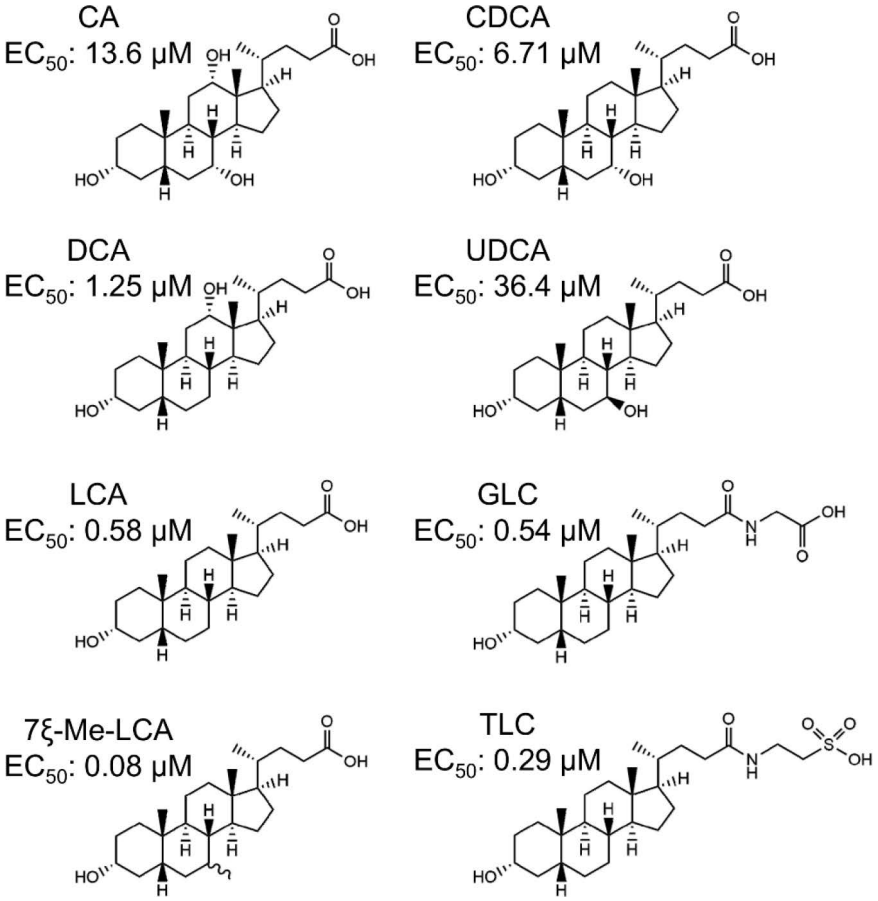


Figure 2 Bile acid agonists alongside their EC_{50} values towards TGR5 as reported in ref.[65] Primary bile acids: CA, CDCA, and UDCA. Secondary bile acids: DCA, LCA, GLC, and TLC. The primary bile acids are generally less effective TGR5 agonists than the secondary bile acids. Among the primary bile acids, the configuration of the hydroxyl group in position seven strongly influences the activity, if comparing CDCA to UDCA. Conjugation of the acid moiety with glycine increases the activity towards TGR5 only slightly, while taurine conjugation increases the activity markedly.

The two bile acids CDCA and UDCA are epimers, as they only differ in the configuration of their hydroxyl group in position seven (**Figure 2**) [65, 76]. In CDCA the hydroxyl group is oriented in the α-position while in UDCA it is oriented in the β-position. This small change

leads to a six-fold reduction in EC_{50} with CDCA being the more potent epimer of the two (**Figure 2**). The reason for this epimeric selectivity of TGR5 was unknown, however the residue in TGR5 causing this epimeric selectivity was discovered in **publication III**. Yet, not only the decoration of the cholane scaffold influences the activity of bile acids towards TGR5. The acid moiety of bile acids can be conjugated with taurine or glycine resulting in a longer and more flexible linker between the cholane scaffold and the acidic moiety. In the case of taurine conjugates the acid moiety is changed to sulfonic acid (**Figure 2**). Conjugated bile acids are more potent TGR5 activators than unconjugated bile acids. Here, taurine conjugates such as taurolithocholic acid (TLC) are more active than glycine conjugated analogs such as glycolithocholic acid (GLC) which is only slightly more active than LCA (**Figure 2**)[65]. Whether the reason for this is the change of the acid moiety resulting in more favorable contacts to TGR5, the elongation of the linker by one methylene unit, or a combination of both was unknown and has been discovered in **publication III**. For this, the ligand dataset devised by Sato *et al*. was utilized in order to elucidate a binding mode model of TGR5 agonists[65]. Sato *et al.* were able to increase the efficacy of bile acid agonists towards TGR5 via addition of hydrophobic substituents in position seven of the cholane scaffold with their most potent ligand being 7ξ-Me-LCA (**Figure 2**)[65]. Recently, more potent and selective TGR5 agonists have been reported, which used a similar approach to increase the efficacy[109]. Upon activation by its ligands, TGR5 mediates signals through several pathways.

2.1.3 TGR5 signaling pathways

With the discovery of TGR5 it could be shown that TGR5 is most commonly interacting with the G_s-protein. Most TGR5 signaling pathways are mediated via this interaction. It was discovered that TGR5 is additionally able to couple to G_q- and G_{i3}-proteins[59]. This variety of G-proteins utilized for TGR5 downstream signaling enables a wide spectrum of different effects, which TGR5 can exert in different tissues. An overview is given in **Figure 3**, and the effects will be discussed in the following chapters in detail.

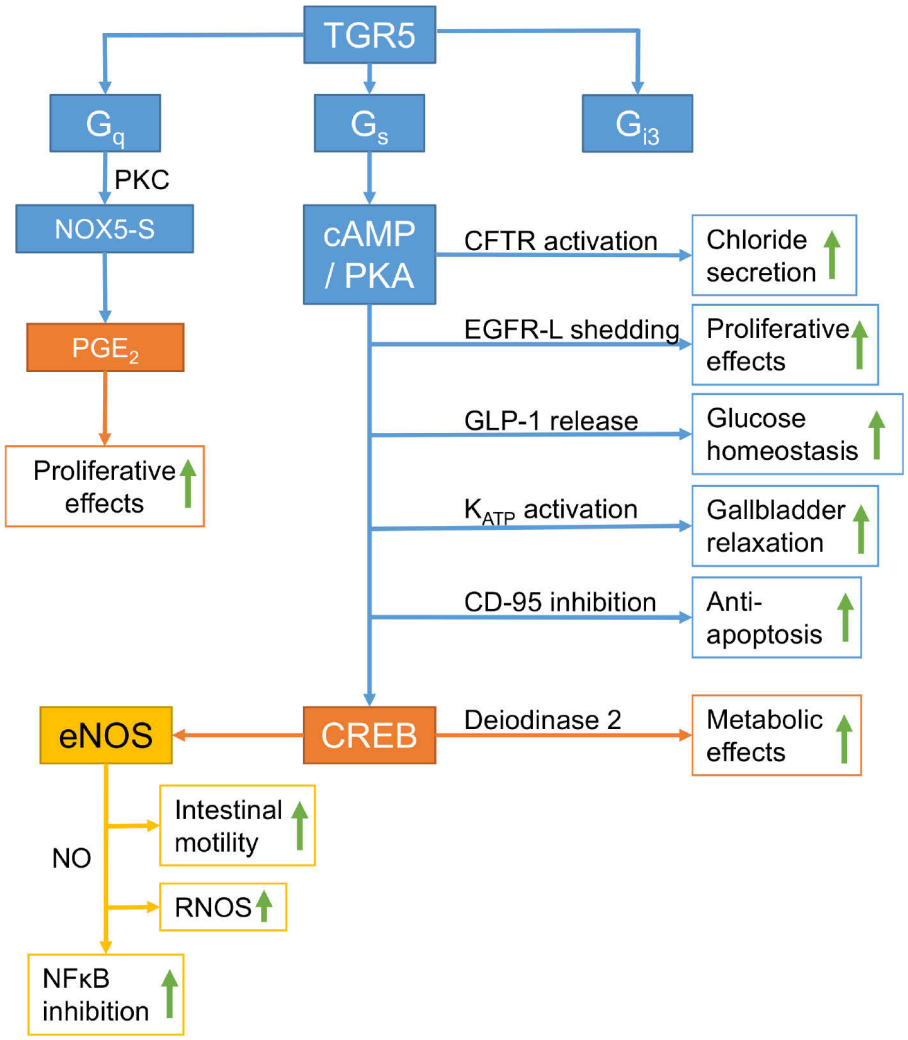


Figure 3 Signal transduction pathways of TGR5. The TGR5 dependent activation of a G_q-protein leads to a protein kinase C (PKC) mediated increase in the NADPH oxidase 5-S (NOX5-S) expression which induces proliferation via prostaglandin E_2 (PGE_2)[59, 110]. Furthermore, TGR5 dependent G_s-protein coupling with subsequent cAMP production and protein kinase A (PKA) activation has a myriad of effects in different tissues: In cholangiocytes it can activate the cystic fibrosis transmembrane conductance regulator (CFTR) facilitating chloride secretion into the bile. It stimulates a disintegrin and metalloproteinase 17 (ADAM17) activity[111]. This sheddase releases the soluble ectodomain of the EGFR ligand (EGFR-L), which is in return able to activate EGFR itself, leading to proliferative effects[111]. In enteroendocrine L-cells, it furthers the release of glucagon-like peptide 1 (GLP-1), which enhances the effect of insulin on cells thus positively influencing glucose homeostasis[112]. Again in cholangiocytes it increases the activation of ATP dependent potassium channels (K_{ATP}) resulting in a relaxation of the adjacent smooth muscle cells and therefore of the gallbladder[113]. The phosphorylation and inactivation of the death receptor cluster of differentiation 95 (CD-95) in cholangiocytes also mediates anti-apoptotic effects[56, 114]. If cAMP binds to the cAMP responsive element binding factor (CREB) this can on the one hand increase the expression of deiodinase 2, which produces thyroxin and thus has a positive impact on metabolism[99]. On the other hand, CREB increases the expression of the endothelial nitric oxide synthase (eNOS) increasing the production of nitric oxide (NO)[96, 115]. Consequentially, NO increases the intestinal motility, turns into reactive NO species (RNOS) and could also be shown to decrease oxidative stress on cells by NFκB inhibition. In various cell types, such as endothelial cells and macrophages, NFκB inhibition decreases the expression of cytokines[52, 96]. G_{i3}-protein coupling of TGR5 could be shown but no signaling pathways are identified to date[59, 110].

After the discovery of TGR5, scientists were interested in the cells and tissues TGR5 is expressed in. Considering the spectrum of agonists TGR5 recognizes, which are most prevalent in the gastro-intestinal tract and the brain, it is unsurprising that TGR5 is found in high expression levels in astrocytes of the brain and gut, in sinusoidal epithelial cells,

gallbladder epithelium, and Kupffer cells in the liver[43, 44, 46, 48]. Furthermore, the mRNA of TGR5 has been found in various cell types as enteroendocrine L-cells of the intestine, CD14-positive alveolar macrophages of the lung, the thyroid gland, subcutaneous adipose tissue, and brown adipose tissue and skeletal muscles[38, 112, 116, 117]. The expression of TGR5 in such a variety of tissues opens many opportunities to influence body functions in which TGR5 is involved.

2.1.4 TGR5 in metabolism

One of the major physiological roles of TGR5 is the regulation of blood glucose homeostasis and metabolism. For the former, TGR5 regulates the blood glucose levels in the body via the release of glucagon-like peptide-1 (GLP-1) in enteroendocrine cells[36, 39, 52, 55, 118-122]. GLP-1 enhances the effect of insulin on cells and thus increases their glucose uptake. The release of GLP-1 is controlled by the farnesoid X receptor (FXR), the fibroblast growth factor 15/19 receptor (FGF15/19), and TGR5[15, 36, 39, 55, 118]. Moreover, TGR5 has been shown to increase energy expenditure and oxygen consumption by increasing the amount of active thyroxine (T_4) in thyroid, skeletal muscle, and brown adipose tissue[56, 99, 112].

2.1.5 TGR5 in the immune response

TGR5 is expressed in a variety of cells partaking in the immune response. Those cells encompass macrophages, including alveolar macrophages and Kupffer cells, monocytes, sinusoidal endothelial cells, and dendritic cells[43, 52, 53, 96, 97]. The latter two cell types are antigen-presenting cells, which differentiate from monocytes upon stimulation with granulocyte-colony stimulation factor (G-CSF) and interleukin-4[96, 123]. However, if TGR5 expressing monocytes are treated with a TGR5 specific agonist, they differentiate into less functional cells, producing lower amounts of cytokines as IL-12. Macrophages, which have been exposed to TGR5 specific agonists, showed lower expression of TNF-α, and a higher inhibition of NF-κB rendering them less active[43, 52, 53, 96]. This TGR5 dependent increase in NF-κB inhibition is also present in endothelial cells, which in return express less adhesion molecules necessary for macrophage adhesion and subsequent infiltration[53, 96, 115, 124, 125]. This reduces the number of macrophages present at the focus of inflammation[53, 96, 115, 124, 125]. Also in hepatic encephalopathy, TGR5 has been shown to tone down the response of microglia and thus alleviate neurological damages[96, 126].

With its impact on metabolism via GLP-1 and T_4 signaling, TGR5 could be a valuable target for the treatment of metabolic diseases[98, 99]. Its ability to reduce the immune response in inflammatory processes could also be harnessed in inflammatory bowel diseases like

Crohn's disease[56, 96, 127, 128]. Yet, the regulation of metabolism and inflammation are not the only important physiological roles TGR5 orchestrates.

2.1.6 TGR5 in the liver

One of the physiological functions of TGR5 in the liver is the relaxation of smooth muscle cells in the gallbladder[45, 57, 113]. Thus, it promotes the filling of the gallbladder with bile[45, 57, 113]. This is the reason why TGR5 agonists might not be used in the treatment of metabolic diseases, as mice fed with TGR5 specific agonists showed extensive swelling of the gallbladder[45, 57, 113]. Additionally, the stimulation of TGR5 was shown to induce itch and analgesia in mice, so the former being mediated by the activation of transient receptor potential ankyrin 1 (TRPA1) ion channels[62, 63]. This gallbladder swelling and the itch could be very unpleasant in patients treated for metabolic or inflammatory diseases. Another very important physiological task of TGR5 is the protection of bile duct and gallbladder cells.

Bile acids exert surfactant properties which can damage the cell membrane and thus induce stress in cells and lead to cell death[98, 129]. It has been shown that hydrophobic bile acids are able to unspecifically trigger the activation of the death receptor CD-95 and cause apoptosis[130, 131]. Consequently, the sinusoidal and gallbladder endothelial cells need to possess ways to protect themselves from the cytotoxic effect of bile acids to avoid apoptosis. These are provided by several TGR5 signaling pathways.

The activation of TGR5 in cholangiocytes enhances the expression of the multidrug resistance protein 3 (MDR3), an ATPase important for the transport of phospholipids into the bile-ducts[129]. Those form mixed micelles with bile acids, which reduces the amount of free bile acids able to activate CD-95. TGR5 induces secretion of HCO_3^- and Cl^- via the anion exchanger 3 and the cystic fibrosis transmembrane conductance regulator, respectively[48, 129]. This prevents the protonation of the acid moieties of bile salts, which hampers their membrane diffusion into cells. Upon activation, the TGR5 induced production of cAMP leads to the phosphorylation and thus inactivation of the death receptor CD-95[56, 114]. TGR5 has also been shown to transactivate the endothelial growth factor receptor (EGFR), which promotes cell survival[52, 56, 96, 132]. Consequently, the bile acid-induced activation of the CD-95 receptor is counterbalanced by the bile acid-induced activation of TGR5 via anti-apoptotic and proliferative effects.

2.1.7 TGR5 in malignancies

While the protection of the sinusoidal endothelial cells and cholangiocytes is an important physiological role of TGR5, this can come with a downside. If extensive signaling

targeting the proliferative and anti-apoptotic pathways is present, this can lead to the formation of cancer from which TGR5 is not exempted[58-60, 114]. Cholangiocarcinoma cells have been found to overexpress TGR5[58]. This cholangiocarcinoma is a malignancy that exploits the TGR5 induced inactivation of CD-95 to overcome programmed cell death. A reason for this could be the CD-95 suppression of TGR5[58, 133]. TGR5 exerts the same influence in the gastric as well as the esophageal adenocarcinoma[59, 98, 134-136]. Unfortunately, the chances of surviving a cholangiocarcinoma are very low, with a five-year survival rate of only 2% at later stages[137, 138]. Currently, the therapy of the cholangiocarcinoma is difficult and usually involves partial hepatectomy and radiotherapy. Here, blocking the signaling of the overexpressed TGR5, which leads to an indefinite prolongation of the cholangiocarcinoma lifecycle, via an antagonist could present a new treatment option. This would also increase the rate of macrophage infiltration of the cancerous tissue by inhibition of the anti-inflammatory effect of TGR5. Yet, no TGR5 antagonists are known, and their design is not trivial. Small changes to the structure of an agonist can often lead to the discovery of an antagonist. But without knowing which part of the agonist is important for the activation of the receptor and which part is responsible for the affinity of the ligand, a period of tedious trial and error is required. The knowledge of a binding mode of TGR5 agonists could guide the rational design of antagonists, considerably reducing the time and resources needed for the accomplishment of this task. For the discovery of a TGR5 binding mode via integrative modeling, the in-depth knowledge of the ligand recognition of other GPCRs is required. This knowledge is important to evaluate probable binding modes and anticipate the impact of mutations on GPCR activity. Therefore, the structure and ligand binding of GPCRs in general will be discussed in the following chapters.

2.2 Structural determinants and the function of GPCRs

Upon binding of their ligand GPCRs undergo structural changes which allow the G-protein to bind, subsequently exchange GDP by GTP, which triggers the dissociation of the G-protein into the α and the βγ subunits, leading to further downstream signaling[139]. The mechanism of the ligand binding to its binding site in the GPCR, the exact structural changes, and its implications are reviewed in the following chapter. Comprehensive information about this can be found in refs. [18, 78, 140, 141].

2.2.1 The ligand recognition of GPCRs

GPCRs are a family with more than 900 different members[1, 2, 18, 78]. They can be divided into five subfamilies based on their sequence similarity and ligand recognition spectrum[1, 2,

[18, 78]. Despite this, the sequence identity within one group can be as low as 14% between different members, e.g. in class A GPCRs to which TGR5 belongs[1, 2, 18, 78, 142]. However, we know from X-ray crystal structures that, despite their low sequence identity, GPCRs exhibit a very similar fold[18, 28, 140]. Although there are structural differences between the subfamilies like big extracellular domains in class C GPCRs, all possess seven TMs and most of them a short (~10 residues) membrane-proximal intracellular helix at the C-terminus (helix 8) [18, 28, 33, 140, 143]. Those membrane spanning α-helices form a group resulting in a bundle as can be seen in **Figure 4**. The binding site is located in the upper third of the GPCR in an extracellular opening[18, 28, 140, 144, 145]. The EL2 partially covers the opening and may directly interact with the ligands in some of the GPCRs[18, 140, 144-146]. In the $S1P_1$ receptor it even blocks the ligand entry from the top so that $S1P_1$ ligands diffuse in between TMs 1 and 7[18, 140, 147]. Generally, the binding site of class A GPCRs consists of the seven TMs which form a small cavity, open towards the extracellular side[18, 28, 140, 144, 145]. The most prominent differences between the subfamilies are the kinks, bulges, and other variations of the TMs allowing for different shapes of the binding pocket[18, 28, 140, 144, 145]. This, together with the varying residues lining the binding pocket resulting in different shapes and electrostatic potentials, allows for a huge variety in the ligand recognition spectra[18, 28, 140, 144, 145]. The high structural similarity between GPCRs indicates an underlying conserved mechanism of activation between GPCRs.

Knowledge of the common ligand recognition of GPCRs considerably increases the likelihood of the identification of a binding mode model for other GPCRs via integrative modeling. During this process, several more or less likely binding modes can occur whereupon the most probable binding mode has to be selected for further validation. Here, the knowledge of common interaction patterns can be exploited to increase the success rate by discarding binding mode models that show unusual interactions between agonist and receptor.

Despite the huge varieties of ligands which are recognized by individual GPCRs, the binding poses of agonists in all GPCRs crystallized to date are very similar[18, 140, 148]. This circumstance substantiates that the activation of GPCRs always follows the same mechanism[18, 140, 148]. Hence, agonists must address certain residues within the TMs in order to activate a GPCR. If the binding of an agonist always follows a similar pattern, this can be exploited in the discovery of binding mode models as done in **publication III**. Interestingly, in all of the GPCRs, which have been crystallized so far, agonists always address residues in TM6[18, 30, 31, 33, 140, 141, 144, 145, 148-153].

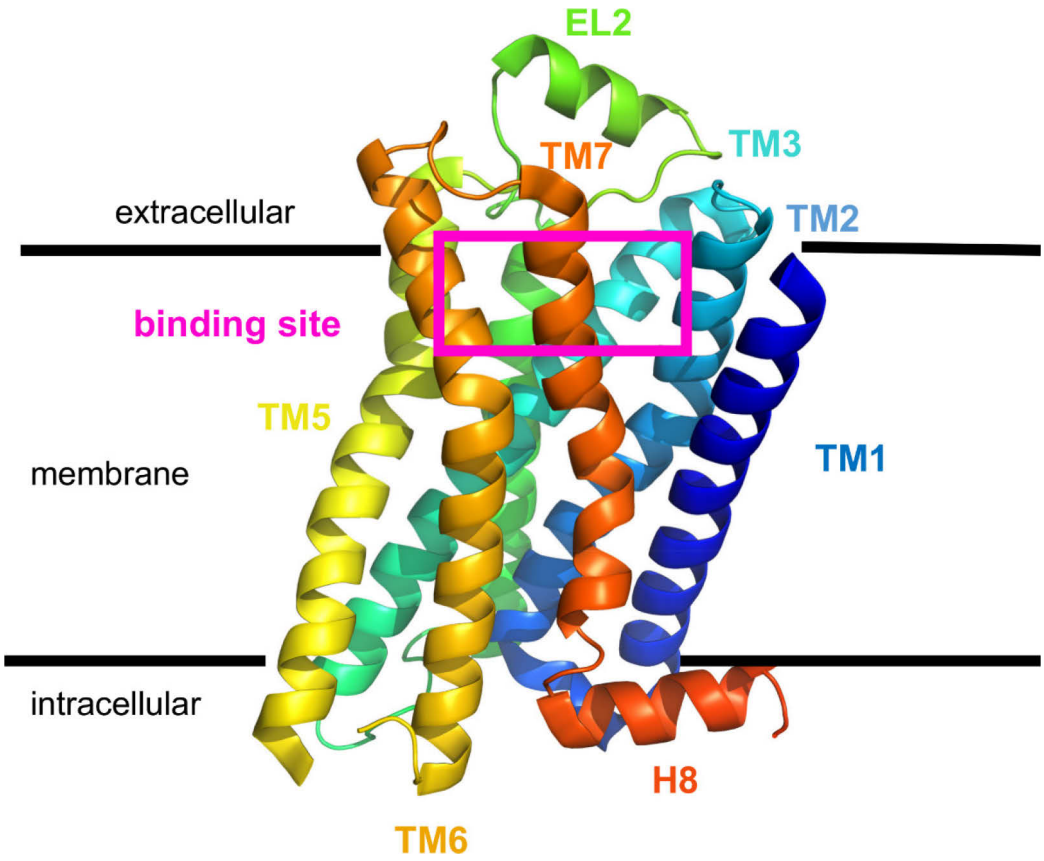


Figure 4 General structure of a class A GPCR. The inactive structure of the β_2-adrenergic receptor (PDB-ID: 3D4S) is shown in cartoon representation with rainbow coloring. The seven TMs are forming a helix bundle while their extracellular ends forming the ligand binding site. Residues of EL2 in proximity to the binding pocket can interact with bound ligands. Several helices, especially TM3, TM5, TM6, and TM7 show prominent kinks in their TMs. These kinks are conserved among the crystallized GPCRs and are a result of conserved proline residues in the positions of the kinks. Helix 8 (H8) is present in nearly all GPCR crystal structures and resides below the membrane embedding some of its residues within it. The binding site (magenta) is located at the extracellular side and is formed by the TM bundle.

For a better comparison between different GPCRs Ballesteros and Weinstein have devised a system which assigns a number to each TM residue[154]. This B&W number describes the helix and the position within this helix based on the highest conserved residue in the format X.YY where X is the helix and YY the position[154]. Thus, even though the sequence identity between GPCRs is very low and the number of residues within the loop regions can differ substantially, residues in the same position interacting with agonists can be easily identified. This is important since similar residue positions should be addressed by agonists in order to mechanically activate a GPCR provided that the underlying mechanism for activation is conserved. Indeed, all agonists which have been co-crystallized with their GPCRs so far, always address position 6.51 and very often 6.52 and 6.55[18, 30, 31, 33, 140, 141, 144, 145, 148-153, 155, 156]. For example, in the β2-adrenergic receptor agonists form hydrogen bonds with N293$^{6.55}$ and hydrophobic contacts with F289$^{6.51}$ and F290$^{6.52}$ (**Figure 5**). These interactions have been shown to be important for the activation of the receptor by mutational

analysis as mutating these residues results in a severe functional impairment of the receptors[18, 141].

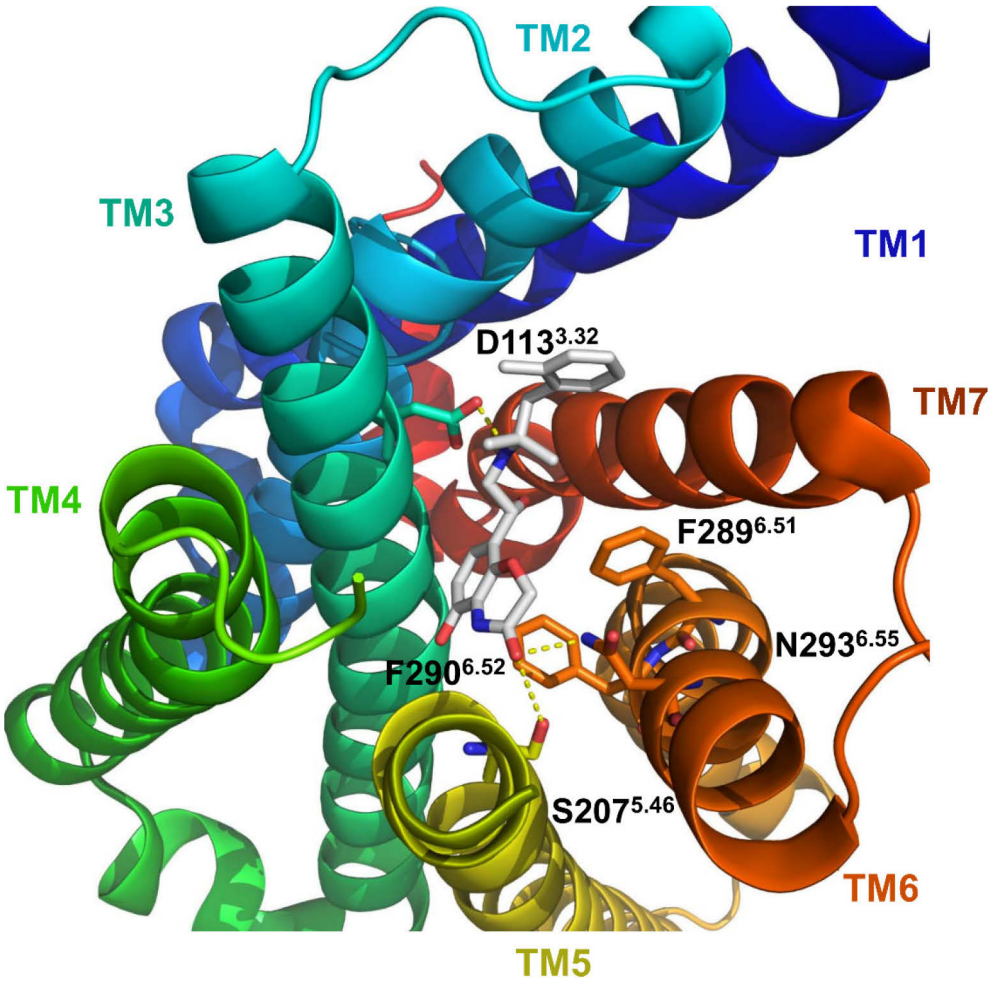


Figure 5 Ligand binding example with residues often involved in agonist binding. The active structure of the β_2-adrenergic receptor (PDB-ID: 3SN6[157]) is shown in the cartoon representation in rainbow coloring. Residues often involved in agonist binding are shown in the stick representation and labeled with their residue number as well as their B&W number, and hydrogen bonds are shown in yellow dotted lines. One of the most commonly addressed residue positions in TM6, N293$^{6.55}$, is forming a hydrogen bond to an oxygen of the co-crystallized agonist. The same oxygen forms a hydrogen bond to S207$^{5.46}$ in TM5, which stabilizes the agonist in the binding pocket. A third hydrogen bond is formed between the most commonly addressed residue position in TM3 D113$^{3.32}$ and a nitrogen in the agonist. Furthermore, phenylalanines in positions 6.51 and 6.52 form hydrophobic interactions with the agonist.

TM3 is another transmembrane helix which is always addressed by agonists in GPCR crystal structures[18, 30, 31, 33, 140, 141, 144, 145, 148-153]. The most commonly addressed residue positions inside TM3 are 3.32 to 3.37, with the most commonly addressed position in this helix being 3.36 (D113$^{3.32}$ in the β_2-AR in **Figure 5**)[18, 141]. As agonists simultaneously bind to TM3 and TM6, they essentially bridge those TMs across the binding pocket. Similar to residues in TM6, the mutation of interacting residues in TM3 hampers the ability of agonists to activate GPCRs[144, 158-163]. Other TMs partake in the binding of an agonist by stabilizing it in the binding pocket. Especially TM5 plays a pivotal role to increase the binding affinity of agonists and antagonists alike[144, 158-163]. Because of their close proximity to the vital residues

in TM6, transmembrane positions 5.42 to 5.46 can stabilize agonists in their binding conformation[18, 141], e.g. S207$^{5.46}$ in the β_2-AR in **Figure 5**. Unsurprisingly, mutating binding pocket residues in TM5 reduces the affinity of the agonists, which leads to a right shift of the dose-response curve[155, 156, 161-163]. In addition, in TM7 residues in transmembrane positions 7.39 and 7.43 stabilize the binding of an agonist in a similar fashion to TM5[18, 141].

The knowledge of the TMs and residues therein usually addressed by agonists in crystal structures is crucial for the discovery of a binding mode model via integrative modeling and has been extensively used in **publication III**. A crystal structure which is extremely valuable in the derivation of a binding mode model for TGR5 is the sphingosine-1-phosphate 1 ($S1P_1$) receptor[147]. Its close homologue, the $S1P_2$ receptor, is known to be activated by bile acids like TGR5, and the physical chemical properties and size of the $S1P_1$ agonists imitate those of TGR5[76, 147, 164]. Therefore, it is highly likely that the agonists' binding mode in TGR5 and $S1P_1$ are very similar. Indeed, as was shown in **publication III**, the binding mode model of TGR5 is nearly identical to the binding mode found in the $S1P_1$ crystal structure (**Figure 6**), confirming this hypothesis.

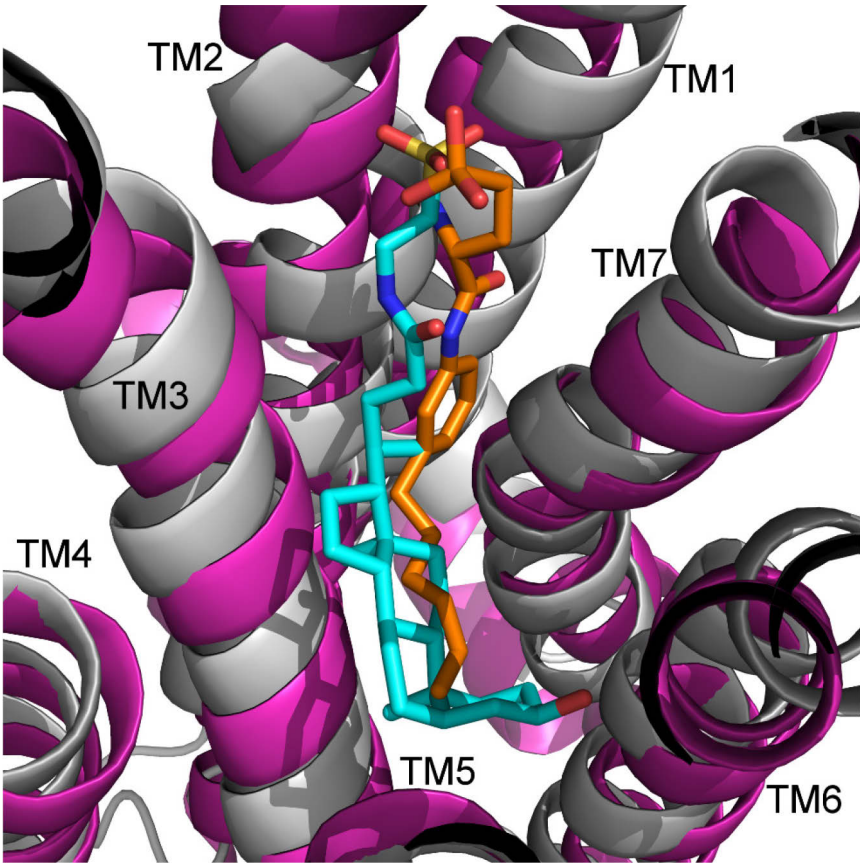


Figure 6 Comparison of the binding mode model of TLC (cyan) in TGR5 (gray) and the binding mode of a sphingolipid mimic (orange) found in the crystal structure of the $S1P_1$ receptor (magenta) (PDB-ID:3V2Y[147]). The sulfonic acid and phosphonate moieties of TLC and the sphingolipid mimic, respectively, occupy the same region, and the hydrophobic scaffolds also exhibit a high positional overlap. Text and picture adapted from **publication III**.

2.2.2 GPCR G-protein complex formation

The interactions between agonists with the commonly addressed residues mentioned in the previous paragraphs lead to characteristic structural changes in GPCRs in order to activate the G-protein.

The recent advances in crystallography and spectroscopy allow for deeper structural insight into the mechanism of GPCR activation[18]. Especially the crystallization of both the active and inactive state of rhodopsin, the A_{2A}-adenosin (A_{2A}-AR), and the β_2-adrenergic receptor, the latter even co-crystallized with a G_s-protein, was a major discovery[18, 157, 158, 165, 166]. From the differences between the active and inactive structures the reason for the importance of TM3 and TM6 can be explained (**chapter 2.2.1**). The binding of agonist to TM3 and TM6 exerts a small inwards force to the extracellular end of TM6[167]. The most prominent structural change arising from this is an outward movement of the intracellular end of TM6 by about 14 Å in a rotating, tilting movement while the extracellular end remains mainly unchanged[18, 145, 157, 158, 165, 168-170]. This movement is of utmost importance as this allows the binding of a G-protein and its subsequent activation[18, 151]. Without this movement, the binding of the G-protein is impossible, as in the inactive state TM6 occupies the binding pocket of the G-protein (**Figure 7A**).

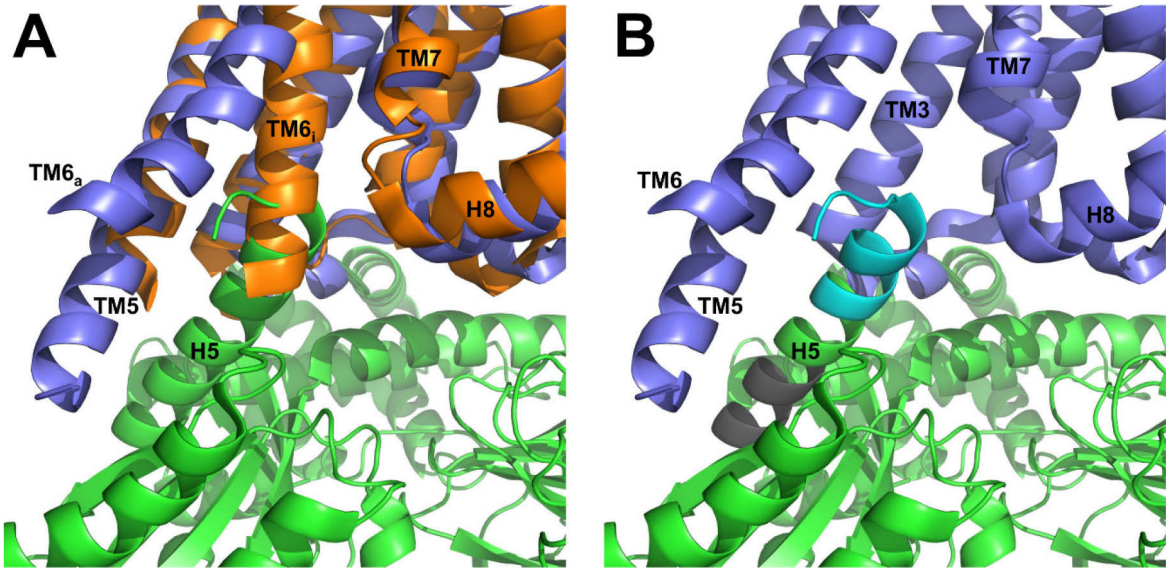


Figure 7 Overlay of the bound G-protein with the active and inactive state of the β_2-AR. The active state (navy) and the G-protein (green) have been co-crystallized (PDB-ID: 3SN6[157]), the inactive state (orange; PDB-ID: 3D4S[147]) has been aligned to the active structure. **A** TM6 in its inactive conformation ($TM6_i$) clearly clashes with helix 5 (H5) of the G-protein's α-subunit and thus prevents the G-protein from binding to the GPCR. In contrast, TM6 in the active conformation ($TM6_a$) allows for G-protein binding. **B** The G-protein binds to the active conformation of the GPCR interacting with TM3, TM5, and TM6[157]. The C-terminal end of H5 (cyan) interfaces with the GPCR while the N-terminal end (gray) activates the G-protein upon binding to the GPCR[157].

Antagonists exploit this by preventing the exertion of the force onto the intracellular end of TM6 via two mechanisms. For one, they bind into the binding pocket, while only making contact to TM3 but not TM6[18, 141]. This way, the binding pocket is occupied, such that no agonist can bind, while no force is exerted on TM6 so that no activation takes place despite a ligand being bound[18, 141, 153, 167, 171, 172]. On the other hand, antagonists bind to both TM3 and TM6 but increase the distance between those two helices, effectively acting as a strut inside the binding pocket, which also prevents the inward force onto TM6 and thus the activation of the receptor[106, 158, 163, 167].

The binding of the G-protein to the GPCR in its active state is mediated by helix 5 (H5) of the G-protein's Gα-subunit, which binds to conserved residues in TM3, TM5, and TM6[141, 157, 173-175]. H5 of the G-protein can be divided into two different sections: The C-terminal end with the task to interface with the GPCR and the N-terminal end with the task to activate the G-protein by reorganization of intra-G-protein residues (**Figure 7B**)[141, 173, 174, 176]. Considering that there are ~900 GPCRs and only 21 G-protein isoforms, the G-proteins bind quite promiscuously[177]. The residues on the side of the GPCRs, however, are diverse with the exception of a few conserved residues specific for each GPCR. This could be exploited to discover new antagonists specifically binding to the G-protein binding site of a GPCR, preventing its G-protein complex formation and activation.

In addition to interacting with G-proteins, GPCRs can bind to GPCR kinases (GRKs) and arrestins[18, 178-180]. Upon binding of GRKs, the kinases phosphorylate serine or threonine residues of the GPCRs making them susceptible to subsequent binding of β-arrestins[18, 180, 181]. The binding of β-arrestins inhibits further G-protein signaling by inhibiting the binding of a G-protein and targets GPCRs for internalization[18, 182, 183]. GPCRs which have been internalized as a result of β-arrestins' binding are subsequently trafficked to clathrin coated pits, where they are degraded[182, 183]. Yet, it has also been demonstrated that GPCR bound β-arrestins can alternatively activate signaling cascades independent of G-protein activation[183-185]. TGR5, however, is not a target of β-arrestins, and signaling from vesicles after internalization of the receptor has been shown[132]..

2.2.3 GPCR membrane trafficking

The binding of effector G-proteins is not the only important G-protein interaction of GPCRs. It is also crucial for their membrane trafficking from the endoplasmic reticulum (ER)[108, 186, 187]. After GPCRs have been expressed via the ribosome they await their plasma membrane trafficking in the membrane of the ER. Upon binding of specialized Rab-

GTPases, which are monomeric G-proteins belonging to the Ras superfamily, GPCRs are transported to the cell membrane[108, 186-189]. Interestingly, unlike the hetero-trimeric G-proteins, which are the GPCR downstream effectors, Rab-GTPases apparently do not rely on the movement of TM6 and also do not bind to the conserved residues in TM3, TM5, and TM6[108, 186-189]. Rather, GPCRs bear sorting motifs inside their C-termini, which are thought to be recognized by those Rab-GTPases enabling the GPCR membrane trafficking. They range from two to ten amino acids[107, 190-193]. The known sorting motifs can be either more hydrophilic, e.g. the DXE and the $E(X)_3LL$ motif, or hydrophobic, e.g. the LL and the $F(X)_3F(X)_3F$ motif, with the $F(X)_6LL$ motif being the most prominent motif among GPCRs[107, 190-193]. The diversity of the sorting motifs and the high number of unspecified residues (X) indicate a structure, rather than a sequence dependent mechanism for GPCR membrane localization. Additionally, many GPCRs such as TGR5 do not contain any of the known sorting motifs[194]. This raises the question what the determinant of the membrane localization of those receptors is. Intriguingly, nearly all crystallized GPCRs with the exception of the CXCR4 receptor show an α-helix formation in their C-terminus[106]. Therefore, Rab-GTPases could recognize the α-helix in the C-terminus of GPCRs and subsequently traffic them to the plasma membrane. In this case, it is possible that the sorting motifs facilitate α-helix formation of the C-terminus rather than being directly recognized by the Rab-GTPases. An indicator for this is that the longer motifs contain hydrophobic residues in *n*+4 positions, most prominently the $F(X)_3F(X)_3F$ motif. In α-helices 3.6 residues are needed for one turn, which would place the hydrophobic residues of the sorting motifs on one side of the helix[195]. The hydrophobic residues could then act as membrane anchors of the C-terminus, which interacts with the membrane (**Figure 4**). This would facilitate α-helix formation of the C-terminus. The determinants of membrane localization in TGR5 have been uncovered in **publication I.**

2.2.4 GPCR dimer and oligomer formation

GPCRs have been found to form homo- and heterodimers up to higher-order oligomeric arrays[18, 196-204]. The dimerization of GPCRs can have a strong influence on their activity in several ways. For several receptors, it has been shown that in the ER homo- or heterodimerization is necessary for the plasma membrane trafficking of the receptor[18, 205-208]. In the cell membrane, GPCR dimerization can have a profound impact on their activity. For several GPCR heterodimers it was shown that the inhibition of one of the receptors with an antagonist led to the inhibition of the other receptor of the heterodimer[18, 160, 209]. Thus, both receptors of a heterodimer were hampered in their function upon the inactivation of

one. Similar effects have also been observed for the propagation of agonistic effects across heterodimers, which was shown to expand the downstream signaling repertoire of GPCRs[18, 210-213]. Astonishingly, some receptors such as the dopamine D_2 receptor only function in homodimers. It was shown that the activation of only one D_2 receptor protomer with an agonist activated both receptors of the dimer via an allosteric mechanism[18, 214]. The protomers of GPCR homo- and heterodimers have recently been targeted simultaneously by bivalent ligands reaching from the binding site of one protomer to the other[18, 214-216]. The advantage of this is the possibility to shape the response by targeting several different receptors simultaneously with one ligand, or potentially increasing the affinity of a ligand towards a homodimer[18, 214-216]. However, even after the activation of GPCRs homo- and heterodimerization still impacts their fate. Some GPCRs such as the $P2Y_{11}$ receptor have shown a requirement of heterodimerization for receptor internalization[18, 217-220]. As TGR5 internalization has been shown to be independent of β-arrestin binding[132], homo- or heterodimerization may play a role, too. Hitherto, the dimerization interfaces of TGR5 are unknown, the knowledge of which could help to understand its function and internalization. As crystal structures of GPCR dimers are already known, it is possible to infer likely dimerization interfaces of TGR5 from those. Until now, three different interfaces have been identified in GPCR crystals.

In the crystal structure of the κ-Opioid receptor, the protomers interface via TM1 and helix 8 (the 1-8 interface, **Figure 8A**)[221, 222]. In this interface TM6 can move unobstructed as it is not part of the interface, leaving the activation of a GPCR unimpaired. The contact area between the protomers is with 615 Å² per protomer quite low as only helices 8 and the extracellular ends of TM1 are interacting with one another[222]. Hence, dimerization involving this interface is expected to be less stable[215, 223].

In the crystal structure of the CXCR4 receptor, the protomers interface via TM4 and TM5 (the 4-5 interface, **Figure 8B**) [222]. Also in this interface, TM6 can move unobstructed, not impairing the activation of GPCRs. Compared to the 1-8 interface the contact surface between the protomers is with 784 Å² about 30% larger[222]. However, as has been shown in the β_2-adrenergic and μ-opioid receptor, TM4 and TM5 form a shallow recess to which cholesterol preferably binds[196, 224-226]. The cholesterol can act as a facilitator for dimerization, rendering this dimerization interface stronger than judging from the protein-protein interaction alone[224, 225]. Interestingly, in this interface the tyrosine residue of the conserved (D/E)RY motif in TM3 could mediate the dimerization due to its proximity to the interface.

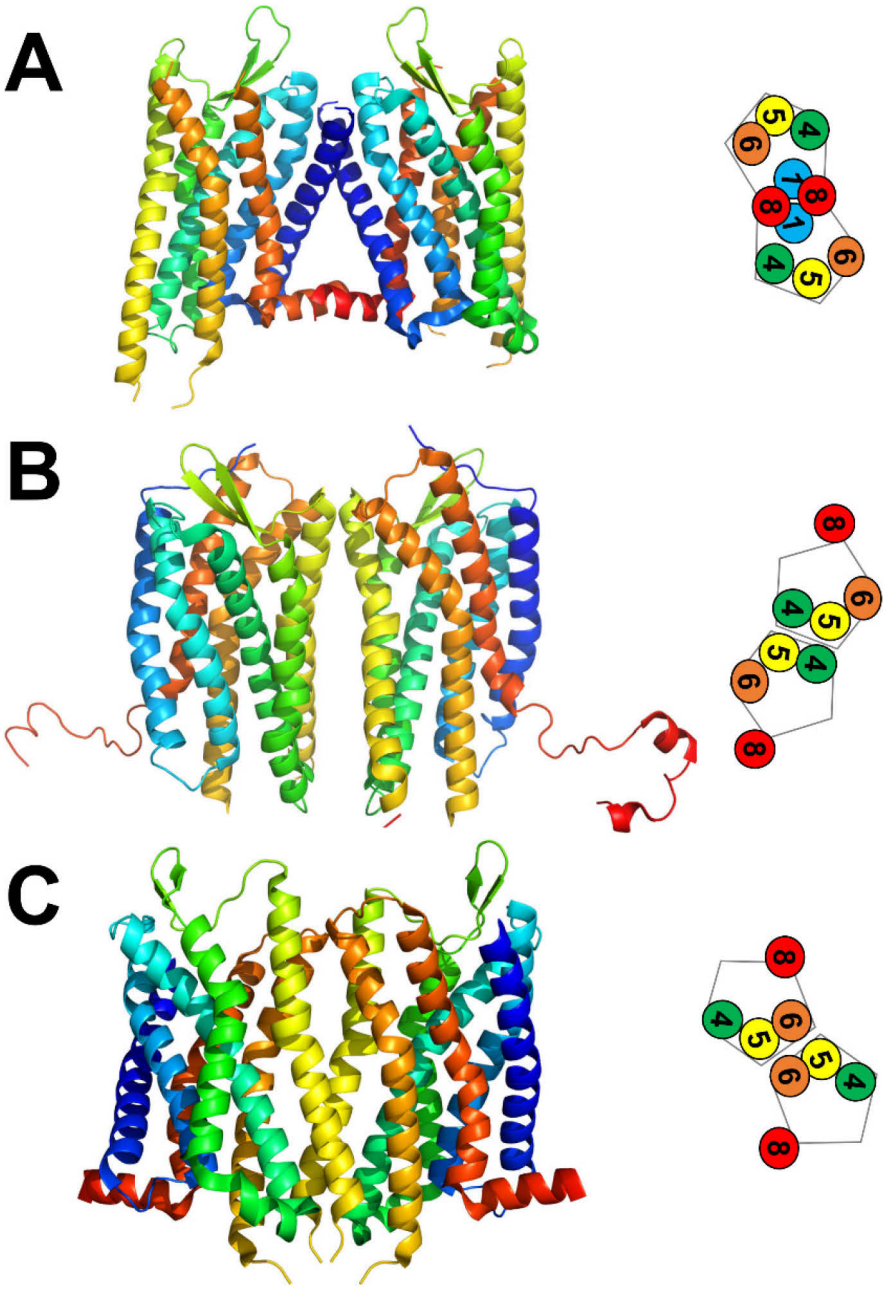


Figure 8 Structures of known GPCR dimers. Shown are two protomers in cartoon representation with rainbow coloring. On the right are schematics indicating the orientation of the protomers and the dimerization interface. **A** The 1-8 interface as found in the κ-Opioid receptor (PDB-ID: 4DJH). The contact areas are mainly between the extracellular ends of TM1 and helix 8. **B** The 4-5 interface as found in the CXCR4 receptor (PDB-ID: 3ODU). The CXCR4 receptor is one of the few GPCR crystal structures in which helix 8 shows loop formation. The gap between the protomers is much wider than in the μ-Opioid receptor. However, the contact could be mediated by cholesterol when embedded in a biological membrane. **C** The 5-6 interface as found in the μ-Opioid receptor (PDB-ID: 4DKL). The contact area in the 5-6 interface is about twice as big as in the other interfaces. Yet, as TM6 directly interacts with TM5 of the other protomer, the outward movement of TM6, which is necessary for the activation of the receptor, is blocked. Figure adapted from **publication II**.

Until now, no reason for the conservation of this tyrosine among nearly all GPCRs is known, as it points into the membrane and plays no obvious role in the activation mechanism of a single GPCR. Yet, promoting the dimerization of GPCRs, which could facilitate cross-activation of homodimers, could be an explanation for the presence of this tyrosine in TM3.

In the crystal structure of the μ-Opioid receptor[222], the protomers interface via TM5 and TM6 (the 5-6 interface, **Figure 8C**) [222]. As the inactive state of the μ-Opioid receptor was crystallized in this dimer interface, it is unclear whether the protomers are able to be activated in this orientation[222]. The reason for this is that as an outward movement of TM6 is blocked by the other protomer, the authors conclude. Hence, it is unclear whether dimers adopting this interface could be activated without rearrangement of their mutual positions. However, the 5-6 interface has the highest contact area between the protomers found so far which indicates a high stability of this interface[222]. With 1492 Å² contact area per protomer it possesses around twice the contact area of other interfaces[222]. Despite the possible hindrance of the activation mechanism in the 5-6 interface, the high contact area indicates that the 5-6 interface could be the most abundant interface found in cells[215, 222, 223].

In this context, it has to be taken into account, that an antagonist has been co-crystallized with the crystal structure of this GPCR dimer. As mentioned earlier, antagonists can propagate their inhibitory effect across GPCR dimers via an allosteric effect. However, the exact mechanism of the allosteric inhibition is unknown. Possibly, the binding of antagonists induces a rearrangement of the protomers, forcing them to adopt the 5-6 interface. In this arrangement, both protomers would be unable to be activated, as the outward movement of TM6 is impossible due to steric hindrance. Thus, the second protomer would be inhibited despite no antagonist presence in its binding pocket. This is only a hypothesis and has to be investigated more thoroughly.

Rhodopsin has been shown to form higher-order oligomers, and other GPCRs are also expected to show oligomerization, which could be mediated by chaperones[201, 202, 227-231]. While the exact orientation of the GPCRs in oligomers remains elusive, it is thought that oligomers are composed of higher-order dimers of dimers, resulting in oligomers[201, 202, 227-230]. Whether GPCRs are always organized in oligomers or whether the oligomerization only plays a role in some GPCRs is not known, so far. For the formation of oligomers from dimers several interface combinations are possible. Combining interfaces 4-5 and 5-6 would be impossible due to their proximity. Yet, combining either of those interfaces with the 1-8 interface results in tetramers, which could be extended (**Figure 9**). Identifying the dimerization interfaces of a receptor may help uncover possible oligomerization states of this receptor. For this, the distances between the C-termini of the protomers could be measured e.g. via Förster resonance energy transfer (FRET) spectroscopy[232]. The experimentally measured distance can then be used to discern between several computational dimerization models. Here, an accurate model is required to discriminate dimer interfaces

where the expected distance between the C-termini is similar, e.g. the 4-5 and 5-6 interface (see **publication II**).

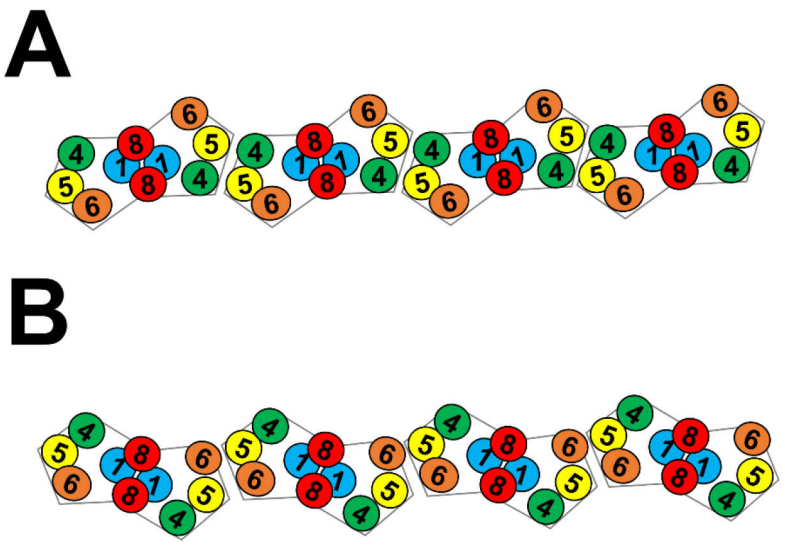


Figure 9 Schematic view of possible GPCR oligomers. **A** Extracellular view on GPCR oligomers using the interfaces 1-8 and 4-5 to form oligomeric arrays. **B** Extracellular view on GPCR oligomers using the interfaces 1-8 and 5-6 to form oligomeric arrays. Figure adapted from **publication III**.

2.3 Integrative modeling

Comprehensive information on integrative modeling can be found in ref. [233].

In integrative modeling, experimental alongside theoretical information is used to build and improve structural models of macromolecules[233]. The advantage of this approach is that the structure of proteins or other polymers can be elucidated that are hard to crystallize because of their size, solubility or other reasons[233-236]. Often, experimental or theoretical information is used to select the structural model that adheres to most of this data[233-236]. X-ray crystallography is part of the integrative modeling process as models of a protein with a known sequence are modeled into electron density maps[233]. Thus, structural models are built with the help of a high amount of experimental data in the tens of experimental observations per heavy atom of the macromolecule[233]. In addition, integrative modeling comprises the use of many other experimental methods in order to refine or gain insight into different aspects of the structure[233, 237, 238]. For example, hydrogen-deuterium exchange with subsequent mass spectrometry can be used to detect the solvent accessible surface area of a protein[233, 239, 240]. The binding site of a protein can be mapped with NMR spectroscopy, FRET, and, as extensively done in **publication III**, by mutating residues and subsequently characterizing the functional consequences[233, 241, 242]. This information can be used to gain atomistic insight into the binding of a ligand to its receptor via homology modeling and docking[233, 241, 242]. Physical proximity between several (macro-)molecules can be determined via co-

precipitation[233, 243, 244]. A more sophisticated method for determining interactions is FRET. Here, not only an apparent distance can be inferred from several fluorescence parameters, but FRET is also capable of obtaining time resolved data[233]. Thus, changes over time can be studied to investigate complex processes in cells[233]. Dynamic processes of molecules can also be assessed with MD simulations, which can be used to interpret the data gained from FRET spectroscopy[233].

2.3.1 Förster resonance energy transfer

FRET spectroscopy is a physical method which makes use of the energy transfer from one fluorophore to another via the emission and absorption of light[245, 246]. The donor fluorophore is excited with a laser using the excitation wavelength of the donor, which upon falling back into the ground state emits light with a lower energy and thus a higher wavelength than the excitation light[245, 246]. The acceptor fluorophore is chosen such that the emission spectrum of the donor corresponds to the excitation spectrum of the acceptor[245, 246]. Thus, the acceptor absorbs the light emitted by the donor and emits at an even higher wavelength[245, 246]. The more donor emissions are absorbed by the acceptor fluorophore, the less donor and the more acceptor emissions can be detected[245-248]. Hence, the FRET efficiency is defined as the ratio of the donor-to-acceptor emission intensity[245-248]. The FRET efficiency depends on the distance between the donor and the acceptor, with a smaller distance resulting in a higher efficiency[245-248]. The lifetime of the donor fluorescence $\varepsilon(t)$ also depends on the FRET efficiency and can be used to calculate apparent distances $R_{DA,app}$[247, 248]. This can be exploited to determine the distance within one or between several molecules by strategic attachment of fluorophores[245-248]. This approach can be applied in a variety of different scenarios in order to answer biological questions. In its most rudimentary form, FRET spectroscopy can determine whether two (macro-)molecules are interacting measuring the presence or absence of FRET. In its more sophisticated forms, FRET spectroscopy is able to accurately track distances over timescales from ns to hours[231, 233]. Here, fluorescent labeled ligands can be used for competition assays and to determine oligomerization of proteins[231]. Thus, FRET spectroscopy is a valuable tool in integrative modeling as it is able to provide distance restraints for structural modeling or to guide conformational selection in a large structural ensemble[245-247].

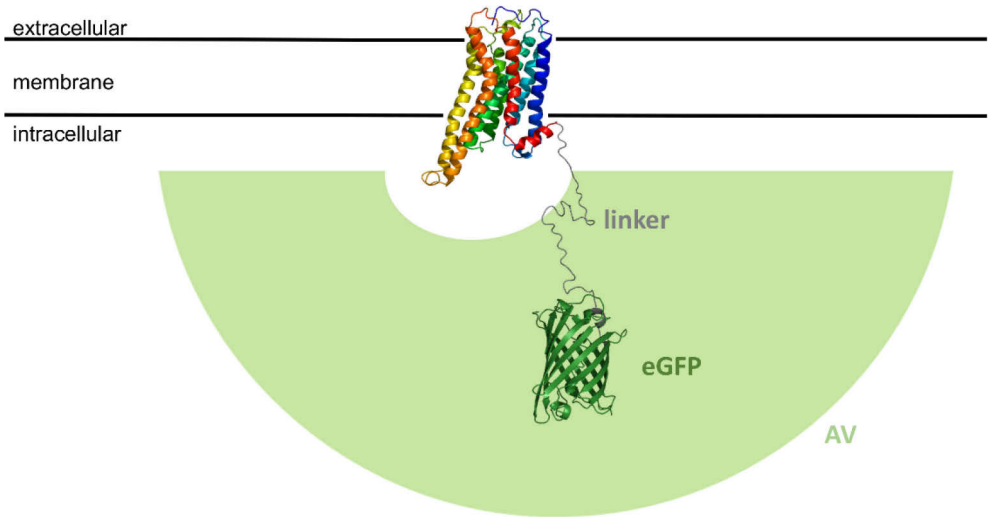


Figure 10 Schematic view of the accessible volume (AV) of a fluorescent dye (eGFP). eGFP (dark green), which is attached to the C-terminus of a GPCR (rainbow) by a linker (dark gray), can move inside a limited area, the AV (light green). The size of the AV is restricted by the shape of the protein, the average length of the linker determined by Gaussian chain approximation[249], and, as the fluorophore is located inside the β-barrel in eGFP, the distance from the fluorophore to the edges of the β-barrel. Because of this, the fluorophore cannot occupy the space close to the membrane or the receptor.

The FRET efficiency depends on the orientation and distance of the dyes[245-247]. As the dyes are moving, many donor-acceptor pairs in different states are measured simultaneously yielding a distribution of efficiencies[245-247]. This hampers the accurate prediction of the correct distance between the dye anchor points[245-247]. However, the positioning of the dyes can be simulated from which a theoretical FRET efficiency distribution can be calculated[245-247]. The comparison of the experimentally measured efficiency distribution and the theoretical distribution then allows the accurate measurement of the distance between the attachment points of the fluorophores[245-247]. Here, the accessible volume (AV) of the dyes is calculated via a Monte Carlo approach, which provides a probability of the dye position in space at any given point of time (**Figure 10**), from which the efficiency distribution is then calculated[245-247]. This is a very fast method and is therefore ideally used on a large structural ensemble generated by MD simulations. Thus, the structures with a theoretical efficiency distribution consistent with the experimental distribution can be identified[247, 250]. The downside of the use of AV simulations is that in those simulations the complete AV of the dye is simulated regardless of possible dye-surface interactions and preferred linker conformations. To compensate this, a Gaussian chain model approximation can be employed. Yet, especially with long peptide linkers, as used in **publication II**, this can lead to discrepancies, which hamper an accurate prediction.

To counteract this, the behavior of the dye and the linker can be simulated with all-atom explicit solvent MD simulations. This method excels at predicting the secondary structure formation of the linker. The high accuracy of this method was successfully used in

publication I to predict the secondary structure formation of the membrane-proximal C-terminus of TGR5. Therefore, I applied this method to determine the preferred conformations of the peptide linker and fluorescent dye in **publication II**. The mean length of the linker could be determined more accurately by all-atom MD simulations compared to AV simulations. However, MD simulations were computationally much more expensive than the AV simulations. MD simulations allowed the identification of a dimerization interface of TGR5 when combined with TGR5 dimer models based on the interfaces reviewed in **chapter 2.2.4**.

2.3.2 Homology modeling

In homology modeling structural models of proteins are created based on the X-ray crystal structures of homologues proteins according to an alignment of their sequences[251-253]. Homology modeling exploits the fact that structure is higher conserved than sequence[251-255]. Because of this, two proteins from the same family usually exhibit a very similar fold although the identity of their sequences might be as low as 20%[251-255]. This is why homology modeling is able to create rather accurate models of proteins as long as crystal structures of homologues proteins are available[251-255]. The sequence of the target protein, which is to be modeled, and the template protein, on which the target is modeled, are aligned according to the similarity of their residues[251-253]. The higher the exact match, i.e. the identity, between the corresponding residues, the easier is the correct alignment of the residues resulting in accurate models[251-253]. Hence, the alignment of the sequences is the most crucial step in homology modeling as an incorrect alignment leads to an incorrectly modeled structure. The sequence alignment is then used as positional restraints in the modeling process so that the backbone of the residues of the target are modeled in the same position of the corresponding template residues as determined by the alignment (**Figure 11**) [251-253]. The side chains are subsequently added to the backbone from a rotamer library[251-253, 256], and the model is subjected to refinement including relaxation via MD simulation[251-253]. Modeling the target based on multiple templates at once instead of a single template, i.e. a multi-template modeling approach, usually results in more accurate homology models[257]. In multi-template modeling, the structural diversity of the templates can be taken into account, and the use of many constraints per residue increases the amount of experimental data included in the modeling process[257].

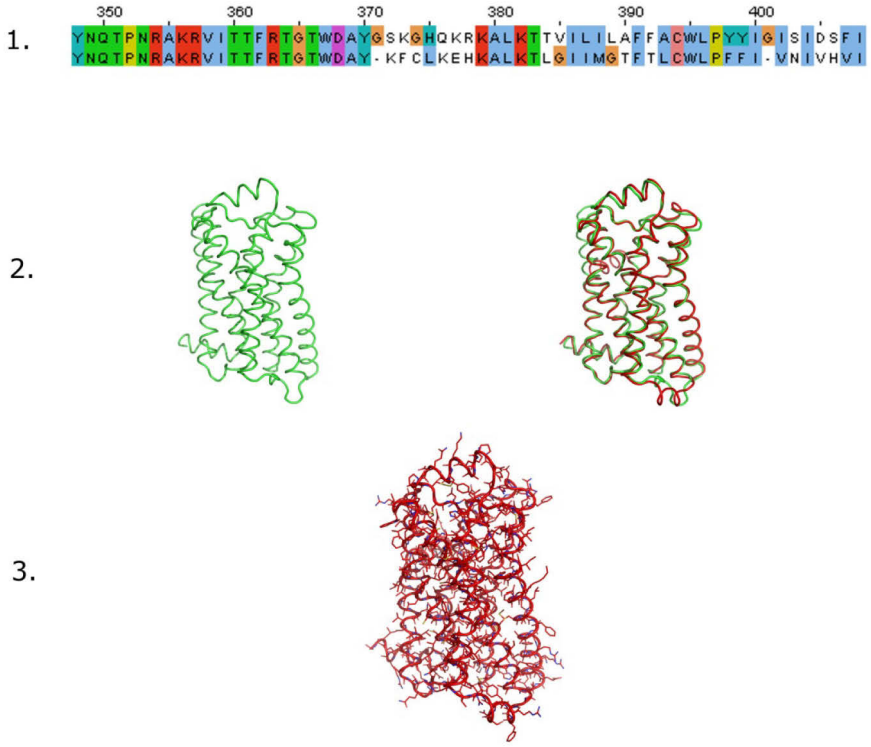


Figure 11 Steps of homology modeling using a single template. **1.** Sequence alignment of the template and target protein, residues of similar physicochemical properties are color coded in the alignment. The sequence alignment is used for positional restraints for the residues in step two. **2.** Based on the structure of the template (green) and the information obtained from the sequence alignment the backbone of the target (red) is modeled. **3.** The side chains of the target are added to the backbone, according to a rotamer-library, in an energetically favorable position. Figure adapted from ref. [258].

All class A GPCRs, which have been crystallized so far, exhibit a very similar fold (**chapter 2.2.1**) with the biggest variations in the EL2 although their sequence identity is as low as 20%, which makes it possible to model other class A GPCRs[18, 79, 87, 88, 104, 105]. The low sequence identity between the GPCRs hypothetically makes GPCRs hard targets, especially because the length of the loops in between the TMs can differ immensely[18, 79, 87, 88, 104, 105]. However, GPCRs possess conserved motifs within their TMs that can guide the alignment of the sequences to overcome this limitation.

In TM3, for example, the (D/E)RY motif is conserved among nearly all GPCRs[18, 104, 105]. Another microswitch in the GPCR superfamily is the tyrosine in the conserved NPXXY motif at the end of TM7, which is, except for the proline residue, not present in TGR5[18, 104, 105]. This absence made the modeling of TGR5 more difficult as it gave several possible alignments of TM7, which had to be tested in **publication III**.

An additional conserved motif is the CWXP motif in TM6[18, 104, 105]. The other TMs contain more or less conserved residues instead of motifs[18]. In TM1 an asparagine[1.50] is conserved among all GPCRs, in TM2 reside a conserved aspartic acid[2.50] and glycine[2.54], in TM4 a conserved tryptophane[4.50], and TM5 a conserved proline[5.50] [18, 104, 105].

The presence of conserved residues or even motifs in every TM in GPCRs simplifies a correct alignment of the TMs. Unsurprisingly, in recent modeling competitions (GPCR Dock) the structure of TMs was correctly predicted within 1 Å root mean square deviation (RMSD), which is well within the experimental error of a crystal structure[79, 80, 87]. For these competitions, novel crystal structures of GPCRs were temporarily withheld so that the structure and the binding mode of a co-crystallized ligand could be predicted[79, 80, 87]. Thus, an objective determination of the best performing strategies and attainable accuracy could be achieved[79, 80, 87]. With a high modeling accuracy of the TMs, only the structurally diverse loops remain a difficult target in modeling GPCRs[79, 80, 87]. The TMs constitute the majority of the binding pocket (**chapter 2.2.1**) so the loops pose a minor problem when predicting binding modes of ligands in homology models of GPCRs[79, 80, 87]. The exact binding pocket conformation of all residues remains elusive despite high modeling accuracy of the TM regions[79, 80, 87]. This can be overcome by conducting binding mode prediction in a variety of homology models covering a range of different binding pocket conformations[79, 80, 87, 259]. Consequently, the binding mode of a ligand with up to 82% correct interactions between ligand and receptor could be predicted[80].

2.3.3 Molecular docking

The prediction of unknown binding modes is usually done with molecular docking approaches. The aim of molecular docking is to identify the energetically most favorable ligand conformation inside the binding pocket[260-262]. However, this poses a challenge of global optimization, as several local minima can be present, which do not represent the binding mode[260]. Here, it is problematic to identify the global minimum among the local minima, and to actually sample the global minimum[260-262]. The solution to the latter problem is to cover the energy landscape as thoroughly and rapidly as possible[260]. For this, many different algorithms can be employed. For example, genetic algorithms capable of rapidly identifying local minima can be used from different starting points on the energy landscape[260-262]. As many energy minima throughout the energy landscape are identified, the global minimum is found with a high chance. One of those algorithms is the Lamarckian genetic algorithm[261, 262]. Here, the ligand is translated and rotated inside the proposed binding pocket while conformers of the ligand are generated to create a set of possible binding poses[261, 262]. These are then evaluated with a scoring function, and the energetically best combinations of translation, rotation, and conformation found are used to generate a new set of possible binding poses [261, 262]. This is done until a local energy minimum is reached, then the process is restarted, eventually finding another or the same energy minimum[261, 262]. After

multiple repetitions the most populated binding pose with at least 20% of all conformations is usually considered to be a valid solution. The energy evaluation of the binding poses can be achieved in three main ways. Force fields, as used in MD simulations, empirical scoring functions, often taking into account the solvent accessible surface area, and knowledge-based scoring functions, deriving an energy term from a statistical assessment of interactions found in crystal structures[81-86, 263-265]. Molecular docking approaches have proven to be valuable tools for the binding mode prediction of ligands among a variety of target systems[79, 80, 87, 94, 266-270].

Although the scoring functions are often able to predict the correct binding mode of a ligand, an energetically favorable binding pose does not necessarily resemble the true binding mode. One of the reasons is the influence of the binding pocket conformation, where small differences can highly impact the outcome of the docking. Following the integrative modeling approach, the binding mode identification should be bolstered with as much information as possible[233]. Viable binding poses interact with residues identified to be crucial for receptor activation (see **chapter 2.2.1**). Unsurprisingly, groups utilizing as much of this information as possible generally outperformed others in the GPCR Dock competition[79, 80, 87]. Another way to imbue the binding mode prediction with information is the mutation of residues and subsequent experimental characterization. This can be accomplished with methods such as cAMP reporter gene assays in the case of GPCRs and other proteins that lead to the production of cAMP, or radioligand assays[271-274]. The disadvantage of the radioligand assay is that it is costly and requires special laboratories, but it is able to accurately capture influences on ligand binding. The disadvantage of the cAMP reporter gene assay is that a reduction in activity upon mutation does not necessarily occur due to a worse binding of the ligand, as it can identify residues that are important for receptor activation independent of ligand binding.

The importance of the correct interpretation and thorough acquisition of information becomes apparent if looking at mutations in the aforementioned CWXP motif in TM6 (**chapter 2.3.2**). The tryptophan residue in this motif lines the bottom of the binding pocket and is considered to be an essential switch for the activation of GPCRs but mostly does not interact with agonists[18]. However, upon mutating this residue resulting in severe impairment of the receptor, and subsequent functional readout it is often misinterpreted to be an essential interacting residue[275]. Yet, careful consideration of the available theoretical and experimental data and its incorporation into the modeling process can lead to highly accurate binding mode models.

2.3.4 AFMoC analysis

Another way to implement experimental data into binding mode prediction is the use of structure-based 3D-QSAR approaches such as Adaptation of fields for molecular comparison (AFMoC). A detailed overview of the AFMoC methodology can be found in refs. [276-278]. AFMoC is usually used to calculate the structure-activity relationships of a set of ligands resulting in Stdev*Coeff maps, which show the favorability of the presence of a specific ligand atom type at positions inside a binding pocket[276-278]. These maps can be interpreted to guide the development of more potent ligands[279-281]. In the AFMoC approach, the individual interactions between a receptor and a set of ligands are correlated to the experimental bioactivity of each ligand to derive a structure-activity relationship model for this ligand set[276]. The atom type specific interaction fields are calculated by multiplying knowledge-based potential fields, distance-dependent functions to evaluate the positioning of a given ligand atom within a binding pocket, with ligand atom probes represented as a Gaussian function (**Figure 12**)[276-278]. For incorporating the information about the structural environment of the ligands, DrugScore pair-potentials are used to calculate the potential fields[276-278]. AFMoC requires the ligands to be structurally aligned with respect to a bioactive conformation and energetically relaxed inside a binding pocket for the calculation of the interaction fields[276-278]. The quality of the AFMoC analysis is heavily dependent on the structural alignment of the ligands in the binding pocket[276-278]. This quality can be measured from the predictivity of the AFMoC model[276-278]. Here each ligand of the dataset is left out once and a QSAR model is generated using the remaining ligands[276-278, 282, 283].

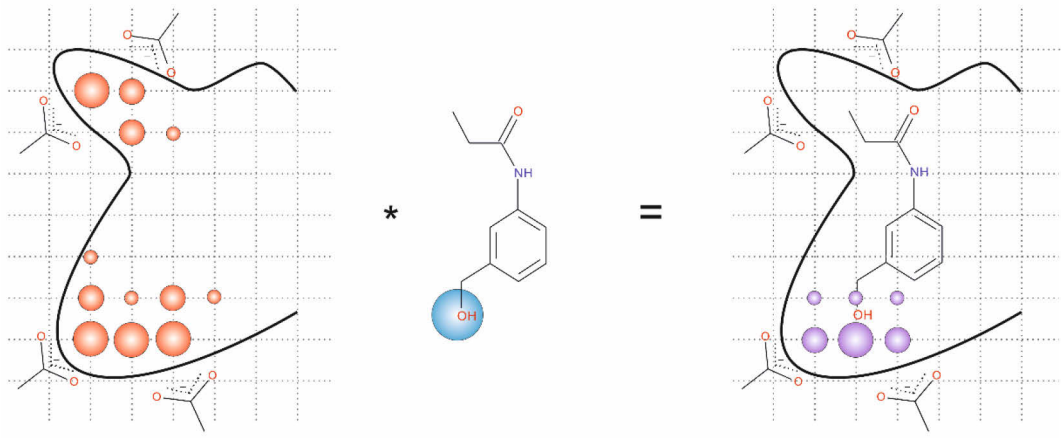


Figure 12 Example for the calculation of AFMoC interaction fields. The DrugScore potential fields (red) inside the binding pocket for the favorable presence of sp^3-hybridized oxygen are convoluted with the ligand atom probe (blue) to result in interaction fields (violet) for this atom type. Figure adapted from ref. [277] with permission (see **REPRINT PERMISSIONS**).

The bioactivity of the left-out ligand is then predicted and compared to the experimentally determined data. This results in the q^2-value, a cross-validated r^2, which is an estimate for

the predictivity of the QSAR model[276-278, 282, 283]. The q^2 can be also seen as an estimate of how well the ligand interactions with the receptor reflect their biological activity[276-278, 282, 283]. In combination with a high sensitivity towards the correct structural alignment of the ligands, as mentioned in the previous paragraph, this can be used to evaluate binding mode models. Only models in which the presence or absence of an interaction is reflected in a higher or lower potency of the ligand will result in a significant binding mode model. Thus, the validity of a binding mode model can be judged by its q^2-value, as employed in **publication III**.

3 SCOPE OF THE THESIS

GPCRs are a diverse protein family with about 900 different members and are among the most important targets on the current drug market. Despite their importance, little is known about their di- and oligomerization, and membrane trafficking. For the latter, several sorting motifs in the C-terminal helix 8 of GPCRs are known but their recognition and function is not well understood. Some GPCRs as TGR5 do not even possess any known sorting motifs in their C-terminus (**chapter 2.2.3**), which raises the question what are the determining factors for their membrane localization. TGR5 is the first known bile acid-sensing GPCR and is involved in many mechanisms controlling energy homeostasis and inflammation in the body (**chapters 2.1.4, 2.1.5**). It is therefore considered a significant factor in the formation of the cholangiocarcinoma, and the gastric and esophageal adenocarcinoma (**chapter 2.1.7**). Hence, it is an attractive target for the development of antagonists for cancer therapy and for the development of agonists for metabolic and inflammatory control. An accurate binding mode model of TGR5 agonists, which is unknown at present, could direct the development of more potent and selective agonists as well as antagonists. A binding mode model of TGR5 agonists might also explain the epimeric selectivity of TGR5 towards bile acid agonists with a 7-hydroxyl group in α-position (**chapter 2.1.2**). Furthermore, many GPCRs are known to dimerize through which agonistic and antagonistic effects can be transmitted across protomers. However, it is currently unknown which dimerization interfaces TGR5 prefers upon di- or oligomerization.

This poses the following questions:

- What are the determinants of TGR5 membrane localization if no known sorting motif is present in its C-terminus?
- Which dimerization interfaces does TGR5 accommodate?
- What is the binding mode of TGR5 agonists and how is the epimeric selectivity mediated?

Those questions have been addressed in this thesis, which led to the following publications.

4 PUBLICATION I - A Membrane-proximal, C-terminal α-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5

Lina Spomer[§], Christoph G.W. Gertzen[§], Birte Schmitz, Dieter Häussinger, Holger Gohlke, and Verena Keitel

[§] Both authors contributed equally to this work.

J. Biol. Chem. (**2014**), 289, 3689-3702.

Original publication, see pages 54-69; contribution: 30%

4.1 Background

In order to exert their function, GPCRs have to be trafficked from the ER, where they are synthetized, to the membrane of the cell. Many GPCRs display sorting motifs in their C-termini, which are recognized by Rab-GTPases transporting the GPCRs to the cell membrane (**chapter 2.2.3**). These sorting motifs range from the DXE motif to the $E(X)_3LL$, $F(X)_3F(X)_3F$, and $F(X)_6LL$ motif, which are quite diverse. The way those sorting motifs are recognized by the limited number of Rab-GTPases is currently unknown. What is more, several GPCRs such as TGR5 do not possess known sorting motifs. This leaves two possibilities: Either the number of sorting motifs is far greater than previously assumed, or the known sorting motifs merely facilitate the adoption of a specific secondary structure, which does not necessarily require their presence. Discovering the determinants of the membrane localization of a GPCR without a known sorting motif as TGR5 could help to understand how GPCRs are recognized by Rab-GTPases.

Here we set out to determine whether variants of the TGR5 C-terminus, which lead to membrane localization of the receptor, specifically adopt an α-helical fold, as seen in the majority of GPCR crystal structures, opposed to the variants retained in the ER. Additionally, we analyzed chimeras of TGR5 with the C-termini of other GPCRs for which an α-helical fold has been confirmed in crystal structures.

4.2 Determinants of the membrane localization of TGR5

The naturally occurring truncation variant Q296X of the C-terminus identified in the lab of Prof. Dr. V. Keitel leads to a retention of TGR5 in the ER. In comparison to the wildtype protein this Q296X variant failed to activate adenylate cyclase after stimulation with the

agonistic bile acid TLC. Similar findings, i.e. a reduced functionality and retention in the ER, have been demonstrated for other GPCRs with a truncated membrane-proximal, intracellular C-terminus, such as the luteinizing hormone/chorionic gonadotropin receptor[284], the vasopressin 2 receptor (V2R)[285], and the A1 adenosine receptor[286]. This indicates that the cellular signaling response is determined by the amount of functionally active receptors in the plasma membrane. Based on the naturally occurring TGR5 truncation mutation Q296X and guided by my secondary structure predictions from MD simulations, we generated nine deletion and substitution variants within the membrane-proximal C-terminus to identify the amino acid motifs/structural determinants that facilitate plasma membrane localization of this bile acid receptor. Using these variants and three chimeras of TGR5 with the membrane-proximal C-terminus of the β_2-AR, the $S1P_1$, or the kappa-type opioid receptor (κ-OR), respectively, we demonstrate that the formation of a membrane-proximal α-helix (helix 8) is essential for anterograde trafficking of TGR5 from the ER to the PM and thus for receptor function.

Immunofluorescence staining of the truncation variants D284X and R297X showed a reticular, intracellular fluorescence pattern (**Figure 13A**), which was identified as the ER by double-labeling with an antibody against the ER marker protein disulfide isomerase. The S310X variant was mainly localized in the PM (**Figure 13A**). Truncation at amino acid 297 led to a significant reduction in TLC-dependent luciferase activity with a remaining increase of 2.6 ± 0.2-fold (n = 10) at 10 µM TLC (**Figure 13B**). While the Q300X variant showed a similar TLC responsiveness as the WT at concentrations above 2.5 µM, no significant rise in luciferase activity was detected after stimulation with 0.1 µM TLC. The loss of the last 20 amino acids in the TGR5 variant S310X had no effect on receptor responsiveness towards TLC (**Figure 13B**). These results suggest that amino acids 284-297 are essential for localization of TGR5 in the PM. To elucidate the role of these residues in more detail, additional variants were generated: deletion of amino acids 285-290, 285-297, and 291-297, and alanine, proline and glycine substitution of amino acids 285-290 and 291-297.

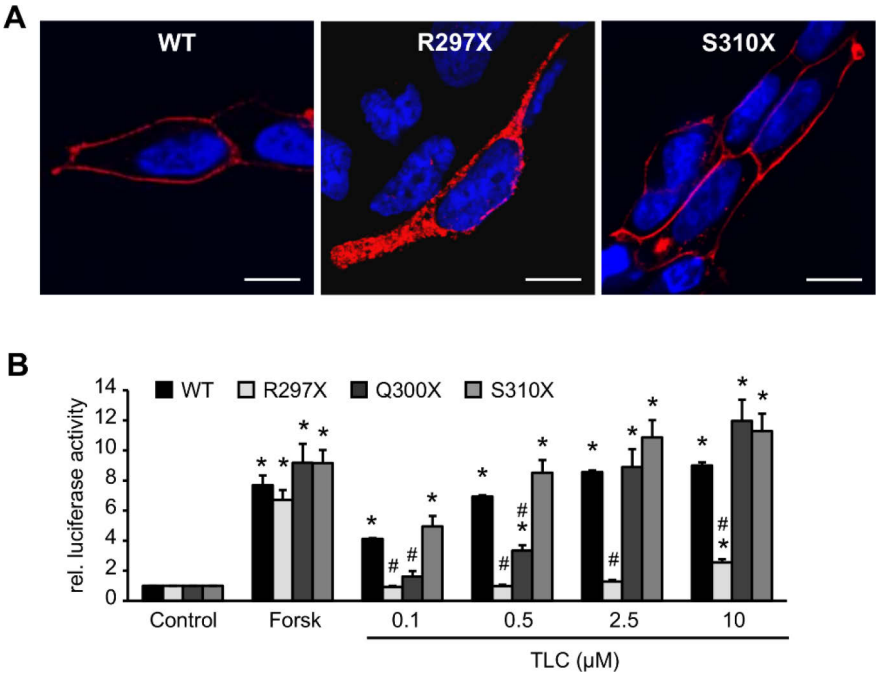


Figure 13 Localization and function of TGR5 truncation variants in HEK293 cells determined by immunofluorescence microscopy. **A** Localization of the FLAG-TGR5-YFP truncation variants. While the WT and the S310X variant are present in the PM, the R297X variant is retained in the ER. **B** TLC responsiveness measured by a fluorescence increase using a cAMP reporter gene assay. While the S310X variant shows no impairment compared to the WT and the Q300X variant is rescuable at high concentrations of TLC, the R297X variant shows a significantly decreased TLC response at all tested concentrations. Figure adapted from **publication I**.

In order to identify similarities and differences in the secondary structure of the TGR5 WT and the aforementioned substitution and deletion variants on a per-residue level, I performed MD simulations of the 18 membrane-proximal amino acid of their C-termini. I then pooled all conformations of the last 500 ns of all MD trajectories of the WT and all variants and hierarchically clustered them according to their secondary structure sequence. As the most outstanding result, a clear correlation between the secondary structure sequence of the respective peptide and the localization of TGR5 and, thus, its function emerged. Variants with a high membrane localization and TLC responsiveness predominantly appear in clusters with a high α-helix content (clusters 1 and 5; **Table 1**, **Figure 14**), while variants showing ER retention appear in clusters with high amounts of loop or β-sheet formation (clusters 2-4; **Table 1**, **Figure 14**). For example, we could show that the WT peptide encompassing residues 285-297 preferentially forms an α-helix. In contrast, the 285-290A variant, which was retained in the ER, showed an exclusive β-sheet formation within the first 120 ns of the simulation. This finding was unexpected because alanine has a high helix propensity. However, the β-sheet formation seemed to be favored in this case by interactions between the alanine residues in positions 285-290 with naturally occurring hydrophobic

alanines in positions 294-296. These initial analyses suggested that a high β-sheet content in the membrane-proximal C-terminus prevents ER to PM trafficking of TGR5.

Table 1 Results of clustering according to secondary structure sequence, function, and protein localization of the TGR5 membrane-proximal C-terminus**.** Table adapted from **publication I**.

	Cluster[1]						
Variant	1	2	3	4	5	Function[2]	Membrane localization[3]
TGR5WT	56	1	8	10	25	100.0 ± 7	90.9 ± 0
285-290A	5	85	7	1	2	13.3 ± 1	52.8 ± 3
285-290G	57	1	42	0	0	30.3 ± 3	69.7 ± 4
285-290P	31	0	16	53	0	14.4 ± 1	55.0 ± 3
Δ285-290	16	0	71	3	10	36.6 ± 2	70.8 ± 3
291-297A	23	0	15	0	62	174.1 ± 23	80.3 ± 4
291-297G	13	76	5	1	5	11.9 ± 1	48.1 ± 4
291-297P	0	0	12	30	58	10.5 ± 1	40.7 ± 4
Δ291-297	6	0	38	25	31	107.5 ± 6	70.3 ± 2

[1]Percentage of the cluster distribution for each variant.
[2]Function at 10 µM TLC as percent of wildtype ± SEM.
[3]Percentage of cell membrane localization determined by FACS analysis ± SEM.

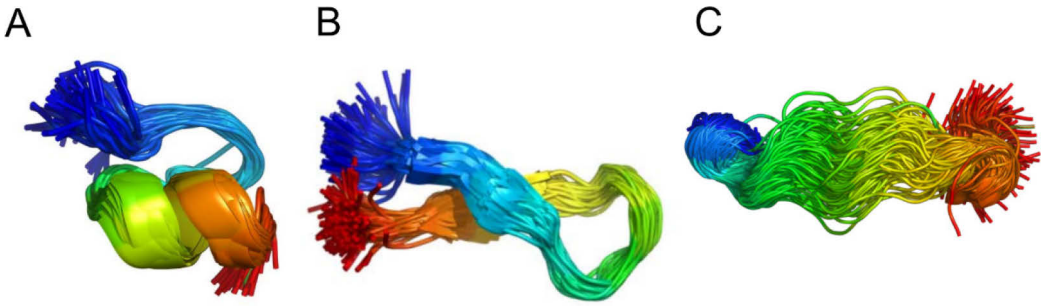


Figure 14 Clustering of conformations from MD simulations according to secondary structure. **A** Conformation with the most frequently occurring secondary structure sequence of the WT in cluster 1. **B** Conformation with the most frequently occurring secondary structure sequence of the 285-290A variant in cluster 2. **C** Conformation with the most frequently occurring secondary structure sequence of the 291-297P variant in cluster 5, which except for an α-helical turn at the N-terminus is unstructured in contrast to other variants in cluster 5. The coloring indicates the sequence from the N-terminus (blue) to the C-terminus (red). Figure adapted from **publication I**.

As a proof of principle, chimeras of TGR5 containing the membrane-proximal amino acids of the β_2-AR, the $S1P_1$, and the κ-OR were generated. The respective amino acid sequence of the receptors form α-helices as shown in high resolution crystal structures. MD simulations of the membrane-proximal 18 amino acids of the C-terminus of the TGR5β_2AR chimera, which contains 13 amino acids of the membrane-proximal C-terminus of β_2AR, reproduced this α-helix character, which demonstrates the quality of the setup of my simulations. Unsurprisingly, the TGR5 chimeras were correctly sorted to the PM and showed similar functional activity in response to 10 µM TLC as WT TGR5. However, the membrane-proximal part of the β_2AR contains an $F(X)_6LL$ motif, which has previously been identified as an important GPCR ER export motif. To assess the mechanism of the $F(X)_6LL$ motif promoting membrane trafficking of the TGR5β_2AR chimeric receptor, we evaluated

chimera variants containing mutations in this motif. This again revealed a strong correlation between the secondary structure content of the respective variant and its localization and function.

The double alanine mutant LL294/5AA revealed the lowest α-helical content but pronounced β-sheet formation. As in the case of the TGR5 variants, this chimera variant showed marked retention in the ER. In the MD simulations, I could pinpoint the β-sheet formation to a hydrophobic interaction between the mutated residues and F291. In order to test whether a disruption of this interaction restores α-helicality, I subjected the F291A//LL294/5AA variant to MD simulation. Indeed, this variant had a lower β-sheet and a higher α-helical content. In experimental validations of these findings, the F291A//LL294/5AA variant showed a PM localization level and luciferase activity in response to 10 µM TLC indistinguishable from the TGR5β_2AR chimera. This suggests, that the F$(X)_6$LL motif might facilitate α-helix formation to promote PM localization of GPCRs.

In summary, these results demonstrate that PM trafficking and, thus, function of TGR5 are determined by the α-helical structure of the membrane-proximal C-terminus rather than a sorting motif.

4.3 Conclusion and significance

In this study, I clustered MD simulations of TGR5 C-terminus variants according to their secondary structure formation. The combination with an experimental characterization by immunofluorescence, FACS sorting and cAMP reporter gene assay revealed a strong correlation between secondary structure and PM localization and function. For a proof of principle, I selected the C-termini of the three GPCRs for the generation of TGR5 chimeras. One of those C-termini contained a sorting motif, which was subsequently substituted with alanine residues. I detected β-sheet formation in an ER retained alanine variant in which I could restore α-helicality, as shown in MD simulations.

The principal results of this study are:

- In this study, it was shown for the first time, that the secondary structure of the C-terminus determines the PM localization of a GPCR, rather than a sorting motif.
- Variants of TGR5 with high PM localization and function were found in clusters with high α-helical content, while variants with low membrane localization were found in clusters with β-sheet or loop formation.

- As a proof of principle, three TGR5 chimeras with the α-helical C-termini of other GPCRs were correctly sorted to the PM.
- One of those chimeras contained a sorting motif, which, when mutated to alanine, led to β-sheet formation and ER retention of this variant. However, the PM localization could be rescued by introduction of an additional mutation, which restored α-helicality in MD simulations.
- This is the first evidence that sorting motifs might promote PM localization by facilitating α-helix formation in the C-terminus.

The membrane localization of a GPCR can be dependent on homodimerization in the ER. As the 1-8 interface, utilizing the membrane-proximal C-terminus, has been identified in a GPCR (**chapter 2.2.4**), its secondary structure formation may influence dimerization. Hence, the dimerization interfaces of TGR5 have been investigated in **publication II** with the expectations to identify the 1-8 interface as one of the dimerization sites of TGR5.

5 Publication II - Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study

Annemarie Greife, Suren Felekyan, Qijun Ma, Christoph G.W. Gertzen, Lina Spomer, Mykola Dimura, Thomas O. Peulen, Christina Wöhler, Dieter Häussinger, Holger Gohlke, Verena Keitel, Claus A.M. Seidel

Sci. Rep. (2016), 6, 36792
Original publication, see pages 70-124; contribution: 10%

5.1 Background

I could show in the previous **publication I** that the secondary structure of the TGR5 C-terminus has a profound impact on the membrane localization of the receptor. For several GPCRs, it has been discovered that homodimerization in the ER is required for the correct sorting to the PM (**chapter 2.2.4**). Among the three different dimerization interfaces identified in X-ray crystal structures of GPCRs to date, the C-terminus is the main contributor of the protomer interaction in the 1-8 interface. A disruption of the secondary structure of the C-termini and, thus, their interaction could explain the ER retention of TGR5 variants with β-sheet or loop formation in their C-termini. Hence, the 1-8 interface is expected to be a dimerization site of TGR5.

As the dimerization of TGR5 could not only influence the membrane localization of the receptor but also affect its activation, we investigated possible di- and oligomerization interfaces of TGR5. I did integrative modeling in a combined strategy in which we applied cellular biology, and Multiparameter Image Fluorescence Spectroscopy (MFIS) for quantitative FRET analysis to obtain structural information about dimerization and higher-order oligomerization assemblies of TGR5. Particularly, a naturally occurring Y111A mutant was investigated, as it showed 60% less dimerization in co-immunoprecipitation assays than the TGR5 WT.

5.2 Structural assembly of TGR5 di- and oligomers

FRET between TGR5 molecules C-terminally fused to enhanced GFP as a donor or mCherry as an acceptor was measured for three different TGR5 variants: TGR5 WT, Y111A and Y111F. Stimulation of the WT, Y111A, or Y111F with TLC led to a significant dose-

dependent increase in luciferase activity in all three cases, which shows that the three variants are fully functional. FRET was detected in all TGR5 variants, indicating at least homodimerization. Interestingly, the TGR5 variants showed differences in their FRET properties: Upon titration of the acceptor, the energy transfer efficiency did not change significantly in Y111A in contrast to WT and Y111F. This indicates that the Y111A variant forms high amounts of dimers but not oligomers, as fluorescence quenching cannot occur in monomers, while the efficiency changes in the Y111F variant and the WT suggest that higher-order oligomers, at least tetramers, are present in those variants.

To quantify this, we formally describe the fluorescence decays by two FRET-rate constants, which are for convenience given in units of apparent distances $R_{DA,app}$. For all TGR5 variants, this k_{FRET} fit resulted in a short apparent distance $R_{DA,app-1}$ with a small fraction and a long apparent distance $R_{DA,app-2}$ with a large fraction. As shown in **Figure 15**, in the WT and Y111F both apparent distances $R_{DA,app-1}$ and $R_{DA,app-2}$ became shorter ($R_{DA,app-1}$ = 40-20 Å; $R_{DA,app-2}$ = 75-50 Å) with increasing acceptor concentration. Furthermore, the species fractions also changed: The short distance-fraction increased from 7% to 30% in an acceptor-dependent manner, leading at the same time to a strong reduction of the long distance-fraction from 39% to 12%. This change is only possible in oligomers, as its limited range disallows FRET between distant, i.e. not oligomerized, dimers.

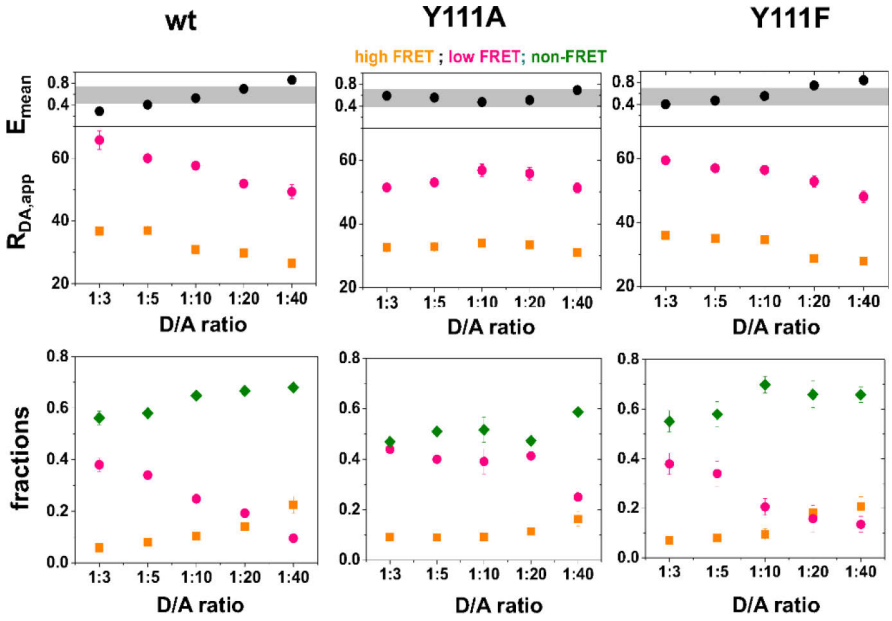


Figure 15 FRET-decays from sub-ensemble analysis at different donor-to-acceptor ratios were fitted with a two-k_{FRET} fit to obtain two apparent distances $R_{DA,1}$ and $R_{DA,2}$ (upper row) with their corresponding FRET fractions (lower row) and to calculate the mean transfer energy efficiency E_{mean}. E_{mean} increased in an acceptor-dependent manner in TGR5 wt and TGR5 Y111F, whereas E_{mean} changed only slightly in TGR5 Y111A. These changes in E_{mean} correlate with a reduction of both apparent distances $R_{DA,1}$ and $R_{DA,2}$ in TGR5 wt and TGR5 Y111F: In the lower row, the $R_{DA,1}$ fractions increase, whereas the $R_{DA,2}$ fractions decrease in an acceptor-dependent manner. Orange: $R_{DA,1}$ and $R_{DA,1}$ fraction, pink: $R_{DA,2}$ and $R_{DA,2}$ fraction, green: non-FRET fraction, the gray bar in E_{mean} represents average E_{mean} for TGR5 Y111A. Figure adapted from **publication II**.

These results were used to determine the di- and oligomerization interfaces of TGR5. For this, the $R_{DA,app}$ was correlated to the theoretical $R_{DA,app}$ calculated from dimer models of TGR5. The $R_{DA,app}$ of the Y111A variant was used, as the titration experiments suggest predominant homodimer formation of this variant, so that the absence of oligomerization allows the exact measurement of the dimerization interface. I built dimer models of TGR5 based on the interfaces structures known from the CXCR4, the μ-OR, and the κ-OR (**Figure 16**). Then I simulated the movement of the linker and a fluorophore by MD simulations. Subsequently, I calculated the conformational free energy and entropy contribution in combination with the dimer models. I used this to obtain a Boltzmann-weighted distribution of the fluorophore position in relation to the dimers. The average length of the linker in this approach is about 5 Å less than in AV simulations, which are considered to be less accurate. The $R_{DA,app}$ for the 1-8 interface of TGR5 calculated from this distribution showed a remarkable similarity with the $R_{DA,app}$ of the Y111A variant. Thus, the primary site for TGR5 homodimerization is the 1-8 interface.

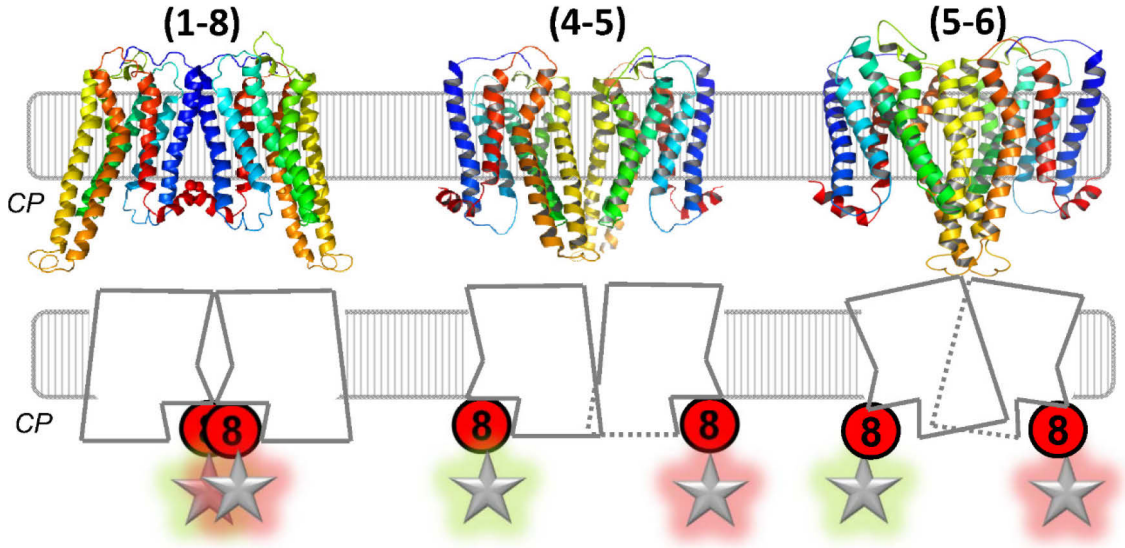


Figure 16 Homodimerization models with the following interfaces from left to right: (1/8), (4/5) and (5/6). TGR5 monomer chains are rainbow colored starting with TM1 in blue to H8 in red. Top row: membrane view of models displayed in PyMol. Bottom row: Schematic models. The attachment point for the fluorescent proteins (FP) at the cytoplasmic H8 is labeled with red circles, and FP's are presented as glowing stars in green for donor and red for acceptor. Abbreviation: CP = cytoplasm. Figure adapted from **publication II**.

In contrast to the Y111A variant, the titration experiments strongly suggest that the WT and Y111F variant form dimers and higher-order oligomers. The concentration dependence indicates that oligomers – (with a formation of tetramers as first step) - are formed from dimers (dimer of dimer model). The absence of a concentration dependence in the Y111A variant, a mutation in the ERY motif (**chapter 2.3.2**), implies the presence of at least a second interface for TGR5 homo-oligomer formation, which involves the ERY motif. We suggest that the

TGR5 oligomers must resemble a one-dimensional array mediated by a single oligomerization interface, because one mutation in the ERY motif Y111A in intracellular loop (ICL2) affects the oligomerization significantly. As shown in **Figure 17** the Y111 residue can interact with TM5 or/and TM6 dependent on its structural environment, which could be either helical or flexible. Hence, both the 4-5 and 5-6 interface could be potential interaction sites for oligomerization. We suggest that the TGR5 oligomers must resemble a one-dimensional array mediated by a single oligomerization interface forming either a (1-8):5-6:(1-8) or (1-8):4-5:(1-8) pattern (see **chapter 2.2.4**). One-dimensional arrays forming mainly (1-8):4-5:(1-8) oligomers have also been found for rhodopsin[287, 288] and (1-8):5-6:(1-8) oligomers for the μ-opioid receptor[222]. However, due to a high similarity of the expected $R_{DA,app}$ of the 4-5 and the 5-6 interface we cannot distinguish between those interfaces at present.

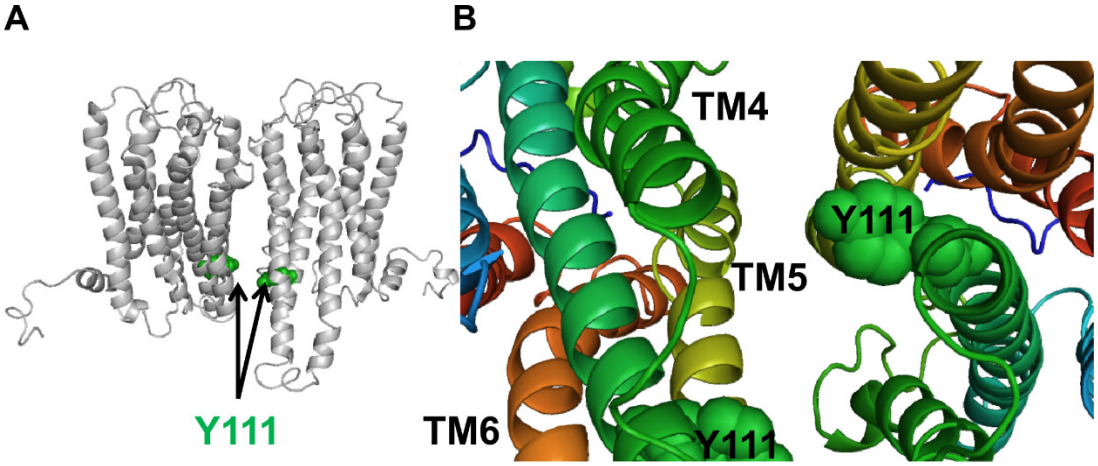


Figure 17 Influence of Y111A on dimerization. **A** The dimerization model (4/5) is displayed as a gray colored cartoon viewed from the membrane. Residue Y111 located in ICL2 is depicted as a green sphere in each TGR5 monomer. **B** Blow-up of the region around residue Y111 to show possible interactions between Y111 from one TGR5 molecule with residues in TM4 (green) and TM5 (yellow) in a second TGR5 molecule. Figure adapted from **publication II**.

5.3 Conclusion and significance

In this study, I built dimer models of TGR5 and calculated the position of the fluorophore relative to the protomers using an all-atom MD simulation. I enhanced the accuracy by calculating the probability distribution over all positions via a Boltzmann-weighing of their conformational free energy I obtained from MM-PBSA calculations including their entropy contribution. This is, to my knowledge, the first time conformational free energies in combination with their entropic contribution were used to enhance the sampling of fluorescent dye movements. My results were necessary to identify the 1-8 interface as a primary dimerization interface of TGR5, as suggested by C-terminus variants in **publication I**. Furthermore, I identified binding partners for Y111 in TM5 and TM6 on the basis of my TGR5

dimer models. Based on these interactions, I predicted the 4-5 and the 5-6 interfaces to be possible oligomerization interfaces of TGR5, as the Y111A variant disrupts its oligomerization.

The principle results of this study are:

- Combined molecular biology, fluorescence microscopy approaches, as well as bioinformatics modeling and simulations identified the 1-8 interface as the primary TGR5 dimerization site.
- TGR5 was shown to form higher-order oligomers.
- The Y111A variant, a mutation in the conserved ERY motif, abolished oligomerization of TGR5.
- This is the first indication that the tyrosine residue in the (D/E)RY motif might be important for GPCR oligomerization.

In GPCRs, a bound ligand can mediate effects via allosteric mechanisms to other protomers in a dimer. This could be dependent on the exact binding mode of the ligand in the binding pocket, as agonists and antagonists can both mediate their effects. Knowledge of the binding mode of bile acids combined with dimerization models of TGR5 could also be used to develop bivalent TGR5 ligands, targeting both protomers of a dimer. Hence, we generated a binding mode model of TGR5 agonists in **publication III**.

6 Publication III - Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists

Christoph G.W. Gertzen[§], Lina Spomer[§], Sander H.J. Smits, Dieter Häussinger, Verena Keitel, Holger Gohlke

[§] Both authors contributed equally to this work.

Eur J Med Chem (**2015**), 104, 57-72.

Original publication, see pages: 125-163; contribution: 30%

6.1 Background

In the previous **publication II**, I could show that TGR5 dimerizes via the 1-8 interface. GPCR signaling can be influenced by dimerization with effects transmitted to other protomers, which may be dependent on the ligand binding mode (**chapter 2.2.4**). Additionally, TGR5 could be a target for the treatment of metabolic diseases, which requires potent and selective agonists (**chapter 2.1.4**), and specialized cancer treatment, which requires potent and selective antagonists (**chapter 2.1.7**). TGR5 antagonists are currently unknown but could be derived from TGR5 agonists by introduction of small chemical modifications in specific positions. A binding mode of TGR5 agonists can direct these changes to simplify antagonist discovery, and could also be used in the development of more potent and selective agonists.

In this publication we used a combination of homology modeling (**chapter 2.3.2**), molecular docking (**chapter 2.3.3**), 3D-QSAR, MD simulations, and site directed mutagenesis with subsequent evaluation of TGR5 ligand responsiveness and membrane localization to elucidate a binding mode model of TGR5 agonists. 68 TGR5 agonists including natural and synthetic bile acids as well as neurosteroids were used to evaluate the binding mode model using the AFMoC approach (**chapter 2.3.4**).

6.2 Binding Mode Prediction of TGR5 Agonists

We considered two alternatives for TM7 of TGR5 when generating the multiple sequence alignment with the templates: In the first alternative, priority is given to the alignment of the conserved (D/E)X(K/R) motif (positions 8.48 to 8.50); in the second alternative, conserved Pro residues in TM7 are preferentially aligned. Both alignments cause a different orientation

of residues in TM7: It has the same length in both cases but residues at position 7.*n* in the first alternative are located at position 7.(*n*+2) in the second one. The second alignment based on the conserved proline residue in TM7 (**chapter 2.3.2**) should be more reliable than the first one. To confirm this hypothesis, we built models based on both alignments and predicted agonist binding modes in those models.

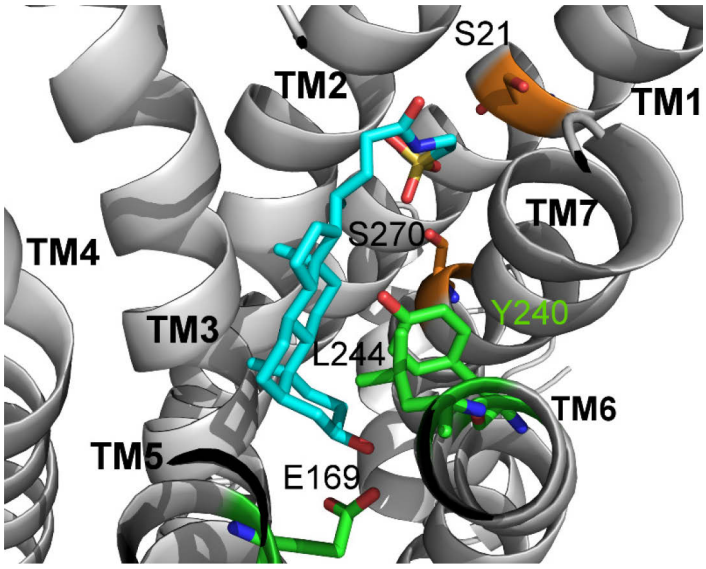


Figure 18 Binding mode of TLC predicted by molecular docking into the initial homology model of TGR5. TLC is shown with cyan sticks, all other TGR5 agonists with cyan lines, and TGR5 in gray cartoon representation. Amino acids subjected to mutational analysis are shown in sticks representation; they are colored according to having a negative effect (green) or no effect (orange) on receptor activity upon stimulation with TLC when mutated to alanine. In the docked binding mode, the 3-hydroxyl group of TLC forms a hydrogen bond with E169$^{5.44}$ and the sulfonic acid moiety of TLC forms a hydrogen bond with S270$^{7.43}$, whereas TLC does not interact with S21$^{1.39}$. Furthermore, TLC makes hydrophobic interactions with Y240$^{6.51}$ and L244$^{6.55}$. Figure adapted from **publication III**.

Binding modes found in models of the first alignment alternative generally showed no significant AFMoC model. The exchange of binding positions between models of the first and proline-centered alignment, to check for a model-independent valid binding poses, also yielded no valid AFMoC model. Only a binding mode found in a model of the proline-centered alignment resulted in a significant AFMoC model (q^2 = 0.37 for six components), which was expected because the alignment on the conserved proline residue should result in the correct orientation of TM7. This pose was used as our initial binding mode model. In the initial binding mode model TLC forms a hydrogen bond with E169$^{5.44}$ using its 3-hydroxyl group and hydrophobic interactions with Y240$^{6.51}$ and L244$^{6.55}$ in TM6 (**Figure 18**). Furthermore, TLC binds with its sulfonic acid moiety in the vicinity of TMs 1 and 7 deep inside the pocket to S270$^{7.43}$ in TM7 (**Figure 18**). The interacting residues were mutated to alanine and experimentally investigated regarding their response to TLC stimulation and membrane localization to validate the binding mode model. Additionally, S21$^{1.39}$ was mutated to alanine as a negative control and to invalidate binding mode models based on the

first alignment alternative. Here, bile acids bind to S21$^{1.39}$ with their sulfonic acid moiety, so that no response to the alanine mutation of this residue would invalidate these binding modes while consolidating our initial binding mode model. The alanine mutation of E169$^{5.44}$ in TM5 was predicted to influence ligand binding (see **chapter 2.2.1**) while the mutations of Y240$^{6.51}$ and L244$^{6.55}$ in TM6 were predicted to prevent the activation of TGR5.

As expected, the TLC-dependent luciferase activity of S21A$^{1.39}$ was comparable to that of wildtype TGR5 (**Figure 19**). In contrast, E169A$^{5.44}$ showed a significantly reduced activity at TLC concentrations between 0.1 and 2.5 µM, but not 10 µM indicating an influence on TLC binding, as predicted. In L244A$^{6.55}$ the dose-response was significantly reduced at all concentrations compared to the WT, while the Y240$^{6.51}$ variant showed nearly no response to TLC stimulation. These results confirmed our initial binding mode model.

We subjected TLC in the initial binding mode model to MD simulations to incorporate ligand and receptor flexibility. The most notable change was the breaking of the hydrogen bond between S270$^{7.43}$ and the sulfonic acid moiety of TLC, which reoriented towards R79^{EL2} to form a salt bridge with this residue. Furthermore, TLC showed a tendency to form a hydrogen bond with its 3-hydroxyl group to the hydroxyl group of Y240$^{6.51}$, in addition to a hydrogen bond with E169$^{5.44}$.

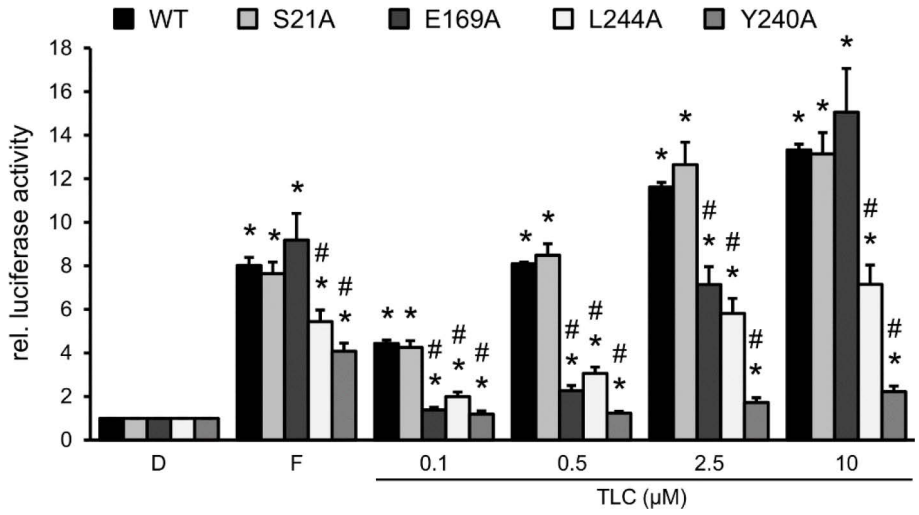


Figure 19 Experimental validation of the initial binding mode. Receptor activity towards taurolithocholate (TLC) was measured using a cAMP responsive luciferase construct, and luciferase activity served as a measure of the rise in intracellular cAMP following activation of TGR5. Forskolin (F, 10 µM) was used as TGR5 independent positive control. Dimethyl sulfoxide (D) was used as a negative control. The variant S21A$^{1.39}$ did not affect receptor responsiveness. E169A$^{5.44}$ and L244A$^{6.55}$ showed reduced luciferase activity at lower TLC concentrations (0.1 – 2.5 µM), while retaining activity like the WT at 10 µM TLC. The variant Y240A$^{6.51}$ almost abolished TLC-dependent luciferase activity at all concentrations tested and also significantly reduced forskolin-mediated rise in cAMP. Results (WT n = 21; S21A$^{1.39}$, E169A$^{5.44}$ n = 8; L244A$^{6.55}$, Y240A$^{6.51}$ n = 7) are expressed as mean ± SEM. *, # = significantly different ($p \leq 0.01$) from DMSO and TGR5 WT, respectively. Figure adapted from **publication III**.

An AFMoC model based on this refined binding mode showed an even higher predictivity ($q^2 = 0.50$ with one component) than our initial binding mode. In the refined binding mode,

we observed that agonists with a 7-hydroxyl group preferred a configuration shifted by about 3 Å towards helix 3 compared to TLC. For agonists with a 7-hydroxyl group in α-position, such as TCDC, this occurred due to hydrogen bond formation with Y89$^{3.29}$ in TM3 (**Figure 20**). Agonists such as TUDC with a 7-hydroxyl group in β-position formed a hydrogen bond with N93$^{3.33}$ instead.

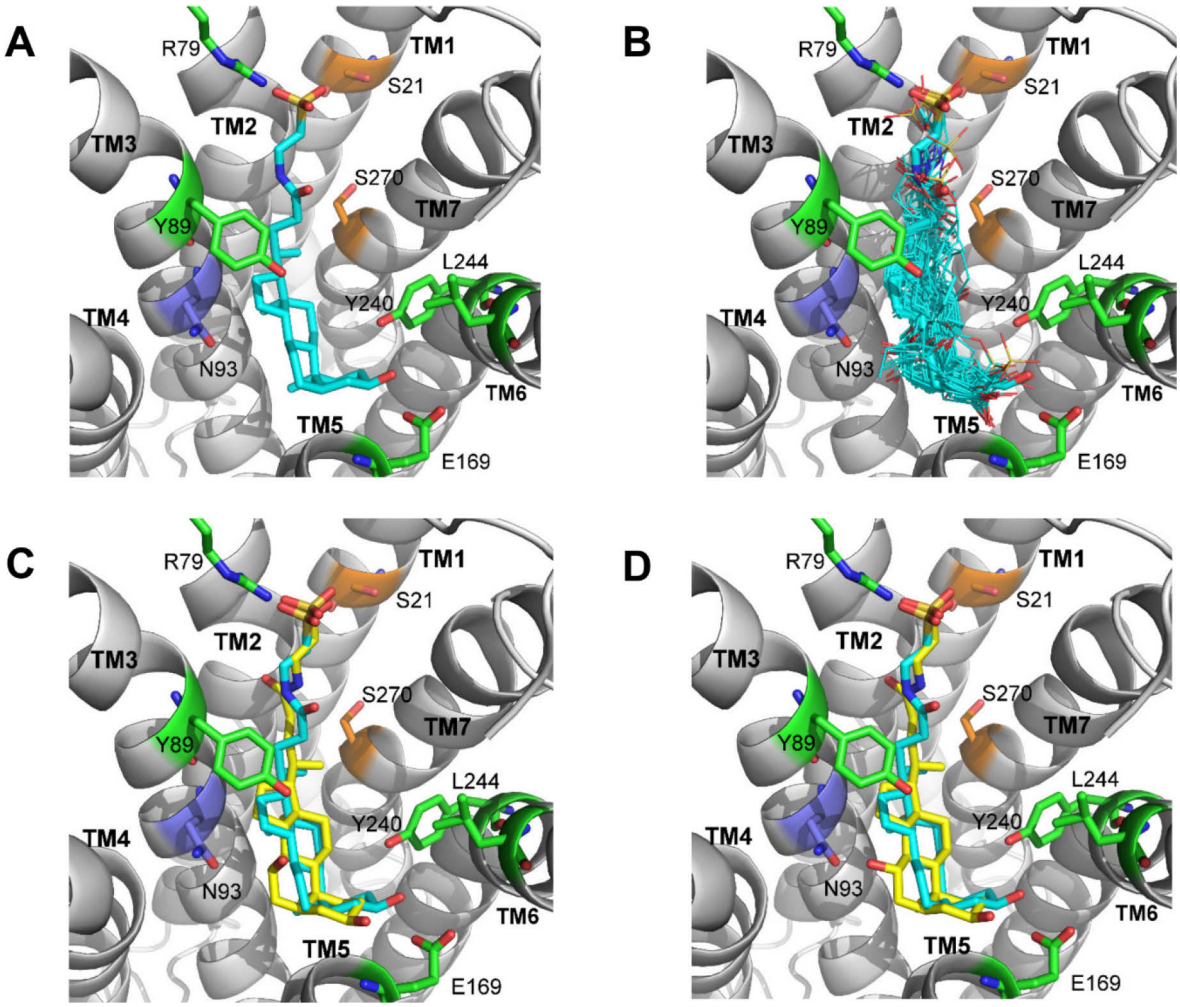


Figure 20 Binding mode of TLC after refinement of the TLC/TGR5 complex by MD simulations (A), corresponding alignment of TGR5 agonists used for the second AFMoC analysis (B), binding mode of TCDC (C) and TUDC (D) as compared to TLC. TLC is shown with cyan sticks, all other TGR5 agonists with cyan lines, and TGR5 in gray cartoon representation. Amino acids subjected to mutational analysis are shown in sticks representation; they are colored according to having a negative effect (green) or no effect (orange) on receptor activity upon stimulation with TLC when mutated to alanine. The N93A$^{3.33}$ variant (navy) was mainly retained intracellularly. In the refined binding mode (panel A), TLC forms a salt bridge to R79 in the EL1 with its sulfonic acid moiety. Furthermore, it forms hydrogen bonds with E169$^{5.44}$ and Y240$^{6.51}$ with its 3-hydroxyl group. Hydrophobic interactions are formed between TLC and Y89$^{3.29}$ and L244$^{6.55}$. Binding mode of TCDC (sticks representation, yellow) with its 7-α-hydroxyl group forming a hydrogen bond to Y89$^{3.29}$, moved towards TM3 relative to TLC (C), and binding mode of TUDC (sticks representation, yellow) with its 7-β-hydroxyl group forming a hydrogen bond to N93$^{3.33}$, moved towards TM3 relative to TLC (D). Figure adapted from **publication III**.

Consequently, mutations of the following residues should yield experimental support to the refined binding mode, including the shifted configurations for agonists with a 7-hydroxyl group: R79A^{EL1} and Y240F$^{6.51}$, as TLC forms hydrogen bonds with these residues; S270A$^{7.43}$

as a negative control, as TLC does not form hydrogen bonds with this residue; Y89A$^{3.29}$ and N93A$^{3.33}$, as these residues are predicted to have an influence on the activation of TGR5 by TCDC and TUDC, respectively, but not TLC (**Figure 20**).

As only 33.0±2.2% of N93A$^{3.33}$ reached the PM of transfected HEK293 cells as measured by FACS analysis, its impact on TUDC binding cannot be measured. In contrast, other variants reached the PM in sufficient amount to test their influence on ligand binding and receptor activation. The activity of S270A$^{7.43}$ was not significantly different to that of WT TGR5 at high concentrations of TLC, while R79A^{EL1} showed a dose-dependent increase in TLC-dependent luciferase activity, which was significantly reduced compared to WT TGR5. The effects of the S270A$^{7.43}$ and the R79A^{EL1} variant reflect the instability of the interaction of the TLC sulfonic acid moiety with S270A$^{7.43}$ which is given up in favor of a salt bridge with R79A^{EL1}, as seen in MD simulations. Y240F$^{6.51}$, effectively removing the hydroxyl group binding to TLC in the refined binding mode, almost completely abolished TLC induced luciferase activity. Finally, Y89A$^{3.29}$ showed the highest impact on TCDC activity, with which it was predicted to form a hydrogen bond to its 7α-hydroxyl group. TLC with no 7-hydroxyl group and TUDC, which does not form a hydrogen bond in the refined binding mode due to the β-configuration of its 7-hydroxyl group, showed a less pronounced reduction in TGR5 stimulation in the Y89A$^{3.29}$ variant. The hydrogen bond formation of Y89$^{3.29}$ with TCDC (EC_{50} = 2.3 μM) rather than with TUDC (EC_{50} = 50.5 μM) (**chapter 2.1.2**), as seen in our refined model, explains the epimeric selectivity of TGR5 towards agonists with a 7-hydroxyl group in α-position.

All in all, we could validate our binding mode model of 68 TGR5 agonists by predicting the effects of nine mutations to either influence agonist binding or TGR5 function including negative controls. The predicted effects were verified by mutagenesis studies with subsequent localization and functional assays. Our binding mode model is highly accurate as it could not only predict the importance of the Y240$^{6.51}$ hydroxyl group for TGR5 activation but also identified the epimeric selectivity determining residue Y89$^{3.29}$.

6.3 Conclusion and significance

In this publication, I created an initial binding mode model of 68 TGR5 agonists using homology modeling, molecular docking, and AFMoC analysis. From this initial model, I predicted the influence of four mutations on TLC response, including a negative control, to validate this binding mode model. Experimental examination of the TGR5 variants corroborated the model. I then subjected TLC in the initial binding mode with TGR5 to MD

simulations. This led to an improved binding conformation with which I could create an even more predictive AFMoC model. For this refined binding mode model, I suggested five additional mutations for which I correctly predicted their influence, including the importance of the Y240$^{6.51}$ hydroxyl group and the epimeric selectivity being mediated by Y89$^{3.29}$.

The principle results of this study are:

- This is the first binding mode model of TGR5 agonists which is in line with nine mutations to TGR5.
- Furthermore, it explains the structure-activity relationships of 68 TGR5 agonists and is the first binding mode model including neurosteroid agonists of TGR5.
- The sulfonic acid moiety of TLC binds to R79^{EL2} while its 3-hydroxyl group forms hydrogen bonds to E169$^{5.44}$ and Y240$^{6.51}$.
- The binding mode model is highly accurate to a degree that it explains the epimeric selectivity of 7α-hydroxyl groups, which is mediated by Y89$^{3.29}$.

This binding mode model could be used for the development of more potent and selective agonists, and the identification of antagonists. In combination with the TGR5 dimerization models identified in **publication II** bivalent ligands of TGR5 could be developed, which target both protomers of a dimer. Furthermore, this binding mode model might be used to explain possible receptor crosstalk in combination with those dimer models.

SUMMARY

In the present work, I first computationally investigated the secondary structure formation of the membrane-proximal C-terminal helix 8 of TGR5 with the aim of correlating secondary structure formation to membrane localization (**publication I**). This should show whether α-helicality of the C-terminus as observed in crystal structures of GPCRs induces membrane localization. I conducted MD simulations of the TGR5 C-terminus and mutants thereof and clustered the structures into groups of similar secondary structure. The results for each mutant were correlated with a characterization of their function and localization done in the lab of Prof. Dr. V. Keitel. This revealed that the C-termini of all membrane localized mutants indeed adopted α-helices while β-sheet or loop formation led to ER retention.

As homodimerization can be another reason for the membrane trafficking of GPCRs, and helix 8 lies within a prominent interface for GPCR dimerization, (**chapters 2.2.3** and **2.2.4**) the contact between these α-helices could promote TGR5 dimerization. Hence, I investigated possible dimerization interfaces of TGR5 (**publication II**). For this I built homology models of TGR5 based on known GPCR dimer interfaces which were used to determine the theoretical FRET efficiency for comparison to FRET measurements done in the lab of Prof. Dr. C. Seidel and Prof. Dr. V. Keitel. This allowed the identification of the 1-8 interface, which was suspected on the results based on **publication I**, as a dimerization site of TGR5.

Finally, I predicted a binding mode model of 68 TGR5 agonists including natural and synthetic bile acids, and neurosteroids, which was experimentally validated in the lab of Prof. Dr. V. Keitel (**publication III**). Here, I created homology models of TGR5 and docked the most potent natural agonist TLC into the TGR5 binding site (**chapter 2.3.2**). Subsequent improvement of the initial binding mode by incorporating ligand and receptor flexibility via MD simulations of the complex showed a reorientation of the sulfonic acid moiety of TLC. Remarkably, the effects of nine mutations suggested based on the model were in perfect agreement with the binding mode model. In combination with the experiments the binding mode model shows an astonishing level of detail. Based on a hydrogen bonding of the hydroxyl group of $Y240^{6.51}$ to TLC, I could correctly predict an abolishment of receptor activation in the Y240F variant, which effectively removes the Y240 hydroxyl group. Furthermore, the binding mode model explains the epimeric selectivity of TGR5 towards bile acids with a 7-hydroxyl group in α-position (e.g. TCDC) rather than in β-position (e.g. TUDC) mediated by $Y89^{3.29}$. All in all, this binding mode model is precise accurate enough to further the development of specific TGR5 agonists as well as antagonists.

PERSPECTIVES

TGR5 is involved in the formation of the cholangiocyte carcinoma, and the gastral and esophageal adenocarcinoma by the mediation of proliferative and anti-apoptotic effects (**chapter 2.1.7**). Antagonists inhibiting those signaling pathways (**chapter 2.1.3**) could pose a new therapy for these types of cancer. However, no antagonists of TGR5 are known to date. Based on the binding mode of agonists identified in **publication III**, I built a pharmacophore model (**Figure 21**) aiming at abolishing the interactions to TM6, which are vital for receptor activation (**chapter 2.2.2**). Thus, most of the interactions with TGR5, especially the salt bridge to $R79^{EL2}$, can be maintained while no inward force is exerted to TM6, which could lead to the activation of TGR5. With this pharmacophore model 48 potential antagonists of TGR5 were identified by a virtual screening, which will be tested in the lab of Prof. Dr. V. Keitel towards their inhibitory potential.

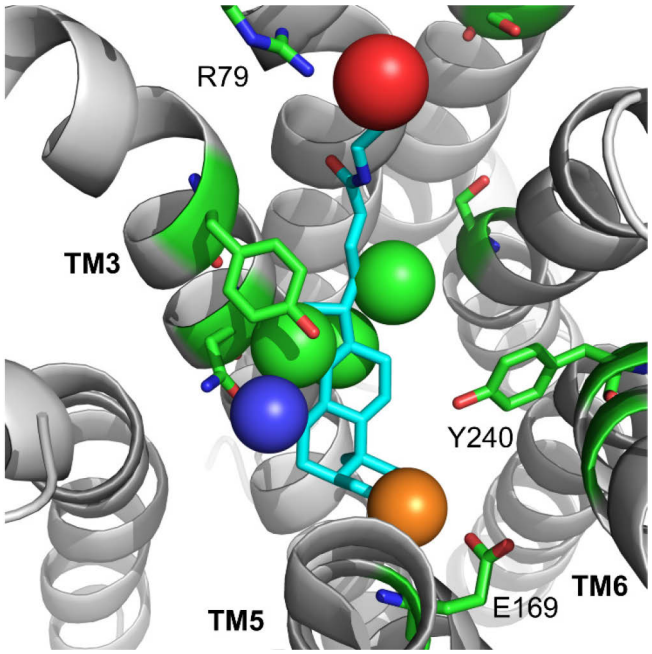


Figure 21 Pharmacophore model for the identification of potential TGR5 antagonists targeting the orthosteric pocket. The pharmacophore model was chosen from the binding pose of TLC (line representation) and includes the presence of anionic groups (red), hydrophobic groups (green), hydrogen bond donors (blue) and excluded volumes (orange).

Similarly, I aim at disrupting the TGR5/G-protein interaction in order to inhibit TGR5 downstream signaling without having to compete with high concentrations of bile salt agonists. For investigating the binding energetics of the TGR5/G-protein complex, I first built a homology model of the TGR5/G_s-protein complex based on our TGR5 homology model (**publication III**) and the β_2-adrenergic receptor/G_s-protein complex[157]. I then performed 160 ns of MD simulations of the complex in a POPC membrane (**Figure 22**).

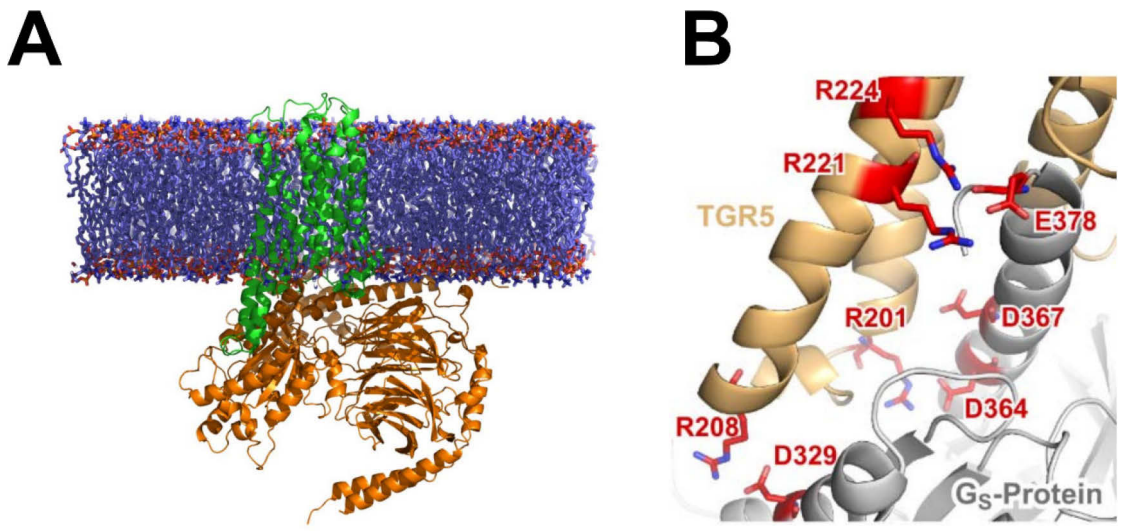


Figure 22 Identification of important residues mediating the TGR5/G_s-protein complex formation via MM-PBSA calculations. **A** Starting structure of the TGR5 (green) / G-protein (orange) complex model in a POPC membrane (navy). **B** Residues in the interface of TGR5/G_s-protein that contribute most to the binding affinity ("hot spots") (red) of the complex. E378 is at the C-terminal end of helix 5, and D367 and D364 at the N-terminal end.

Conformations of the complex were extracted from the MD trajectory and subsequently subjected to MM-PBSA calculations in an implicit membrane environment[289] to compute per-residue contributions to the effective energy of complex formation. Here, three residues in the N-terminal and C-terminal ends of helix 5 (H5) of the $G\alpha_s$-protein subunit were identified to contribute most (~5 kcal·mol^{-1}); these residues form salt bridge interactions with corresponding residues in TGR5 (**Figure 22**). This is in line with recent findings, which show that residues in the C-terminal end of H5 are important for the complex formation of the heterotrimeric G_s-protein and the β_2-adrenergic receptor (**chapter 2.2.2**)[174]. The residues identified to mediate the TGR5/G-protein complex formation could be used to screen for antagonists similar to a rational, structure-based approach to inhibit protein-protein interactions (PPI)[290] based on recent advances in the understanding of the energetics and dynamics of protein binding interfaces[291] and methodological developments in our working group[292-295].

Furthermore, it would be very interesting to uncover the exact mechanism of the influence of TGR5 C-terminus α-helix formation on TGR5 membrane localization (**publication I**). As it has been shown that Rab-GTPases are involved in the membrane trafficking of GPRCs[108, 186-189] (**chapter 2.2.3**), studying the interaction between H8 of TGR5 and Rab-GTPases could lead to success. Here, an approach similar to the one described in the previous paragraph can be used.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor Prof. Dr. Holger Gohlke for his continuous and invaluable guidance throughout my work. His encouraging and fruitful discussions and his never-ending enthusiasm always showed me new perspectives to overcome challenges.

Thanks to all of my colleagues and friends of the CPC lab. In particular, I would like to thank: Daniel Mulnaes, Benedikt Frieg, and Michele Bonus, for their constant, valuable and patient help in calculations, scripting, reading this thesis, for the fruitful and sometimes comical discussions brightening up my everyday life, and much more.

Thanks to Prof. Dr. V. Keitel, Dr. Lina Spomer, Prof. Dr. D. Häussinger, Prof. Dr. C. Seidel, Birte Schmitz, Dr. A. Greife, and Dr. S.H.J. Smits for the valuable discussions and fruitful cooperation to explore the secrets of TGR5, a chapter of my life I will always hold dearly.

Thanks to Prof. Dr. T. Kurz, Dr. F.K. Hansen, D. Diederich, K. Stenzel, Prof. Dr. M.U. Kassack, Dr. A. Hamacher, L.A. Alves Avelar, and V. Syntschewsk for the collaboration that brought success to our efforts to inhibit HDAC isoforms.

Furthermore, I would like to thank (in alphabetical order): Alexander Batzke, Bastian Bluhmki, Berti Botterbrod, Tibor Burton, Maira Deters, Christoph Dürmann, Aki and René Erler, Robert Fröhlich, Rüdiger Fröhlich, Dorothee Gronkowsky, Lisa Hankamer, Hannah and Michael Kokott, Tobias Kröger, Fabian Real, Christoph Sawatzki, and Ines and Sascha Stanicki for the wonderful time spent together.

Thanks to my father and grandmother, who were always there for me.

Last but not least, special thanks to Corinna for your unconditional love and support throughout this chapter of my life. Together with you, I am looking forward to experience the next big adventure, which is the rest of my life.

REPRINT PERMISSIONS

Figure 12 (page 29):
Reprinted from "*DrugScore Meets CoMFA: Adaptation of Fields for Molecular Comparison (AFMoC) or How to Tailor Knowledge-Based Pair-Potentials to a Particular Protein*", Holger Gohlke, Gerhard Klebe, Journal of Medicinal Chemistry, (2002), 45: 4153-4170, DOI: 10.1021/jm020808p, Copyright (2002) with permission from American Chemical Society.

Publication I (pages 53-68):
Reprinted from "*A Membrane-proximal, C-terminal α-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5"*, Lina Spomer, Christoph G.W. Gertzen, Birte Schmitz, Dieter Häussinger, Holger Gohlke, and Verena Keitel, Journal of Biological Chemistry, (2014), 289, 3689-3702, DOI: 10.1074/jbc.M113.502344, Copyright (2014), with permission from The American Society for Biochemistry and Molecular Biology, Inc.

Publication II (pages 53-68):
Reprinted from " *Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study"*, Annemarie Greife, Suren Felekyan, Qijun Ma, Christoph G.W. Gertzen, Lina Spomer, Mykola Dimura, Thomas O. Peulen, Christina Wöhler, Dieter Häussinger, Holger Gohlke, Verena Keitel, Claus A.M. Seidel, Scientific Reports, (2016), 6, 36792, DOI: 10.1038/srep36792, Copyright (2016), with permission from Nature Publishing Group.

Publication III (pages 148-186):
Reprinted from "*Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists"*, Christoph G.W. Gertzen, Lina Spomer, Sander H.J. Smits, Dieter Häussinger, Verena Keitel, Holger Gohlke, European Journal of Medicinal Chemistry, (2015), 104, 57-72, DOI: 104, 57-72, Copyright (2015), with permission from Elsevier B.V..

PUBLICATIONS

Publication I

A Membrane-proximal, C-terminal α-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5

Lina Spomer[§], Christoph G.W. Gertzen[§], Birte Schmitz, Dieter Häussinger, Holger Gohlke, and Verena Keitel

[§] Both authors contributed equally to this work.

J. Biol. Chem. (2014), 289, 3689-3702.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 6, pp. 3689–3702, February 7, 2014

A Membrane-proximal, C-terminal α-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5*

Received for publication, July 15, 2013, and in revised form, November 26, 2013 Published, JBC Papers in Press, December 13, 2013, DOI 10.1074/jbc.M113.502344

Lina Spomer[‡1], Christoph G. W. Gertzen[§1], Birte Schmitz[§], Dieter Häussinger[‡], Holger Gohlke[§2], and Verena Keitel[‡2,3]

From the [‡]Clinic for Gastroenterology, Hepatology, and Infectious Diseases and [§]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich-Heine-University, D-40225 Düsseldorf, Germany

Background: TGR5 is a G protein-coupled bile acid receptor that modulates the immune response, glucose homeostasis, and liver regeneration.
Results: Secondary structure of the receptor C terminus determines plasma membrane trafficking.
Conclusion: TGR5 plasma membrane content and responsiveness to extracellular ligands depends on C-terminal α-helix formation.
Significance: This provides insights into the structure-function relationship of TGR5, which is a potential drug target for metabolic diseases.

The C terminus of G protein-coupled receptors (GPCRs) is important for G protein-coupling and activation; in addition, sorting motifs have been identified in the C termini of several GPCRs that facilitate correct trafficking from the endoplasmic reticulum to the plasma membrane. The C terminus of the GPCR TGR5 lacks any known sorting motif such that other factors must determine its trafficking. Here, we investigate deletion and substitution variants of the membrane-proximal C terminus of TGR5 with respect to plasma membrane localization and function using immunofluorescence staining, flow cytometry, and luciferase assays. Peptides of the membrane-proximal C-terminal variants are subjected to molecular dynamics simulations and analyzed with respect to their secondary structure. Our results reveal that TGR5 plasma membrane localization and responsiveness to extracellular ligands is fostered by a long (≥ 9 residues) α-helical stretch at the C terminus, whereas the presence of β-strands or only a short α-helical stretch leads to retention in the endoplasmic reticulum and a loss of function. As a proof-of-principle, chimeras of TGR5 containing the membrane-proximal amino acids of the β_2 adrenergic receptor (β_2AR), the sphingosine 1-phosphate receptor-1 (S1P1), or the κ-type opioid receptor (κOR) were generated. These TGR5β_2AR, TGR5S1P1, or TGR5κOR chimeras were correctly sorted to the plasma membrane. As the exchanged amino acids of the β_2AR, the S1P1, or the κOR form α-helices in crystal structures but lack significant sequence identity to the respective TGR5 sequence, we conclude that the secondary structure of the TGR5 membrane-proximal C terminus is the determining factor for plasma membrane localization and responsiveness towards extracellular ligands.

TGR5 (Gpbar-1, M-Bar) is a G protein-coupled receptor (GPCR)[4] that is expressed almost ubiquitously in humans and rodents (1–4). The receptor is coupled to a stimulatory G protein. Both unconjugated and conjugated bile acids as well as various steroid hormones have been identified as potent TGR5 agonists (5, 6). In the liver, TGR5 is localized in sinusoidal endothelial cells, Kupffer cells, cholangiocytes, gallbladder epithelial cells, and gallbladder smooth muscle cells (7–13). Here, TGR5 modulates hepatic microcirculation, exerts anti-inflammatory, anti-apoptotic and choleretic effects, and promotes gallbladder filling (7, 12–15). In the intestine, TGR5 is expressed in enteroendocrine L-cells, in immune cells as well as in neurons and astrocytes of the enteric nervous system (16–18). Although the latter suggests a role for TGR5 in intestinal motility, activation of TGR5 in L-cells has been linked to increased glucagon-like peptide-1 secretion and the regulation of glucose homeostasis (18). In animal models administration of TGR5 agonists improved glucose tolerance and reduced liver inflammation and steatosis as well as atherosclerotic plaque formation (14, 18, 19). Thus, TGR5 is a promising drug target for the treatment of metabolic disorders, such as type II diabetes, obesity, atherosclerosis, and non-alcoholic steatohepatitis (7, 19).

Although TGR5 functions in different organs are progressively elucidated, the regulation of the receptor expression, localization, and function has not been studied so far. We previously identified a naturally occurring TGR5 mutation (Q296X) that leads to the truncation of the 35 C-terminal

* This work was supported by the Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 974 Düsseldorf "Kommunikation und Systemrelevanz bei Leberschädigung und Regeneration."
[1] Both authors contributed equally to the study.
[2] Both authors share senior authorship.
[3] To whom correspondence should be addressed: Clinic of Gastroenterology, Hepatology, and Infectious Diseases, Heinrich-Heine-University, Moorenstrasse 5, D-40225 Düsseldorf, Germany. Tel.: 49-211-8115087; Fax: 49-211-8115106; E-mail: verena.keitel@med.uni-duesseldorf.de.

[4] The abbreviations used are: GPCR, G protein-coupled receptor; aa, amino acids; AR, adrenergic receptor; MD, molecular dynamics; PM, plasma membrane; TLC, taurolithocholic acid; WT, TGR5 wild type; ER, endoplasmic reticulum; S1P1, sphingosine 1-phosphate receptor-1; κOR, κ-type opioid receptor; β_2AR, β_2-adrenergic receptor.

amino acids (20). In comparison to the wild type protein the Q296X variant failed to activate adenylate cyclase after stimulation with the agonistic bile acid taurolithocholic acid (TLC). Immunofluorescence staining revealed that the truncated protein was retained in the endoplasmic reticulum (ER) of transfected HEK293 and Madin-Darby canine kidney cells (20). Similar findings, *i.e.* a reduced functionality and retention in the ER, have been demonstrated for other GPCRs with a truncated membrane-proximal, intracellular C terminus, such as the luteinizing hormone/chorionic gonadotropin receptor (21), the vasopression 2 receptor (22), and the A1 adenosine receptor (23). This indicates that the cellular signaling response is determined by the amount of functionally active receptor in the plasma membrane (PM) (24). However, the underlying molecular mechanisms for ER retention varied for different GPCRs. Although the C terminus of some receptors contained an ER export signal composed of specific amino acid residues (25, 26), the C terminus of other receptors required a sequence of hydrophobic amino acids to form a putative helix 8 as a prerequisite for proper folding and anterograde trafficking to the PM (27, 28) (for a recent review, see Ref. 24).

Based on the naturally occurring TGR5 truncation mutation Q296X and guided by secondary structure predictions from molecular dynamics (MD) simulations, we generated further deletion and substitution variants within the membrane-proximal C terminus to identify the amino acid motifs/structural determinants that facilitate PM localization of this bile acid receptor. Using these variants and three chimeras of TGR5 with the membrane proximal C terminus of the β_2 adrenergic receptor (β_2AR), the sphingosine 1-phosphate receptor-1 (S1P1), or the κ-type opioid receptor (κOR), we demonstrate that the formation of a membrane-proximal α-helix (helix 8) is essential for anterograde trafficking of TGR5 from the ER to the PM and thus for receptor function.

EXPERIMENTAL PROCEDURES

Cell culture reagents were from PAA (Coelbe, Germany). Fetal calf serum (FCS) was from Biochrom (Berlin, Germany). TLC and forskolin (Forsk) were from Sigma and Calbiochem, respectively.

Cloning of TGR5 Variants and TGR5 Chimera—Human TGR5 was cloned as described (10, 20). TGR5 mutations were introduced into two different human TGR5 cDNA constructs. One construct contained part of the 5′-UTR and the complete coding sequence of human TGR5 including the stop codon and was cloned into the pcDNA3.1+ vector (Clontech, Palo Alto, CA) (20). The second construct, FLAG-TGR5-YFP, was cloned into the pEYFP-N1 vector (Clontech, Palo Alto, CA) (10) and contained the coding sequence with an N-terminal FLAG tag and a C-terminal yellow fluorescent protein (YFP) tag as described (20). Mutations were generated using the QuikChange Site-Directed Mutagenesis kit (Agilent Technologies, Santa Clara, CA). Deletion variants as well as the substitution variants 285–297A and 291–297P were generated using PCR-based cloning strategies. The receptor chimera TGR5β_2AR was cloned with primers, in which amino acids 285–297 (QRYTAPWRAAAQR) of the TGR5 C terminus were replaced with amino acids 330–342 (PDFRIAFQELLCL) of the β_2AR. The first PCR was carried out using the forward primer of the respective TGR5 pcDNA3.1+/pEYFP constructs (10, 20) and a reverse primer containing part of the sequence of the β_2AR. The second PCR was performed with a forward primer containing the β_2AR sequence fragment as well as the respective reverse primers of the TGR5 pcDNA3.1+/pEYFP constructs (10, 20). The third PCR used the PCR products of the first two PCRs as template and contained the forward and reverse cloning primers of the TGR5 pcDNA3.1+/pEYFP constructs (10, 20). The receptor chimera TGR5S1P1 was cloned accordingly with primers in which amino acids 285–294 (QRYTAPWRAA) of the TGR5 C terminus were replaced with amino acids 316–325 (KEMRRAFIRI) of the S1P1. The receptor chimera TGR5κOR was cloned with primers in which amino acids 285–296 (QRYTAPWRAAAQ) of the TGR5 C terminus were replaced with amino acids 335–346 (ENFKRCFRDFCF) of the κOR. All mutagenesis primer sequences and cloning strategies can be obtained upon request. Successful cloning and mutagenesis was verified by sequencing (GenBank™ accession numbers: TGR5, NM_001077191.1; β_2AR, NM_000024.5; S1P1, NM_001400.4; κOR, NM_000912.3).

Cell Culture and Transfection—Human embryonic kidney 293 (HEK293) cells were cultured in DMEM with 10% FCS and were kept at 37 °C and 5% CO_2. Cells were transiently transfected with cDNA plasmids using Lipofectamine2000 (Invitrogen) according to the manufacturer's instructions. Cells were used for immunofluorescence staining, FACS, or luciferase assays 48 h after transfection.

Immunofluorescence and Confocal Laser Scanning Microscopy—HEK293 cells, grown on glass coverslips, were transfected with the different TGR5 cDNA constructs. After 48 h cells were fixed in 100% methanol (−20 °C) for 30 s and incubated with the following antibody dilutions (2% FCS in PBS without calcium chloride and magnesium chloride ($PBS^{-/-}$): anti-TGR5 amino acids 298–318 (RVLR2) (10), 1:500; anti-FLAG-M2 (Sigma), 1:250; sodium/potassium (Na^+/K^+) ATPase (Sigma), 1:100; protein disulfide isomerase (Thermo Fisher Scientific Inc., Rockford, IL), 1:150. Fluorescent-labeled secondary antibodies (fluorescein, 1:100; cyanine-3, 1:500), were purchased from Dianova (Hamburg, Germany). Nuclei were stained with Hoechst 34580 (1:20,000; Invitrogen). Images were analyzed on a Zeiss LSM510META confocal microscope using a multitracking modus. A 63× objective and a scanning resolution of 1024 × 1024 pixels was used for all samples.

Flow Cytometry—To quantify PM localization, flow cytometry was performed as described (20). In brief, transiently transfected HEK293 cells were washed in $PBS^{-/-}$, detached with FACS buffer ($PBS^{-/-}$, 0.5 mM EDTA, 2% FCS) and centrifuged (2000 × *g*, 3 min, 4 °C). Cells were resuspended in FACS buffer and washed once by centrifugation and resuspension in FACS buffer (2000 × *g*, 3 min, 4 °C). The FLAG tag was detected with the anti-FLAG M2-antibody using the Zenon PacificBlue Label-Kit (Invitrogen) for 30 min (4 °C; dilution 1:250). Subsequently, cells were washed in FACS buffer and measured for forward scatter (FCS), side scatter (SCS), YFP, and PacificBlue fluorescence using a FACS-CANTO-II (BD Biosciences). TGR5 PM localization was determined by the amount of FLAG tag-positive cells divided by the total amount

of TGR5 positive cells as detected by YFP fluorescence. Control experiments were carried out using unlabeled non-transfected (no fluorescence) cells, unlabeled FLAG-TGR5-YFP transfected (only YFP fluorescence) cells, and permeabilized cells (20).

Measurement of TGR5 Activity Using a cAMP-responsive Luciferase Assay—HEK293 cells, grown on 6-wells, were co-transfected with a cAMP-sensitive reporter gene construct (Bayer AG; Leverkusen, Germany; 1.6 μg), a Renilla expression vector (Promega; Madison, WI; 0.1 μg), and human TGR5 variants in pcDNA3.1+ (0.5 μg diluted with 1.1 μg of pEYFP-N1 vector). The cAMP reporter construct contains five cAMP-responsive elements (CREs) in front of the luciferase gene. Luciferase activity was normalized to transfection efficacy. Cell lysis and luciferase assays were performed using the dual-luciferase kit (Promega) according to the manufacturer's instructions. Control experiments were performed with the pcDNA3.1+ vector (20). Luciferase activity after stimulation with DMSO (vehicle) served as control and was set to 1.0, and forskolin and TLC-dependent luciferase activity is represented as multiple thereof.

Molecular Dynamics Simulations of Peptidic Variants of the C Terminus—Peptidic deletion and substitution variants of the C terminus of TGR5 as well as of the TGR5β_2AR chimeras were subjected to MD simulations. Starting structures of all peptides were built using the *leap* program from the Amber 11 suite of molecular simulation programs (29). Each structure was generated in an extended conformation to reduce any bias by the starting structure on the subsequent MD simulations.

MD simulations were performed with Amber 11 (29) using the force field by Cornell *et al.* (30) with modifications suggested by Simmerling *et al.* (31). Each peptide was placed into an octahedral periodic box of TIP3P water molecules (32) and, if required, neutralized by Cl^- ions. The distance between the edges of the water box and the closest atom of the peptide was at least 11 Å, resulting in system sizes of ~16,000 atoms for the smallest Δ291–297 and ~41,000 atoms for the largest system 285–290A. The particle mesh Ewald method (33) was used to treat long-range electrostatic interactions. Bond lengths involving bonds to hydrogen atoms were constrained using SHAKE (34). The time-step for all MD simulations was 2 fs, with a direct-space, non-bonded cutoff of 8 Å. The starting structures were initially minimized by 2500 steps of steepest decent and conjugate gradient minimization applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms. Then NVT-MD, *i.e.* MD simulation with a constant number of particles, volume, and temperature, was carried out for 50 ps during which the system was heated from 100 K to 300 K. Subsequent NPT-MD, *i.e.* MD simulation with a constant number of particles, pressure, and temperature, was used for 50 ps to adjust the solvent density. As the last step of the thermalization procedure, we performed NVT-MD for 200 ps. Of the following 600 ns of NVT-MD at 300 K, the last 500 ns were used for analysis with conformations extracted every 40 ps.

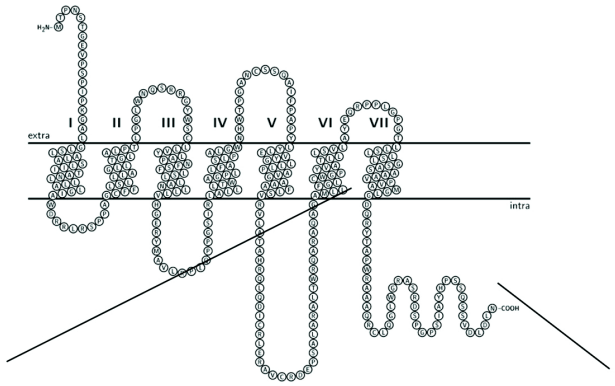
Two additional MD simulations were performed for the wild type (WT), the 285–290A, and the 291–297A and -G variants. When pooled alongside the original simulations, a similar distribution across clusters of secondary structure sequence (see below) was found (data not shown), thus indicating the convergence of the simulations.

Clustering According to Secondary Structure Sequence—For each conformation the secondary structure was assigned to each amino acid using the "continuous extension version" of the DSSP algorithm (35). This resulted in a string consisting of the letters B, E, G, H, T, and "–". To overcome the problem of sparse secondary structure types, we assigned all helical-like secondary structure elements (letter G (3–10-helix), letter T (Turn)) the letter H (α-helix). We also assigned the secondary structure element β-bridge (letter B) the letter E (β-sheet). Amino acids with unassigned secondary structure are marked by –. Based on these assignments, a similarity matrix of all conformations was computed considering only those assignments that occur at least 20 times per amino acid in each mutant. Missing amino acids in deletion variants were substituted with the letter X. The elements of the similarity matrix were computed by summing one for each exact letter match between respective string positions and zero otherwise. A hierarchical clustering was then performed on the similarity matrix applying Ward's method as implemented in the program *R* (36).

Statistical Analysis—Experiments were performed independently at least three times. Results are expressed as the means ± S.E. and analyzed using the two-sided student *t* test. A $p < 0.01$ was considered statistically significant.

RESULTS

PM Localization and Function of Membrane-proximal Deletion and Alanine Substitution Variants of the TGR5 C Terminus—A previous study had demonstrated that the deletion of the 35 C-terminal amino acids (aa) of human TGR5 (Gpbar-1) results in a retention of the mutated protein in the ER and a complete loss of function (20). To determine the role of the TGR5 C terminus for receptor localization and function, several C-terminal truncation variants were generated. Using the TMHMM program, Asp-284 was identified as the first amino acid of the intracellular C terminus (37, 38). Thus, premature stop codons were introduced into the TGR5 cDNA constructs (TGR5 pcDNA3.1+, FLAG-TGR5-YFP) by site-directed mutagenesis at the following residues: Asp-284, Pro-290, Arg-297, Gln-300, and Ser-310 (Fig. 1). Localization of the mutated proteins was studied in transiently transfected HEK293 cells by immunofluorescence staining with an anti-FLAG antibody and confocal laser scanning microscopy as described (20). Immunofluorescence staining of the truncation variants D284*X*, P290*X*, and R297*X* showed a reticular, intracellular fluorescence pattern (Fig. 2*A*) that was identified as the ER by double-labeling with an antibody against the ER marker protein disulfide isomerase (data not shown). The mutated protein Q300*X* was detected both in the ER and in the PM, whereas the S310*X* variant was mainly localized in the PM (Fig. 2*A*). Subsequently, TGR5 responsiveness was analyzed in HEK293 cells co-transfected with the respective TGR5 constructs and a cAMP-responsive luciferase reporter plasmid as described (8, 20). Luciferase activity served as a measure for the rise in intracellular cAMP after TGR5 activation by TLC, which is the most potent endogenous agonist for this receptor (1, 2, 6). Forskolin elevates cAMP levels independently of TGR5 and was used as a



	284	290	300	310	320	330
TGR5 WT	DQRYTAPWRAAAQR	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
D284X	X					
P290X	DQRYTAX					
R297X	DQRYTAPWRAAAQX					
Q300X	DQRYTAPWRAAAQR	CLX				
S310X	DQRYTAPWRAAAQR	CLQGLWGRASRDX				
Δ285-290	D- - - - - -WRAAAQR	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
Δ285-297	D- - - - - - - - - - - - -	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
Δ291-297	DQRYTAP- - - - - - -	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
285-290A	D AAAAAA WRAAAQR	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
285-290G	DGGGGGGWRAAAQR	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
285-290P	D PPPPPP WRAAAQR	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
291-297A	DQRYTAP AAAAAAA	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
291-297G	DQRYTAPGGGGGGG	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
291-297P	DQRYTAP PPPPPPP	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
TGR5β_2AR	D PDFR I A FQELLCL (330–342)	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
TGR5κOR	D ENFKRCFRDFCF R (335–346)	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				
TGR5S1P1	D KEM RRAFIRI AQR (316–325)	CLQGLWGRASRDSPGPSQAYHPSSQSSVDLDLN				

FIGURE 1. TGR5 cDNA constructs used to characterize the role of the C terminus for receptor localization and function. A two dimensional model of human TGR5 was created with TOPO2 software (Johns S. J., TOPO2, Transmembrane protein display software). Using the TMHMM program (37, 38) to predict transmembrane helices, aspartic acid 284 was identified as the first intracellular amino acid of the C terminus. To elucidate the role of the proximal TGR5 C terminus for receptor localization and function and to determine the potential secondary structure of this part, TGR5 truncation variants (D284*X*, P290*X*, R297*X*, Q300*X*, S310*X*), deletion variants (Δ), and alanine (*A*), glycine (*G*), or proline (*P*) substitutions were generated as shown. Furthermore, a receptor chimera (TGR5β_2AR) was generated in which aa 285–297 of TGR5 were replaced by 13 aa of the C-terminal α-helix from the β_2AR. For the receptor chimera TGR5S1P1, aa 285–294 of the TGR5 C terminus were replaced with amino acids 316–325 of the S1P1. In the TGR5κOR chimera, aa 285–296 of the TGR5 C terminus were replaced with amino acids 335–346 of the κ-type opioid receptor isoform-1 (κOR). *X*, premature stop codon.

positive control. As described previously, stimulation of TGR5 WT with TLC showed a dose-dependent increase in luciferase activity (8, 20) (Fig. 2*B*). The truncation of the TGR5 C terminus at amino acids 284 and 290 completely abolished the responsiveness of the receptor towards TLC, whereas the forskolin-induced luciferase activity was unaffected (Fig. 2*B*). Truncation at amino acid 297 led to a significant reduction in TLC-dependent luciferase activity with a remaining increase of 2.6 ± 0.2-fold (n = 10) at 10 μM TLC (Fig. 2*C*). Although the Q300*X* variant showed a similar TLC responsiveness as the WT at concentrations >2.5 μM, no significant rise in luciferase activity was detected after stimulation with 0.1 μM TLC. The loss of the last 20 amino acids in the TGR5 variant S310*X* had no effect on receptor responsiveness towards TLC (Fig. 2*C*). These findings are in line with previous studies demonstrating that the intracellular signaling response of a GPCR is determined by the amount of functionally active receptor in the PM (24). Furthermore, these results suggest that amino acids 284–297 are essential for localization of TGR5 in the PM. To elucidate the role of these residues in more detail, additional cDNA constructs were generated: deletion of amino acids 285–290, 285–297, and 291–297 and alanine substitution of amino acids 285–290 and 291–297 (Figs. 1, 3, and 4).

The deletion variants Δ285–290 and Δ285–297 were localized in the ER but also reached the PM, as demonstrated by the colocalization of the TGR5 fluorescence pattern with the stain-

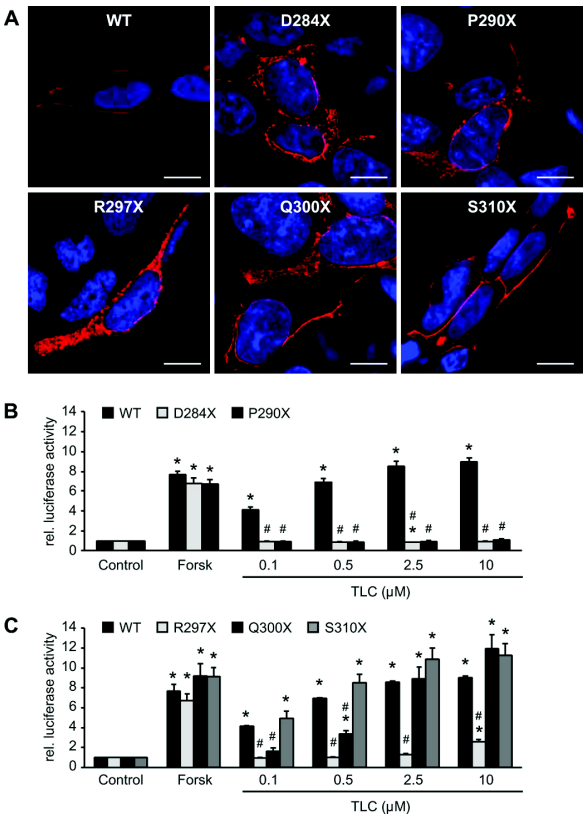


FIGURE 2. **Analysis of TGR5 truncated variants in HEK293 cells.** *A*, HEK293 cells were transiently transfected with the different FLAG-TGR5-YFP truncation variants. The FLAG tag was made visible using an anti-FLAG-M2 antibody (*red*). In contrast to the WT, the truncation variants D284*X*, P290*X*, and R297*X* were not detected in the PM but were localized intracellularly within the ER. The truncated proteins Q300*X* and S310*X* were localized both in the PM and the ER. Nuclei were stained with Hoechst (*blue*). *Bars* = 10 μm. *B* and *C*, HEK293 cells were co-transfected with TGR5 (pcDNA3.1+), a cAMP-responsive luciferase reporter construct, and a Renilla expression vector (8). Luciferase activity served as a measure of the rise in intracellular cAMP after activation of TGR5. Forskolin (*Forsk*, 10 μM) was used as a TGR5-independent positive control. The truncation variants D284*X* and P290*X* completely lost responsiveness to stimulation with the TGR5 agonistic bile acid TLC ($n = 9$ each; *B*). The R297*X* variant showed a 2.6-fold increase after incubation with 10 μM TLC, whereas the Q300*X* mutation was equally responsive to 10 μM TLC as the WT ($n = 10$ each; *C*). The variant S310*X* did not affect receptor responsiveness to TLC ($n = 12$; C). Results are expressed as the mean ± S.E. * and # = significantly different ($p < 0.01$) from DMSO (Control) and TGR5 WT, respectively.

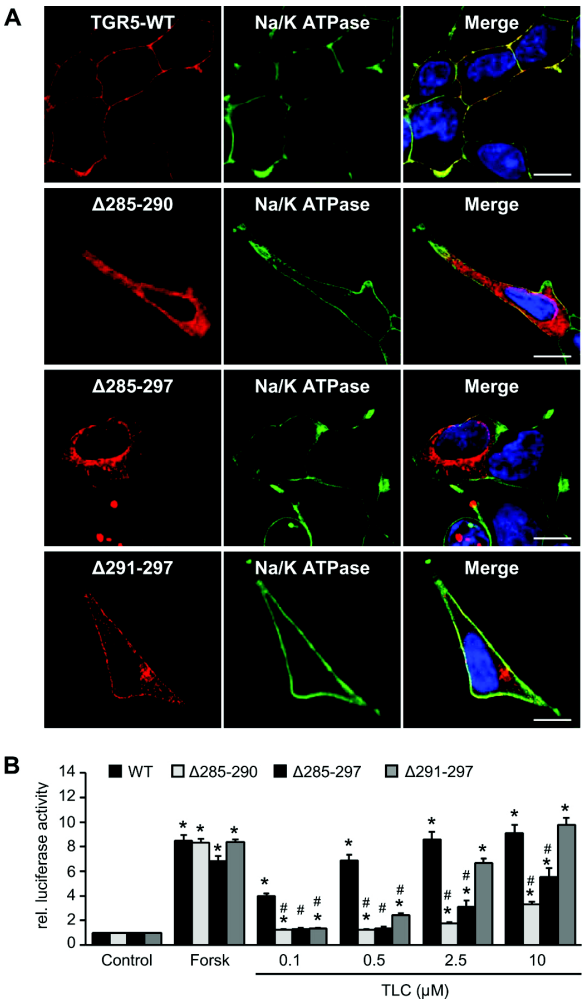


FIGURE 3. **Analysis of TGR5 deletion variants in HEK293 cells.** HEK293 cells were transiently transfected with TGR5 WT and the deletion TGR5 (pcDNA3.1+) variants. *A*, the TGR5 protein was made visible using an anti-TGR5 antibody (RVLR2, *red*). WT TGR5 was almost completely targeted to the PM as demonstrated by the yellow coloring in the overlay image with the PM marker protein sodium/potassium (Na^+/K^+) ATPase (*green*). Although the deletion variant Δ285–290 was mainly retained in the ER, the variant Δ285–297 was present both in the ER and the PM, and the Δ291–297 variant was located predominately in the PM. Nuclei were stained with Hoechst (*blue*). *Bars* = 10 μm. *B*, receptor responsiveness towards TLC was measured using a cAMP-responsive luciferase construct. The deletion variants Δ285–290 and Δ285–297 showed a significantly reduced responsiveness towards TLC stimulation as compared with the TGR5 WT. The Δ291–297 variant showed normal receptor activity after stimulation with 2.5 and 10 μM TLC and reduced activity in response to lower TLC concentrations (0.1 μM, 0.5 μM). Results (Δ285–290, Δ291–297 $n = 8$; Δ285–297 $n = 10$, WT $n = 23$) are expressed as the mean ± S.E. * and # = significantly different ($p < 0.01$) from DMSO (Control) and TGR5 WT, respectively.

ing of the PM marker protein sodium-potassium ATPase (Na^+/K^+ ATPase). In contrast, the Δ291–297 variant was predominately detected in the PM (Fig. 3*A*). The variants Δ285–290 and Δ285–297 showed a significant reduction in luciferase activity in response to TLC compared with the WT protein. However, stimulation of Δ285–290 and Δ285–297 with 10 μM TLC induced a significant rise in luciferase activity by 3.3 ± 0.2- ($n = 8$) and 5.5 ± 0.7-fold ($n = 10$), respectively (Fig. 3*B*). Although deletion of amino acids 291–297 (Δ291–297) resulted in a significantly reduced responsiveness towards low TLC concentrations (0.1 and 0.5 μM) compared with the WT, incubation with higher TLC concentrations (2.5 and 10 μM) led to a similar rise in luciferase activity as observed with WT (Fig. 3*B*).

The substitution of amino acids 285–290 with alanine (285–290A) led to an accumulation of the mutated protein in the ER, whereas the 291–297A variant had no effect on receptor localization as demonstrated by the colocalization of the mutant protein with the PM marker Na^+/K^+ ATPase (Fig. 4*A*). In line

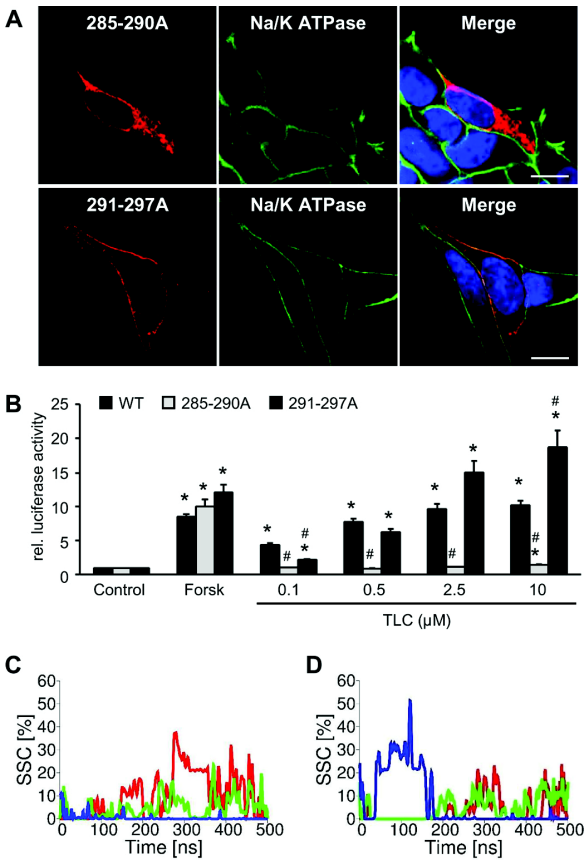


FIGURE 4. **Analysis of TGR5 alanine substitution variants.** HEK293 cells were transiently transfected with the TGR5 (pcDNA3.1+) alanine substitution variants at aa 285–290 and 291–297. *A*, the TGR5 protein was detected with the anti-TGR5 antibody (RVLR2, *red*). The variant 285–290A accumulated in the ER, whereas the 291–297 variant was mainly present in the PM as demonstrated by the yellow coloring in the overlay image with the PM marker protein Na^+/K^+ ATPase (*green*). Nuclei were stained with Hoechst (*blue*). *Bars* = 10 μm. *B*, receptor responsiveness towards TLC was measured using a cAMP-responsive luciferase assay. Alanine substitutions from 285–290 completely abolished receptor activity towards TLC stimulation. Although the 291–297A variant showed similar luciferase activity as the WT protein after stimulation with 0.5 and 2.5 μM TLC, incubation with 0.1 μM led to a significant reduction. Interestingly, stimulation of the 291–297A variant with 10 μM TLC caused a significant increase in luciferase activity as compared with the TGR5 WT. Results (285–290A $n = 9$; 291–297A $n = 13$, WT $n = 43$) are expressed as the mean ± S.E. * and #, significantly different ($p < 0.01$) from DMSO (Control) and TGR5 WT, respectively. *C* and *D*, secondary structure content (*SSC*) of conformations extracted from molecular dynamics trajectories, expressed as the percentage of amino acids of the simulated peptides showing an α-helix (*red*), 3_{10}-helix (a right-handed helix where C═O···H-N hydrogen bonds occur between residues $i \rightarrow i + 3$) (*green*), or β-sheet (*blue*). The secondary structure was assigned using the continuous extension version of the DSSP algorithm (35). For reasons of clarity, the curves were smoothed with a Bezier function. The TGR5 wild type (*C*) shows, on average, 10.2 ± 2.1% (mean ± S.E.) and up to 37.8% α-helical content; the 285–290A variant (*D*) shows, on average, 3.6 ± 0.5% α-helical content and up to 51% β-sheet formation within the first 120 ns.

with the intracellular localization, 285–290A almost completely lost the responsiveness towards TLC. Only when stimulated with 10 μM TLC was an increase in luciferase activity by 1.4 ± 0.1-fold ($n = 9$) observed for this variant. In contrast, after stimulation with TLC, 291–297A showed a similar dose-dependent increase in luciferase activity as WT. At a concentration of 10 μM TLC the rise in luciferase activity measured for 291–297A (18.7 ± 2.4; $n = 13$) was even significantly higher than the luciferase activity obtained with the WT (10.1 ± 0.7; $n = 43$) (Fig. 4*B*).

Secondary Structure Content of Membrane-proximal C-terminal Peptides Predicted from Molecular Dynamics Simulations—Because the membrane-proximal part of the TGR5 C terminus does not contain a known ER exit signal (di-acidic motif, FX_6LL or EX_3LL motifs) (24–26, 39) and because the deletion variants as well as the 291–297A variant were able to reach the PM, we hypothesized that it is the secondary structure of the membrane-proximal C terminus rather than a specific amino acid sequence that determines the localization and function of the receptor. To investigate the secondary structure content of the membrane-proximal C terminus of the WT and the variants, peptides excised from the C termini were subjected to MD simulations in explicit solvent over 600 ns of simulation time, of which the last 500 ns were used for analysis. The WT and the mutant peptides had a length of 18 aa, whereas the Δ285–290 and Δ291–297 variants had a length of 12 and 11 aa, respectively. As a starting structure, a straight peptide conformation was used, in that way avoiding any structural bias in the setup of the simulations. Fig. 4, *C* and *D*, show the secondary structure content over the course of the MD simulations for WT and 285–290A, respectively, which was retained in the ER. For the WT, an α-helix content of, on average, 10.2 ± 2.1% (mean ± S.E.; $n = 3$) with a maximum of 37.8% was observed, with only a low tendency (0.5 ± 0.2%; mean ± S.E.; $n = 3$) to form antiparallel β-sheets (Fig. 4*C*). In contrast, for 285–290A, a predominant β-sheet formation during 120 ns of the MD simulation (maximum content of 51.0%) was observed, with only an overall low tendency to form α-helices (3.6 ± 0.5%; $n = 3$) (Fig. 4*D*). These results suggest that trafficking of TGR5 from the ER to the PM requires the formation of a membrane-proximal α-helix.

PM Localization and Function of Membrane-proximal C Termini with Glycine and Proline Substitution—To further investigate the influence of the secondary structure content of the membrane-proximal TGR5 C terminus on PM localization and function of the receptor, substitution variants were generated where amino acids 285–290 and 291–297 were replaced by glycine, which has a low helix propensity (40), and proline, which is a potent helix breaker (41, 42), respectively. The 285–290P variant accumulated in the ER, whereas the 285–290G variant was present in the ER and the PM (Fig. 5*A*). Similar to 285–290A, the 285–290P and 285–290G variants almost completely lost responsiveness towards TLC as measured by luciferase activity (Fig. 5*B*). Only when stimulated with 10 μM TLC did the 285–290P and 285–290G variants show a small but significant increase in luciferase activity by 1.5 ± 0.1-($n = 8$) and 3.3 ± 0.3-fold ($n = 8$), respectively (Fig. 5*B*). Thus, the exchange of amino acids 291–297 to glycine (291–297G) and proline (291–297P) resulted in retention of the mutated proteins in the ER and abolished their responsiveness towards TLC stimulation (Fig. 5*B*).

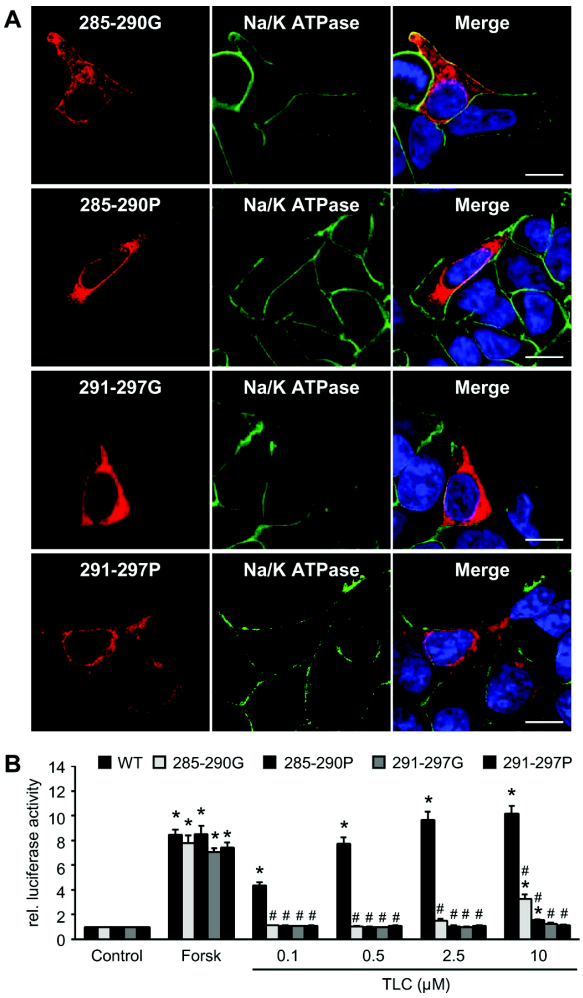


FIGURE 5. **Analysis of TGR5 glycine and proline substitution variants in HEK293 cells.** HEK293 cells were transiently transfected with the different TGR5 (pcDNA3.1+) glycine and proline substitution variants at aa 285–290 and 291–297. *A*, the TGR5 protein was detected with the anti-TGR5 antibody (RVLR2, *red*). An antibody against Na^+/K^+ ATPase (*green*) was used to stain the plasma membrane. The variants 291–297G, 285–290P, and 291–297P were mainly localized in the ER, whereas the 285–290G variant was also present in the PM as demonstrated by the *yellow coloring in the overlay image* with the PM marker protein Na^+/K^+ ATPase. Nuclei were stained with Hoechst (*blue*). *Bars* = 10 μm. *B*, receptor responsiveness towards TLC was measured using a cAMP-responsive luciferase assay. All four substitution variants showed a significantly reduced responsiveness towards TLC. Although stimulation with 10 μM TLC resulted in a 3.3 ± 0.3-fold increase in luciferase activity by the 285–290G variant, the variants 285–290P, 291–297G, and 291–297P caused an almost complete loss of receptor function. Results (285–290G, 285–290P, 291–297P $n = 8$; 291–297G $n = 9$, WT $n = 43$) are expressed as the mean ± S.E. * and # = significantly different ($p < 0.01$) from DMSO (Control) and TGR5 WT, respectively.

Clustering of MD-generated Conformations According to the Secondary Structure Sequence—To identify similarities and differences in the secondary structure of the TGR5 WT and the variants (Fig. 1) on a per-residue level, conformations of the last 500 ns of all trajectories were pooled and then hierarchically clustered according to their secondary structure sequence (Fig. 6*A*). Five main clusters were identified: conformations in cluster 1 mostly show an α-helical content at the beginning and the end of the peptides with a β-bend in between (Fig. 6*B*); conformations in cluster 2 show a short β-sheet at the beginning and the end of the peptides with a short α-helical stretch in between (Fig. 6*C*); conformations of clusters 3 and 4 show almost no defined secondary structure except for an α-helical turn of three amino acids length at the C-terminal end (cluster 4); finally, in cluster 5, most of the conformations are dominated by an α-helix formation over nine amino acids in the C-terminal region along with a turn at the N terminus (Fig. 6*D*). Forming a subcluster, conformations of variant 291–297P share the turn at the N terminus with the other members of cluster 5 but are unstructured in the C-terminal region (Fig. 6*E*).

Regarding the composition of the clusters in terms of conformations of the TGR5 variants, a distinct pattern is obvious that correlates with receptor PM localization and function (Table 1). The WT and 291–297A variants appear predominantly in clusters 1 and 5 (~60 and 20%, respectively); both variants also show the highest PM localization as determined by flow cytometry and the highest responsiveness towards TLC as measured by luciferase activity. Intermediate levels of PM localization and responsiveness towards TLC are found for the variants 285–290G, Δ285–290, and Δ291–297; these variants constitute predominantly cluster 3 (≥38%), with conformations of 285–290G and Δ291–297 also found in clusters 1 and 5. In contrast, conformations of the 285–290A and 291–297G variants occur predominantly in cluster 2 (≥76%); both variants show low PM localizations and an almost complete loss of responsiveness towards TLC. Variants 285–290P and 291–297P were similarly retained in the ER and were non-functional; they constitute predominantly cluster 4 (≥30%), with conformations also found in cluster 1 and 5.

PM Localization and Function of TGR5β_2AR, TGR5S1P1, and TGR5κOR Chimeras—To investigate whether indeed mere secondary structure content rather than a specific amino acid sequence determines PM localization and function, a chimeric receptor was constructed with the 13 membrane-proximal amino acids of the C terminus of the β_2AR (aa 330–342) replacing the respective amino acids in TGR5 (aa 285–297). These β_2AR residues were chosen because (i) they form an α-helix in the β_2AR crystal structure (PDB code 3D4S) (43, 44) and (ii) the sequence identity between the exchanged amino acids from TGR5 and β_2AR is 0%, whereas the sequence similarity is 33%.

In transiently transfected HEK293 cells the TGR5β_2AR chimera was localized in the PM, as demonstrated by the colocalization with the PM marker protein Na^+/K^+ ATPase as well as in the ER, as shown by the colocalization with the ER marker protein disulfide isomerase (Fig. 7, *A* and *B*). Using flow cytometry analysis, the percentage of transfected HEK293 cells with the chimeric receptor in the PM was calculated to be 87.4 ± 1.2% ($n = 21$), which is similar to the WT (91.3 ± 0.8%; $n = 30$) (Fig. 7*C*). MD simulations of the membrane-proximal 18 amino acids of the C terminus of the TGR5β_2AR chimera revealed an α-helix content of, on average, 24.2% (Table 2) with a maximum of 77.7%, which is higher than what has been found for the respective amino acids of the TGR5 WT. Hence, the most frequently occurring conformation of the TGR5β_2AR chimera is

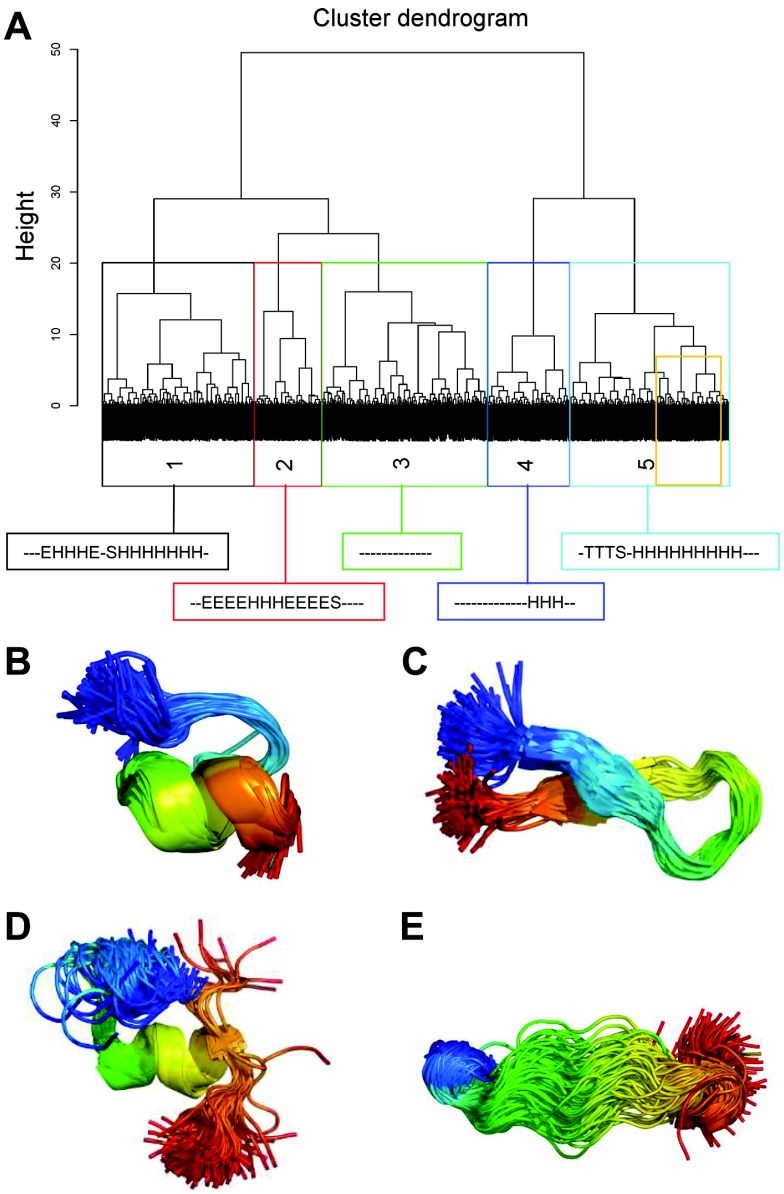


FIGURE 6. **Clustering of conformations from MD simulations according to secondary structure.** *A*, dendrogram obtained by hierarchical clustering of MD-generated conformations of the TGR5 WT and all substitution and deletion variants according to the secondary structure sequence. The most frequently occurring secondary structure sequence is shown beneath each cluster (*E*, extended β-strand; *H*, α-helix; *S*, bend; *T*, α-helical turn). Clusters 1 and 5 contain conformations with α-helix formation predominantly at the C terminus; cluster 2 contains conformations that predominantly form a β-sheet; clusters 3 and 4 contain conformations that largely lack secondary structure. A subcluster of cluster 5 containing mostly conformations of the 291–297P variant is indicated by an *orange rectangle*. *B*, conformation with the most frequently occurring secondary structure sequence in cluster 1. *C*, conformation with the most frequently occurring secondary structure sequence in cluster 2. *D*, conformation with the most frequently occurring secondary structure sequence belonging to the 291–297A variant in cluster 5. This conformation shows a pronounced α-helix formation at the C terminus. *E*, conformation with the most frequently occurring secondary structure sequence in the subcluster of cluster 5. Except for a turn at the N terminus, this conformation is unstructured otherwise. *B–E*, the coloring indicates the sequence from the N terminus (*blue*) to the C terminus (*red*).

an α-helix (Fig. 8*D*). Functional analysis with 10 μM TLC showed a receptor function of the TGR5β_2AR chimera that is similar to the one of WT. However, when TGR5β_2AR was stimulated with lower TLC concentrations, the rise in luciferase activity was significantly lower than in the case of WT (Fig. 7*D*).

Because the exchanged region of β_2AR contains an FX_6LL motif, previously identified as a sorting motif essential for the trafficking of the β_2AR from the ER to the PM (45), chimera variants containing mutations in this motif were evaluated (Fig. 8*A*). Mutation of the phenylalanine 287 to tyrosine (F287Y)

TABLE 1

Results of clustering according to secondary structure sequence, function, and protein localization of the TGR5 membrane-proximal C terminus

Variant	Cluster[1] 1	2	3	4	5	Function[2]	Membrane localization[3]
TGR5WT	56	1	8	10	25	100.0 ± 7	90.9 ± 0
285-290A	5	85	7	1	2	13.3 ± 1	52.8 ± 3
285-290G	57	1	42	0	0	30.3 ± 3	69.7 ± 4
285-290P	31	0	16	53	0	14.4 ± 1	55.0 ± 3
Δ285-290	16	0	71	3	10	36.6 ± 2	70.8 ± 3
291-297A	23	0	15	0	62	174.1 ± 23	80.3 ± 4
291-297G	13	76	5	1	5	11.9 ± 1	48.1 ± 4
291-297P	0	0	12	30	58	10.5 ± 1	40.7 ± 4
Δ291-297	6	0	38	25	31	107.5 ± 6	70.3 ± 2

[1] Percentage of the cluster distribution for each variant.
[2] Function at 10 μM TLC as percent of wildtype ± S.E.
[3] Percentage of cell membrane localization determined by FACS analysis ± S.E.

TABLE 2

α-Helicality, β-sheet content, function, and protein localization of the TGR5β_2AR mutant membrane-proximal C terminus

Chimera variant	α-Helix[1]	β-Sheet[1]	Function[2]	Membrane localization[3]
hTGR5β2AR	24.2	0.0	100.0 ± 13	87.4 ± 1
F287Y	23.0	0.5	94.5 ± 15	86.8 ± 1
L294V	10.4	0.2	125.5 ± 25	88.0 ± 1
L295V	4.5	0.6	86.1 ± 15	88.1 ± 1
L294V/L295V	1.9	6.2	83.5 ± 15	84.2 ± 1
F287Y/L294V/L295V	11.2	0.3	123.9 ± 15	89.2 ± 3
L294A/L295A	0.9	8.0	43.1 ± 10*	79.7 ± 2*
F287A/L294A/L295A	8.0	4.6	29.6 ± 5*	74.2 ± 4*
F291A/L294A/L295A	3.9	1.6	91.6 ± 9	85.1 ± 3

[1] Secondary structure content averaged over all residues throughout the MD simulations; in %.
[2] Function at 10 μM TLC as percent of TGR5β_2AR ± S.E.
[3] Percentage of cell membrane localization determined by FACS analysis ± S.E.
* Significantly different from TGR5β_2AR $p < 0.01$.

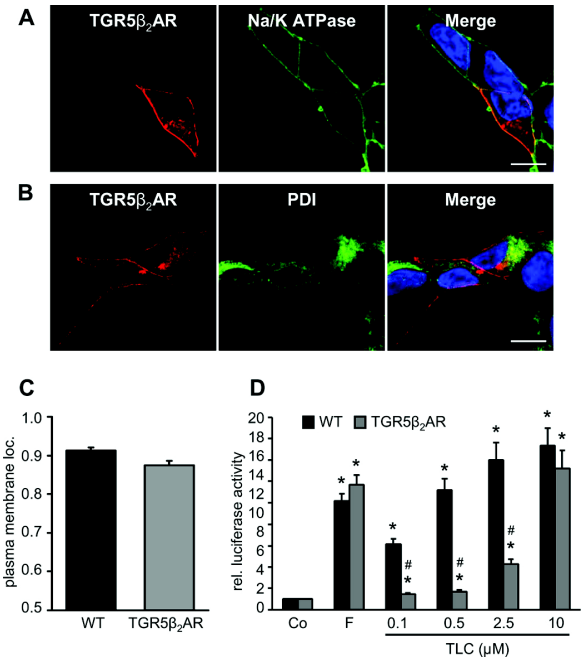


FIGURE 7. **Analysis of TGR5β_2AR chimera in HEK293 cells.** HEK293 cells were transiently transfected with the TGR5β_2AR pcDNA3.1+ (*A*, *B*, and *D*) and FLAG-TGR5β_2AR-YFP (*C*) chimera. *A* and *B*, TGR5 was stained with the anti-TGR5 (RVLR2, *red*) antibody. The TGR5β_2AR chimera was localized in the PM as demonstrated by the colocalization with the PM marker Na^+/K^+ ATPase (*A*, *green*) as well as in the ER, as shown by the colocalization with the ER marker protein disulfide isomerase (*PDI*) (*B*, *green*). Nuclei were stained with Hoechst (*blue*). Bars = 10 μm. *C*, quantification of PM localization of WT and TGR5β_2AR chimera proteins as determined by flow cytometry of unpermeabilized HEK293 cells. The amount of FLAG-TGR5-YFP within the PM corresponds to the amount of FLAG tag-positive labeling (extracellular labeling) divided by the total amount of YFP fluorescence. 91.3% of the WT and 87.4% of the TGR5β_2AR chimera were detected in the PM. Results (WT, $n = 30$; TGR5β_2AR, $n = 21$) are expressed as mean ± S.E. *D*, TGR5 receptor activity was determined by the cAMP-responsive luciferase assay. The TGR5β_2AR chimera showed a similar rise in luciferase activity in response to 10 μM TLC as the WT. However, at lower TLC concentrations (0.1, 0.5, 2.5 μM) the luciferase activity was significantly reduced compared with the WT. Results (WT $n = 29$; TGR5β_2AR $n = 21$) are expressed as mean ± S.E. * and # = significantly different ($p < 0.01$) from DMSO (Co) and WT, respectively.

resulted in a PM localization of 86.8 ± 1.1% ($n = 19$), which was significantly less than the membrane content of TGR5 WT (91.3 ± 0.8%; $n = 30$) but was indistinguishable to the cell surface levels of the TGR5β_2AR chimera (87.4 ± 1.2%; $n = 21$) (Fig. 8*B*; Table 2). This is supported by the MD simulation of the F287Y variant showing α-helix (23.0%) and β-sheet (0.5%) contents almost identical to the ones of TGR5β_2AR. Mutation of either leucine 294 or 295 to valine (L294V or L295V) did not affect PM localization or luciferase activity after stimulation with 10 μM TLC. Mutation of the FX_6LL motif to YX_6VV resulted in a TGR5β_2AR(F287Y/L294V/L295V) chimera that showed PM localization levels and functional responsiveness towards 10 μM TLC indistinguishable from the TGR5 WT and the TGR5β_2AR chimera (Table 2). MD simulations demonstrate α-helix contents between 4.5 and 11.2% for the F287Y/L294V/L295V, L294V, and L295V variants but a low β-sheet content (< 0.6%) as well.

Exchange of both leucines to valine (L294V/L295V) or alanine (L294A/L295A) reduced plasma membrane localization to 84.2 ± 1.1% ($n = 23$) and 79.7 ± 2.0% ($n = 16$), respectively (Fig. 8*C*; Table 2). Mutation of the phenylalanine and the leucine residues in the FX_6LL motif to alanine (AX_6AA, F287A/L294A/L295A) markedly compromised ER to PM trafficking (74.2 ± 3.9% PM localization, $n = 10$) as well as function. For these variants MD simulations revealed a high β-sheet content of 4.6–8.0% (Table 2). Also, the most frequently occurring conformation of L294A/L295A shows a β-sheet formation (Fig. 8*E*).

Visual inspection of the conformations of the L294A/L295A variant with high β-sheet content revealed that the β-sheet is stabilized by a hydrophobic contact between phenylalanine 291 and alanine 295. This suggested the replacement of phenylalanine 291 with alanine to break this contact. Accordingly, MD simulations predicted a low β-sheet content (1.6%) but an increased α-helix content (3.9%) for the variant where phenylalanine 291 and leucines 294 and 295 are mutated to alanine (F291A/L294A/L295A). Its most frequently occurring conformation also shows an α-helical stretch (Fig. 8*F*). Thus, we expected that with this variant PM localization and functional activity will be restored. Indeed, the F291A/L294A/L295A variant showed a PM localization level and luciferase activity in

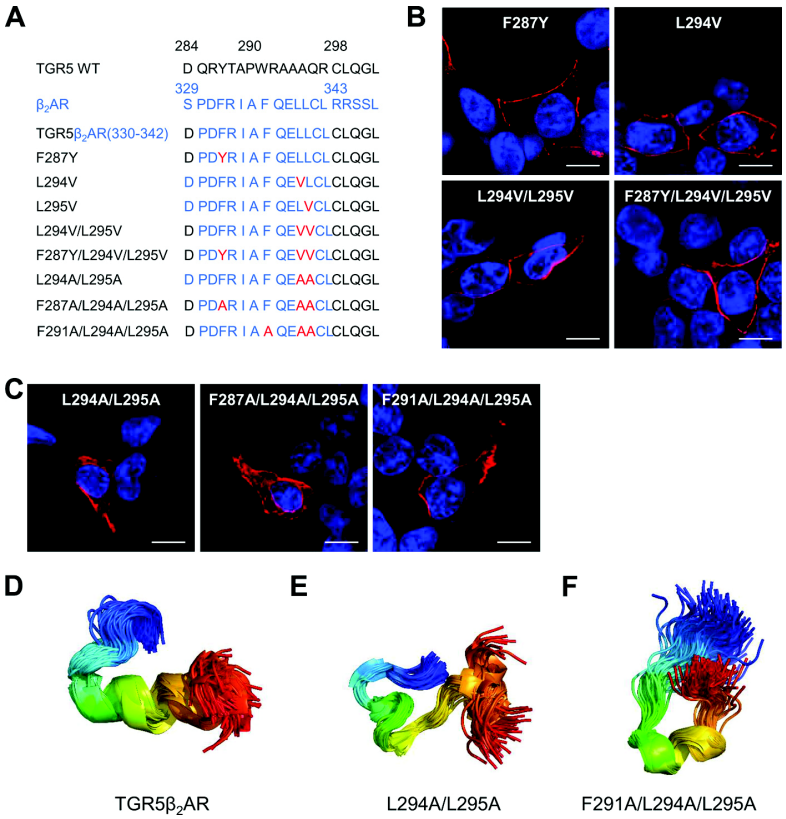


FIGURE 8. **Effect of mutations in the FX_6LL motif on the localization, function, and structure of TGR5β_2AR chimera.** *A*, sequence alignment of the membrane-proximal C terminus regions of TGR5 WT, β_2AR, and TGR5β_2AR chimera. Mutations of the FX_6LL motif within the TGR5β_2AR chimera are marked in *red*. *B* and *C*, localization of the mutated TGR5β_2AR chimeras in transiently transfected HEK293 cells was revealed by detecting TGR5 with the RVLR2 antibody (*red*). Conservative amino acid mutations (Phe to Tyr and Leu to Val) have only minor effects on receptor plasma membrane localization. However, substitution of Leu to Ala in L294A/L295A and Phe to Ala as well as Leu to Ala in F287A/L294A/L295A impaired trafficking of the mutated proteins, which can be detected in both the ER as well as the plasma membrane of transfected cells. Nuclei were stained with Hoechst (*blue*). *Bars* = 10 μm. *D*, most frequently occurring conformation of the TGR5β_2AR chimera C terminus as found by MD simulations. *E*, most frequently occurring conformation of the C terminus of the L294A/L295A variant as found by MD simulations. *F*, most frequently occurring conformation of the C terminus of the F291A/L294A/L295A mutant. *D–F*, the coloring indicates the sequence from the N terminus (*blue*) to the C terminus (*red*).

response to 10 μM TLC indistinguishable from the TGR5β_2AR chimera (Table 2).

To further support the hypothesis that the secondary structure formation of the membrane-proximal C terminus is essential for TGR5 ER to PM trafficking, we generated two additional receptor chimeras in which 12 or 10 amino acids of TGR5 (aa 285–296, aa 285–294) were replaced by membrane-proximal amino acids of the κOR isoform-1 (aa 335–346) or the S1P1 (aa 316–325), respectively. These κOR and S1P1 residues were chosen because (i) they form an α-helix in the crystal structures (PDB codes 4DJH and 3V2Y, respectively) (46, 47), (ii) the sequence identity (similarity) between the exchanged amino acids from TGR5 and those from κOR is 8% (33%) and those from S1P1 is 0% (30%), and (iii) they lack the FX_6LL sorting motif.

In transiently transfected HEK293 cells the TGR5κOR and TGR5S1P1 chimeras were sorted to the PM as shown by the colocalization with the PM marker protein Na^+/K^+ ATPase (Fig. 9, *A* and *B*). Using flow cytometry analysis, the percentage of transfected HEK293 cells with chimeric TGR5κOR and TGR5S1P1 receptors in the PM were 91.1 ± 1.5% (n = 15) and 86.3 ± 1.7% (n = 10), respectively (Fig. 9*C*). Functional analysis with 10 μM TLC showed a receptor function of the TGR5κOR and TGR5S1P1 chimeras that was similar to the one of TGR5 WT. However, when TGR5κOR was stimulated with lower TLC concentrations, the rise in luciferase activity was significantly lower than in the case of TGR5 WT (Fig. 9*D*). In contrast, luciferase activity of the TGR5S1P1 chimera was indistinguishable from the WT when stimulated with 0.1–10 μM of TLC (Fig. 9*E*). Taken together, these results confirm the hypothesis that an α-helical structure in the proximal part of the C terminus is required for TGR5 PM localization and responsiveness toward extracellular ligands.

DISCUSSION

The molecular mechanisms involved in the trafficking of GPCRs from the ER to the PM are incompletely understood (24). Furthermore, the determinants required for ER exit seem to vary between different GPCRs (24). A naturally occurring truncation variant in the TGR5 C terminus (Q296*X*) led to ER

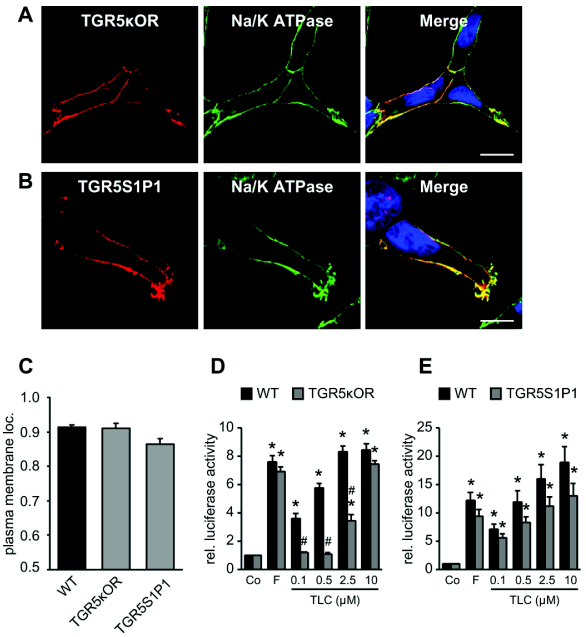


FIGURE 9. **Analysis of TGR5κOR and TGR5S1P1 chimeras in HEK293 cells.** HEK293 cells were transiently transfected with the TGR5κOR pcDNA3.1+ and TGR5S1P1 pcDNA3.1+ chimeras (*A*, *B*, *D*, *E*) or with the FLAG-TGR5κOR-YFP and FLAG-TGR5S1P1-YFP chimeras (*C*). *A* and *B*, the TGR5 chimeric proteins were detected with the anti-TGR5 antibody (RVLR2, *red*). An antibody against Na^+/K^+ ATPase (*green*) was used to stain the PM. Both chimeras TGR5κOR (*A*) and TGR5S1P1 (*B*) were localized in the PM as demonstrated by the *yellow coloring* in the overlay image with the PM marker protein Na^+/K^+ ATPase. *C*, the PM localization of WT and TGR5 chimeras was quantified as described under "Experimental Procedures" using flow cytometry. TGR5κOR and TGR5S1P1 were detected on the cell surface in 91.1 and 86.3% of the transfected cells, which was similar to the TGR5 WT with 91.3% (WT, $n = 30$; TGR5κOR, $n = 15$; TGR5S1P1, $n = 10$). *D* and *E*, receptor responsiveness towards TLC was measured using a cAMP-responsive luciferase assay. Although the TGR5κOR chimera showed a significantly reduced responsiveness toward lower TLC concentrations (0.1, 0.5, 2.5 μM) and no difference at 10 μM TLC compared with the WT (*D*, WT $n = 8$; TGR5κOR $n = 7$), the TGR5S1P1 chimera showed luciferase activity indistinguishable from the WT (*E*, WT $n = 6$; TGR5S1P1 $n = 8$). Results are expressed as mean ± S.E. *, #, = significantly different ($p < 0.01$) from DMSO (Co) and TGR5 WT, respectively.

retention and abolished ligand responsiveness (20), indicating that the C terminus of TGR5 is required for PM localization and function of this bile acid receptor. Using different substitution and deletion variants within the TGR5 C terminus, the present study demonstrates that the secondary structure of the membrane-proximal amino acids 285–297 has a strong influence on ER to PM trafficking of the TGR5 receptor.

Employing MD simulations that are 3-fold longer than the time required for α-helix formation as determined from experiments (48–50), we showed that the wild type peptide encompassing residues 285–297 preferentially forms an α-helix. In contrast, the 285–290A variant, which was retained in the ER, showed an exclusive β-sheet formation within the first 120 ns of the simulation. This finding was unexpected because alanine has a high helix propensity (40). However, the β-sheet formation seemed to be favored in this case by interactions between the alanine residues in positions 285–290 with naturally occurring hydrophobic alanines in positions 294–296. These initial analyses suggested that a high β-sheet content in the membrane-proximal C terminus prevents ER to PM trafficking of TGR5. In contrast, formation of an α-helix, located in TGR5 at a position equivalent to helix 8 in crystallographically determined GPCR structures (44, 51, 52), facilitates it.

To further support our findings, the localization, functional, and MD simulation analyses were extended to additional mutation and deletion variants in the C terminus of TGR5. As the most outstanding result, a clear correlation between the secondary structure sequence of the respective peptide and the localization of TGR5 and, thus, its function emerged. As such, the variant 291–297A most strongly resembles the WT with respect to the secondary structure sequence, as determined by cluster analysis of all MD-generated conformations; both preferentially occur in clusters 1 and 5 (Table 1). 291–297A is also correctly sorted to the PM and shows a functional activity that is even larger than that of the WT when stimulated with 10 μM TLC. As a possible explanation for the increased activity, ~⅔ of the conformations of 291–297A occur in cluster 5, which largely contains conformations of other highly active variants, too. This suggests that the 291–297A variant has a particularly high probability to form a favorable conformation. Likewise, about ⅓ of all generated conformations of the Δ291–297 variant, which shows a similarly high PM localization and functional activity as 291–297A, are found in cluster 5. The conformations in clusters 1 and 5 are dominated by long α-helical stretches at the C termini of the peptides.

A second major fraction of the Δ291–297 conformations is found in cluster 3; conformations in cluster 3 are largely unstructured. Likewise, major fractions of the conformations of the Δ285–290 and 285–290G variants are found in this cluster, with other major fractions found in cluster 1. This reduced conformational preference to form an α-helix of the Δ285–290 and 285–290G variants goes along with experimental findings of a reduced PM localization and functional activity.

Finally, of the least functional variants that also showed the highest retention in the ER, 291–297G mostly strongly resembles the variant 285–290A in terms of the formation of a β-sheet (appearance of the major fraction of the conformations in cluster 2). In contrast, a major fraction of the conformations of the variant 291–297P show a short stretch of α-helix in the C-terminal region (cluster 4), which is apparently not sufficient to lead to ER to PM trafficking. To our surprise, the largest fraction of the conformations of the 291–297P variant was found in cluster 5. Yet these conformations form a subcluster there and are characterized by an N-terminal TTTS motif but lack the long α-helical stretch at the C terminus, in agreement with experimental findings that the 291–297P variant is largely retained in the ER. As to the last variant, most of the conformations of 285–290P are found in cluster 4, too, again in agreement with the findings that this variant is retained in the ER. However, another major fraction is found in cluster 1, which would lead to the false prediction that the 285–290P variant will be sorted to the PM. In that respect, 285–290P is the sole variant investigated that does not fit to the relationship between the secondary structure sequence of a respective peptide and the localization of TGR5.

Taken together, these results reveal that PM localization and, thus, function of the receptor is fostered by a long (≥9 residues)

α-helical stretch at the C terminus of the variants, whereas the presence of β-strands or only a short α-helical stretch leads to ER retention and a loss of function.

As a proof-of-principle, a chimera of TGR5 containing the membrane-proximal amino acids of the β_2AR was generated. The respective amino acid sequence of β_2AR forms an α-helix as shown in high resolution crystal structures (43, 44). MD simulations of the membrane-proximal 18 amino acids of the C terminus of the TGR5β_2AR chimera, which contains 13 amino acids of the membrane-proximal C terminus of β_2AR, reproduced this α-helix character (Fig. 8*D*), which demonstrates the quality of the setup of our simulations. We presumed that the TGR5β_2AR chimera would be mainly localized in the PM due to the α-helical structure of its C terminus. The TGR5β_2AR chimera was indeed correctly sorted to the PM and showed similar functional activity in response to 10 μM TLC as WT TGR5.

However, the membrane-proximal part of the β_2AR contains an FX_6LL motif that has previously been identified as an important GPCR ER export motif facilitating interaction of the β_2AR with the GTPase Rab1, that way promoting ER to PM trafficking (45, 53). To assess whether the FX_6LL motif promotes membrane trafficking of the TGR5β_2AR chimeric receptor, too, we evaluated chimera variants containing mutations in this motif. This again revealed a strong correlation between the secondary structure content of the respective variant and its localization and function. First, the conservative replacement of phenylalanine by tyrosine in the F287Y variant does not change the α-helix content with respect to the WT; in agreement, this variant also showed a high level of membrane localization and retained functional activity. Second, variants with conservative mutations at positions 294 and 295 revealed a reduced α-helical content compared with the WT and negligible β-sheet content (F287Y/L294V/L295V, L294V, and L295V) and showed intermediate levels of membrane localization and functional activity. Third, the alanine mutants F287A/L294A/L295A and L294A/L295A and the valine double mutant L294V/L295V revealed the lowest α-helical contents but pronounced β-sheet formation. As in the case of the TGR5 variants, these chimera variants showed marked retention in the ER. Finally, we created the F291A/L294A/L295A variant guided by insights from structural analysis and MD simulations that this variant has a lower β-sheet and a higher α-helical content. The variant indeed showed PM localization level and luciferase activity in response to 10 μM TLC indistinguishable from the TGR5β_2AR chimera.

Two additional chimeras in which 12 or 10 amino acids of TGR5 were replaced by membrane-proximal amino acids of κOR or S1P1, respectively, showed PM localization and functional activity very similar to the TGR5 WT and the TGR5β_2AR chimera. Although there is only a very low or even no sequence identity between the exchanged residues, respectively, the κOR or S1P1 residues do form an α-helix in the respective crystal structures (46, 47).

In summary, these results demonstrate that PM trafficking and, thus, function of the TGR5 chimera are determined by the secondary structure of the membrane-proximal C terminus. As found for variants of TGR5 itself, a high α-helix content of this region fosters PM localization, whereas a high β-sheet content affects PM localization adversely. Furthermore, the modifications of the FX_6LL motif in the TGR5β_2AR chimera demonstrated that trafficking to the PM occurs as along as the modifications retain a high α-helix content in the membrane-proximal C terminus. Using the PDBeMotif webservice (54) the frequency of occurrence of the FX_6LL motif in secondary structures was analyzed in all non-redundant proteins within the Protein Data Bank. This revealed that the FX_6LL motif occurs in helices in 71% of the cases, whereas it is found in sheets and loops in 5 and 24% of cases, respectively. Taken together our results suggest that the FX_6LL motif itself has an important role for α-helix formation but not as a sequence-specific sorting motif in the TGR5β_2AR chimera.

Similar findings have been described for the human cannabinoid receptor 1 (27, 55, 56). Using overexpression, purification, and circular dichroism spectroscopy analysis of the full-length CB1 C terminus as well as different substitution variants, it was demonstrated that it is the level of hydrophobicity rather than specific amino acids that is critical for helix formation and, thus, targeting to the PM of this GPCR (27, 55, 56). Another example that hydrophobic amino acid residues in the membrane-proximal C terminus allowing α-helix formation are essential for ER exit and PM localization has been described for the human vasopressin II receptor (28). However, in contrast to our finding with the TGR5β_2AR chimera, replacement of the C-terminal tail of the vasopressin II receptor with amino acids 327–413 of the human β_2AR resulted in a chimera that failed to reach the PM (22). It is unclear whether this discrepancy is dependent on special amino acid residues in the membrane proximal C terminus or on the different cloning strategies for the chimera. In the TGR5β_2AR chimera, 13 amino acids (aa 330–342) of the cytoplasmic C-tail of the β_2AR replaced the respective 13 amino acids of TGR5 C terminus (aa 285–297). In contrast, in the vasopression 2 receptor β_2AR chimera, the complete C-terminal tail of the β_2AR (aa 327–413) was fused to alanine 325 of the human vasopressin II receptor, which is predicted to be part of transmembrane domain 7 (22).

Although an important role of the C terminus anterograde trafficking from the ER to the PM has been described for many different GPCRs, it remained unclear whether specific amino acid signals or, rather, a certain conformation are essential for successful ER export (24). Most likely the structural or amino acid sequence determinants required for correct sorting to the PM vary for different GPCRs. For TGR5, which lacks any of the previously described ER exit motifs (di-acidic motif, FX_6LL or EX_3LL motifs) (24–26, 39), our study demonstrates that the secondary structure of the membrane-proximal C terminus, which forms an α-helix according to the MD simulations, is essential for PM localization and, thus, function of this bile acid receptor.

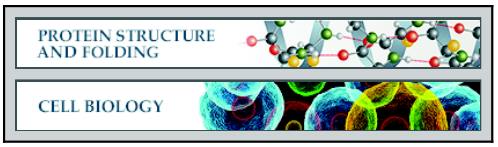
REFERENCES

1. Kawamata, Y., Fujii, R., Hosoya, M., Harada, M., Yoshida, H., Miwa, M., Fukusumi, S., Habata, Y., Itoh, T., Shintani, Y., Hinuma, S., Fujisawa, Y., and Fujino, M. (2003) A G protein-coupled receptor responsive to bile acids. *J. Biol. Chem.* **278**, 9435–9440
2. Maruyama, T., Miyamoto, Y., Nakamura, T., Tamai, Y., Okada, H., Sugiyama, E., Nakamura, T., Itadani, H., and Tanaka, K. (2002) Identifica-

tion of membrane-type receptor for bile acids (M-BAR). *Biochem. Biophys. Res. Commun.* **298,** 714–719
3. Maruyama, T., Tanaka, K., Suzuki, J., Miyoshi, H., Harada, N., Nakamura, T., Miyamoto, Y., Kanatani, A., and Tamai, Y. (2006) Targeted disruption of G protein-coupled bile acid receptor 1 (Gpbar1/M-Bar) in mice. *J. Endocrinol.* **191,** 197–205
4. Vassileva, G., Golovko, A., Markowitz, L., Abbondanzo, S. J., Zeng, M., Yang, S., Hoos, L., Tetzloff, G., Levitan, D., Murgolo, N. J., Keane, K., Davis, H. R., Jr., Hedrick, J., and Gustafson, E. L. (2006) Targeted deletion of Gpbar1 protects mice from cholesterol gallstone formation. *Biochem. J.* **398,** 423–430
5. Keitel, V., Görg, B., Bidmon, H. J., Zemtsova, I., Spomer, L., Zilles, K., and Häussinger, D. (2010) The bile acid receptor TGR5 (Gpbar-1) acts as a neurosteroid receptor in brain. *Glia* **58,** 1794–1805
6. Sato, H., Macchiarulo, A., Thomas, C., Gioiello, A., Une, M., Hofmann, A. F., Saladin, R., Schoonjans, K., Pellicciari, R., and Auwerx, J. (2008) Novel potent and selective bile acid derivatives as TGR5 agonists. Biological screening, structure-activity relationships, and molecular modeling studies. *J. Med. Chem.* **51,** 1831–1841
7. Keitel, V., and Häussinger, D. (2012) Perspective. TGR5 (Gpbar-1) in liver physiology and disease. *Clin. Res. Hepatol. Gastroenterol.* **36,** 412–419
8. Keitel, V., Donner, M., Winandy, S., Kubitz, R., and Häussinger, D. (2008) Expression and function of the bile acid receptor TGR5 in Kupffer cells. *Biochem. Biophys. Res. Commun.* **372,** 78–84
9. Keitel, V., Reinehr, R., Gatsios, P., Rupprecht, C., Görg, B., Selbach, O., Häussinger, D., and Kubitz, R. (2007) The G-protein coupled bile salt receptor TGR5 is expressed in liver sinusoidal endothelial cells. *Hepatology* **45,** 695–704
10. Keitel, V., Cupisti, K., Ullmer, C., Knoefel, W. T., Kubitz, R., and Häussinger, D. (2009) The membrane-bound bile acid receptor TGR5 is localized in the epithelium of human gallbladders. *Hepatology* **50,** 861–870
11. Keitel, V., Ullmer, C., and Häussinger, D. (2010) The membrane-bound bile acid receptor TGR5 (Gpbar-1) is localized in the primary cilium of cholangiocytes. *Biol. Chem.* **391,** 785–789
12. Lavoie, B., Balemba, O. B., Godfrey, C., Watson, C. A., Vassileva, G., Corvera, C. U., Nelson, M. T., and Mawe, G. M. (2010) Hydrophobic bile salts inhibit gallbladder smooth muscle function via stimulation of GPBAR1 receptors and activation of KATP channels. *J. Physiol.* **588,** 3295–3305
13. Li, T., Holmstrom, S. R., Kir, S., Umetani, M., Schmidt, D. R., Kliewer, S. A., and Mangelsdorf, D. J. (2011) The G protein-coupled bile acid receptor, TGR5, stimulates gallbladder filling. *Mol. Endocrinol.* **25,** 1066–1071
14. Pols, T. W., Nomura, M., Harach, T., Lo Sasso, G., Oosterveer, M. H., Thomas, C., Rizzo, G., Gioiello, A., Adorini, L., Pellicciari, R., Auwerx, J., and Schoonjans, K. (2011) TGR5 activation inhibits atherosclerosis by reducing macrophage inflammation and lipid loading. *Cell Metab.* **14,** 747–757
15. Wang, Y. D., Chen, W. D., Yu, D., Forman, B. M., and Huang, W. (2011) The G-Protein-coupled bile acid receptor, Gpbar1 (TGR5), negatively regulates hepatic inflammatory response through antagonizing nuclear factor κ light-chain enhancer of activated B cells (NF-κB) in mice. *Hepatology* **54,** 1421–1432
16. Poole, D. P., Godfrey, C., Cattaruzza, F., Cottrell, G. S., Kirkland, J. G., Pelayo, J. C., Bunnett, N. W., and Corvera, C. U. (2010) Expression and function of the bile acid receptor GpBAR1 (TGR5) in the murine enteric nervous system. *Neurogastroenterol. Motil.* **22,** 814–825
17. Reimann, F., Habib, A. M., Tolhurst, G., Parker, H. E., Rogers, G. J., and Gribble, F. M. (2008) Glucose sensing in L cells. A primary cell study. *Cell Metab.* **8,** 532–539
18. Thomas, C., Gioiello, A., Noriega, L., Strehle, A., Oury, J., Rizzo, G., Macchiarulo, A., Yamamoto, H., Mataki, C., Pruzanski, M., Pellicciari, R., Auwerx, J., and Schoonjans, K. (2009) TGR5-mediated bile acid sensing controls glucose homeostasis. *Cell Metab.* **10,** 167–177
19. Pols, T. W., Noriega, L. G., Nomura, M., Auwerx, J., and Schoonjans, K. (2011) The bile acid membrane receptor TGR5 as an emerging target in metabolism and inflammation. *J. Hepatol.* **54,** 1263–1272
20. Hov, J. R., Keitel, V., Laerdahl, J. K., Spomer, L., Ellinghaus, E., ElSharawy, A., Melum, E., Boberg, K. M., Manke, T., Balschun, T., Schramm, C., Bergquist, A., Weismüller, T., Gotthardt, D., Rust, C., Henckaerts, L., Onnie, C. M., Weersma, R. K., Sterneck, M., Teufel, A., Runz, H., Stiehl, A., Ponsioen, C. Y., Wijmenga, C., Vatn, M. H., IBSEN Study Group, Stokkers, P. C., Vermeire, S., Mathew, C. G., Lie, B. A., Beuers, U., Manns, M. P., Schreiber, S., Schrumpf, E., Häussinger, D., Franke, A., and Karlsen, T. H. (2010) Mutational characterization of the bile acid receptor TGR5 in primary sclerosing cholangitis. *PLoS ONE* **5,** e12403
21. Rodriguez, M. C., Xie, Y. B., Wang, H., Collison, K., and Segaloff, D. L. (1992) Effects of truncations of the cytoplasmic tail of the luteinizing hormone/chorionic gonadotropin receptor on receptor-mediated hormone internalization. *Mol. Endocrinol.* **6,** 327–336
22. Oksche, A., Dehe, M., Schülein, R., Wiesner, B., and Rosenthal, W. (1998) Folding and cell surface expression of the vasopressin V2 receptor. Requirement of the intracellular C terminus. *FEBS Lett.* **424,** 57–62
23. Pankevych, H., Korkhov, V., Freissmuth, M., and Nanoff, C. (2003) Truncation of the A1 adenosine receptor reveals distinct roles of the membrane-proximal carboxyl terminus in receptor folding and G protein coupling. *J. Biol. Chem.* **278,** 30283–30293
24. Dong, C., Filipeanu, C. M., Duvernay, M. T., and Wu, G. (2007) Regulation of G protein-coupled receptor export trafficking. *Biochim. Biophys. Acta* **1768,** 853–870
25. Zhang, X., Dong, C., Wu, Q. J., Balch, W. E., and Wu, G. (2011) Di-acidic motifs in the membrane-distal C termini modulate the transport of angiotensin II receptors from the endoplasmic reticulum to the cell surface. *J. Biol. Chem.* **286,** 20525–20535
26. Duvernay, M. T., Zhou, F., and Wu, G. (2004) A conserved motif for the transport of G protein-coupled receptors from the endoplasmic reticulum to the cell surface. *J. Biol. Chem.* **279,** 30741–30750
27. Ahn, K. H., Nishiyama, A., Mierke, D. F., and Kendall, D. A. (2010) Hydrophobic residues in helix 8 of cannabinoid receptor 1 are critical for structural and functional properties. *Biochemistry* **49,** 502–511
28. Thielen, A., Oueslati, M., Hermosilla, R., Krause, G., Oksche, A., Rosenthal, W., and Schülein, R. (2005) The hydrophobic amino acid residues in the membrane-proximal C tail of the G protein-coupled vasopressin V2 receptor are necessary for transport-competent receptor folding. *FEBS Lett.* **579,** 5227–5235
29. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J Comput. Chem.* **26,** 1668–1688
30. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117,** 5179–5197
31. Simmerling, C., Strockbine, B., and Roitberg, A. E. (2002) All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **124,** 11258–11259
32. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935
33. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald. An N [center-dot] log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092
34. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints. Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23,** 327–341
35. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure. Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637
36. R-Development-Core-Team (2008) R, a language and environment for statistical computing. R-Foundation for Statistical Computing, Vienna, Austria
37. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6,** 175–182
38. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model. Application to complete genomes. *J. Mol. Biol.* **305,** 567–580

39. Schülein, R., Hermosilla, R., Oksche, A., Dehe, M., Wiesner, B., Krause, G., and Rosenthal, W. (1998) A dileucine sequence and an upstream glutamate residue in the intracellular carboxyl terminus of the vasopressin V2 receptor are essential for cell surface transport in COS. M6 cells. *Mol. Pharmacol.* **54,** 525–535
40. Chakrabartty, A., Schellman, J. A., and Baldwin, R. L. (1991) Large differences in the helix propensities of alanine and glycine. *Nature* **351,** 586–588
41. Pace, C. N., and Scholtz, J. M. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75,** 422–427
42. Richardson, J. S. (1981) The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **34,** 167–339
43. Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Kuhn, P., Weis, W. I., Kobilka, B. K., and Stevens, R. C. (2007) High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265
44. Rasmussen, S. G., Choi, H. J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., Schertler, G. F., Weis, W. I., and Kobilka, B. K. (2007) Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **450,** 383–387
45. Duvernay, M. T., Dong, C., Zhang, X., Zhou, F., Nichols, C. D., and Wu, G. (2009) Anterograde trafficking of G protein-coupled receptors. Function of the C-terminal FX_6LL motif in export from the endoplasmic reticulum. *Mol. Pharmacol.* **75,** 751–761
46. Hanson, M. A., Roth, C. B., Jo, E., Griffith, M. T., Scott, F. L., Reinhart, G., Desale, H., Clemons, B., Cahalan, S. M., Schuerer, S. C., Sanna, M. G., Han, G. W., Kuhn, P., Rosen, H., and Stevens, R. C. (2012) Crystal structure of a lipid G protein-coupled receptor. *Science* **335,** 851–855
47. Wu, H., Wacker, D., Mileni, M., Katritch, V., Han, G. W., Vardy, E., Liu, W., Thompson, A. A., Huang, X. P., Carroll, F. I., Mascarella, S. W., Westkaemper, R. B., Mosier, P. D., Roth, B. L., Cherezov, V., and Stevens, R. C. (2012) Structure of the human κ-opioid receptor in complex with JDTic. *Nature* **485,** 327–332
48. Lin, M. M., Mohammed, O. F., Jas, G. S., and Zewail, A. H. (2011) Speed limit of protein folding evidenced in secondary structure dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 16622–16627
49. Gnanakaran, S., Nymeyer, H., Portman, J., Sanbonmatsu, K. Y., and García, A. E. (2003) Peptide folding simulations. *Curr. Opin. Struct. Biol.* **13,** 168–174
50. De Sancho, D., and Best, R. B. (2011) What Is the Time Scale for + −helix nucleation? *J. Am. Chem. Soc.* **133,** 6809–6816
51. Orel, M., Padrós, E., and Manyosa, J. (2012) Structural features of the C terminus from the human neurokinin-1 receptor. *FEBS J.* **279,** 2357–2367
52. Wu, B., Chien, E. Y., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V., and Stevens, R. C. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330,** 1066–1071
53. Hammad, M. M., Kuang, Y. Q., Morse, A., and Dupré, D. J. (2012) Rab1 interacts directly with the β2-adrenergic receptor to regulate receptor anterograde trafficking. *Biol. Chem.* **393,** 541–546
54. Golovin, A., and Henrick, K. (2008) MSDmotif. Exploring protein sites and motifs. *BMC Bioinformatics* **9,** 312–319
55. Ahn, K. H., Pellegrini, M., Tsomaia, N., Yatawara, A. K., Kendall, D. A., and Mierke, D. F. (2009) Structural analysis of the human cannabinoid receptor one carboxyl-terminus identifies two amphipathic helices. *Biopolymers* **91,** 565–573
56. Stadel, R., Ahn, K. H., and Kendall, D. A. (2011) The cannabinoid type-1 receptor carboxyl-terminus, more than just a tail. *J Neurochem.* **117,** 1–18

Protein Structure and Folding:
A Membrane-proximal, C-terminal α-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5



Lina Spomer, Christoph G. W. Gertzen, Birte Schmitz, Dieter Häussinger, Holger Gohlke and Verena Keitel

J. Biol. Chem. 2014, 289:3689-3702.
doi: 10.1074/jbc.M113.502344 originally published online December 13, 2013

Publication II

Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study

Annemarie Greife, Suren Felekyan, Qijun Ma, Christoph G.W. Gertzen, Lina Spomer, Mykola Dimura, Thomas O. Peulen, Christina Wöhler, Dieter Häussinger, Holger Gohlke, Verena Keitel, Claus A.M. Seidel

Scientific Reports, (2016), 6, 36792

www.nature.com/scientificreports

OPEN

Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study

Received: 14 July 2016
Accepted: 21 October 2016
Published: 11 November 2016

Annemarie Greife[1], Suren Felekyan[1], Qijun Ma[1], Christoph G. W. Gertzen[2], Lina Spomer[3], Mykola Dimura[1], Thomas O. Peulen[1], Christina Wöhler[3], Dieter Häussinger[3], Holger Gohlke[2], Verena Keitel[3] & Claus A. M. Seidel[1]

TGR5 is the first identified bile acid-sensing G-protein coupled receptor, which has emerged as a potential therapeutic target for metabolic disorders. So far, structural and multimerization properties are largely unknown for TGR5. We used a combined strategy applying cellular biology, Multiparameter Image Fluorescence Spectroscopy (MFIS) for quantitative FRET analysis, and integrative modelling to obtain structural information about dimerization and higher-order oligomerization assemblies of TGR5 wildtype (wt) and Y111 variants fused to fluorescent proteins. Residue 111 is located in transmembrane helix 3 within the highly conserved ERY motif. Co-immunoprecipitation and MFIS-FRET measurements with gradually increasing acceptor to donor concentrations showed that TGR5 wt forms higher-order oligomers, a process disrupted in TGR5 Y111A variants. From the concentration dependence of the MFIS-FRET data we conclude that higher-order oligomers – likely with a tetramer organization - are formed from dimers, the smallest unit suggested for TGR5 Y111A variants. Higher-order oligomers likely have a linear arrangement with interaction sites involving transmembrane helix 1 and helix 8 as well as transmembrane helix 5. The latter interaction is suggested to be disrupted by the Y111A mutation. The proposed model of TGR5 oligomer assembly broadens our view of possible oligomer patterns and affinities of class A GPCRs.

TGR5 (GPBAR-1, M-BAR) is the first identified G-protein coupled bile acid receptor[1] and is widely expressed in tissues, including liver, intestine, and the central and enteric nervous system[2,3]. Animal studies suggest that TGR5 activation leads to anti-inflammatory effects and influences energy homeostasis and glucose metabolism, thereby playing a role in the pathogenesis of obesity and diabetes[4]. Therefore, TGR5 has emerged as a potential therapeutic target to treat metabolic disorders. The most potent TGR5 bile acid agonist is taurolithocholic acid (TLCA/TLC)[1]. In model cell lines it was shown that TGR5 couples to $G\alpha_s$, leading to stimulation of adenylate cyclase (AC) and formation of cyclic AMP (cAMP)[1].

To date, no high-resolution crystal structure of TGR5 is available, and knowledge on TGR5 regulation and oligomerization is scarce. Homology models of TGR5 have been presented based on template structures of other seven transmembrane (7TM) domain receptors[5–8]. We previously reported that the amino acids 285–294 at the TGR5 C-terminus form an alpha-helical stretch important for plasma membrane localization and thus responsiveness to extracellular ligands[9].

It is now well established that class C GPCRs form homo- and hetero-oligomers[10]. Oligomer formation of GPCRs affects a broad range of biological functions ranging from intracellular trafficking, protein turnover,

[1]Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany. [2]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany. [3]Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany. Correspondence and requests for materials should be addressed to H.G. (email: gohlke@hhu.de) or V.K. (email: verena.keitel@med.uni-duesseldorf.de) or C.A.M.S. (email: cseidel@hhu.de)

receptor function, signal enhancement or blockage upon ligand binding, G-protein independent signaling to internalization and receptor desensitization (for an overview see refs 11 and 12). However, for class A GPCRs such as TGR5, there are controversial data about the functional significance of homo- and hetero-oligomer formation[10]. Studies with rhodopsin[13,14], μ-opioid[15] and β_2-adrenergic receptors trapped as either monomers or dimers in nanodiscs demonstrated that monomers are functional and activate G-proteins; sometimes monomers are even more efficient than homo-dimers[10]. The same GPCRs were also shown to be stable as dimers or tetramers in living cells[13–15]. Many researchers proved at least dimerization by using biophysical approaches such as bioluminescence and Förster Resonance Energy Transfer techniques (BRET and FRET), as well as single molecule analysis[16] and atomic force microscopy in native disc membranes[17]. FRET describes the distance dependent energy transfer from an excited donor (D) to an acceptor (A) fluorophore and is used to study biomolecules in living cells which are fused to genetically encoded fluorescent proteins (FP) for convenience, although other molecular tags are also being used.

Several oligomer models exist for GPCRs, based on predictions of relative stabilities of dimer interfaces by molecular simulations and bioinformatics studies as well as wet-lab techniques. Extended biased molecular dynamics simulations suggested a model in which homo-dimers characterized by stable interactions involving transmembrane helix 1 (TM1) transiently interact with the other protomer via other helices such as TM4[18]. Bioinformatics studies predicted a role for transmembrane helices TM1 and TM4 to TM6 in dimerization; mutation of residues in this region disrupted dimerization[19,20]. AFM, crystallography and FRET studies of the β_1- and β_2-adrenergic receptors[21], muscarinic receptor M_3[22], rhodopsin[17,23] and the μ-opioid receptor[24] suggested that oligomerization interfaces are most probably formed by TM1-TM2-helix(H)8 and TM4-TM5 or TM5-TM6. So far, several spatial arrangements of tetrameric GPCRs are discussed. For muscarinic receptor M_3 a rhombic arrangement of tetramers seems to be preferred rather than linear or squared ones[22], whereas for rhodopsin either a more linear or squared arrangement are discussed[10,17,22]. We will discuss our data in respect to these findings to suggest TGR5 oligomerization models.

To perform protein-protein interaction studies in living cells without disturbance and with high spatial resolution, we applied Multiparameter Image Fluorescence Spectroscopy (MFIS). It combines fluorescence lifetime imaging and fluorescence anisotropy microscopy allowing a comprehensive analysis of the biophysical properties of homo- and heteromeric molecular complexes by FRET. MFIS is based on Multiparameter Fluorescence Detection (MFD), which has been established as a standard tool to investigate biomolecules in *in vitro* experiments[25–27]. Similar to MFD, MFIS-FRET records photons one by one, which allows for parallel recording of all fluorescence parameters (fundamental anisotropy, fluorescence lifetime, fluorescence intensity, time, excitation spectrum, fluorescence spectrum, fluorescence quantum yield, and distance between fluorophores) and additionally pixel/image information over time periods of hours with picosecond accuracy. The multidimensional analysis of correlated changes of several parameters measured by FRET, fluorescence fluctuation, fluorescence lifetime and anisotropy increases the robustness of the analysis significantly. The economic use of photon information even allows detection of fluorescent fusion proteins that are expressed at very low levels. We already showed the reliability of this technique for molecular interaction studies in different environments in human and plant cells[28,29].

The main focus of this study is to use a combined strategy applying cellular biology, co-immunoprecipitation experiments, MFIS-FRET, molecular modelling and simulations to obtain information about oligomerization of TGR5 and the influence of a mutation in the TGR5 ERY domain on oligomerization.

Results

TGR5 forms homo-complexes but the complex affinity differs between TGR5 variants. To characterize the complex formation of TGR5, we used three TGR5 variants, TGR5 wt, TGR5 Y111A, and TGR5 Y111F. The tyrosine residue at position 111 is part of the highly conserved ERY motif, which is important for GPCR function[30] and also predicted to be phosphorylated by EGFR using NetPhos[31].

Immunofluorescence staining in MDCK and HEK293 cells as well as FACS analysis of transfected HEK293 cells demonstrated that all TGR5 variants were correctly localized at the plasma membrane in about 92% of the transfected cells (Fig. 1a,b). Furthermore, TGR5 responsiveness towards TLC was investigated using a cAMP-responsive luciferase assay[9], where luciferase activity served as a measure for the second messenger cAMP following TGR5 activation. Forskolin (F) elevates cAMP independent of TGR5 and was used as positive control. Stimulation of TGR5 wt, TGR5 Y111A, or TGR5 Y111F with TLC led to a significant dose-dependent increase in luciferase activity in all three cases (Fig. 1c).

To analyze the interaction between TGR5 wt proteins or TGR5 wt with TGR5 Y111A, we performed Co-immunoprecipitation experiments (Co-IP). His-tagged TGR5 wt and either TGR5 wt-YFP or TGR5 Y111A-YFP proteins were expressed in HEK293 cells. Immunoprecipitation of His-tagged TGR5 wt was carried out with an anti-His antibody. The interaction of TGR5 proteins was visualized using an anti-GFP antibody, which recognized the TGR5 C-terminal YFP (Fig. 2a lane 3). Co-IP clearly showed that TGR5 forms homo-complexes. Compared to the interaction between TGR5 wt proteins, the interaction between TGR5 wt and TGR5 Y111A is significantly reduced by about 40% as measured by densitometry (Fig. 2b).

Pixel-wise MFIS-FRET analysis demonstrates remarkable differences in FRET properties between TGR5 variants. To further analyze differences in the complex formation found by Co-IP we used the genetically encoded fluorescent proteins GFP and mCherry attached to the C-terminus of TGR5 to measure FRET by MFIS-FRET in live cells. GFP and mCherry are commonly used as a FRET pair with a Förster radius $R_0 = 52$ Å[32]. As shown in Fig. 3a and SI Fig. 1a, all TGR5-GFP and TGR5-mCherry variants (wt, Y111A and Y111F) are strongly co-localized at the cell membrane of HEK293 cells. To visualize the heterogeneity within and between cells, the MFIS-FRET images were accurately analyzed in a pixel-wise manner to compute all relevant fluorescence parameters. During this procedure, photons are pixel-wise selected, grouped according to

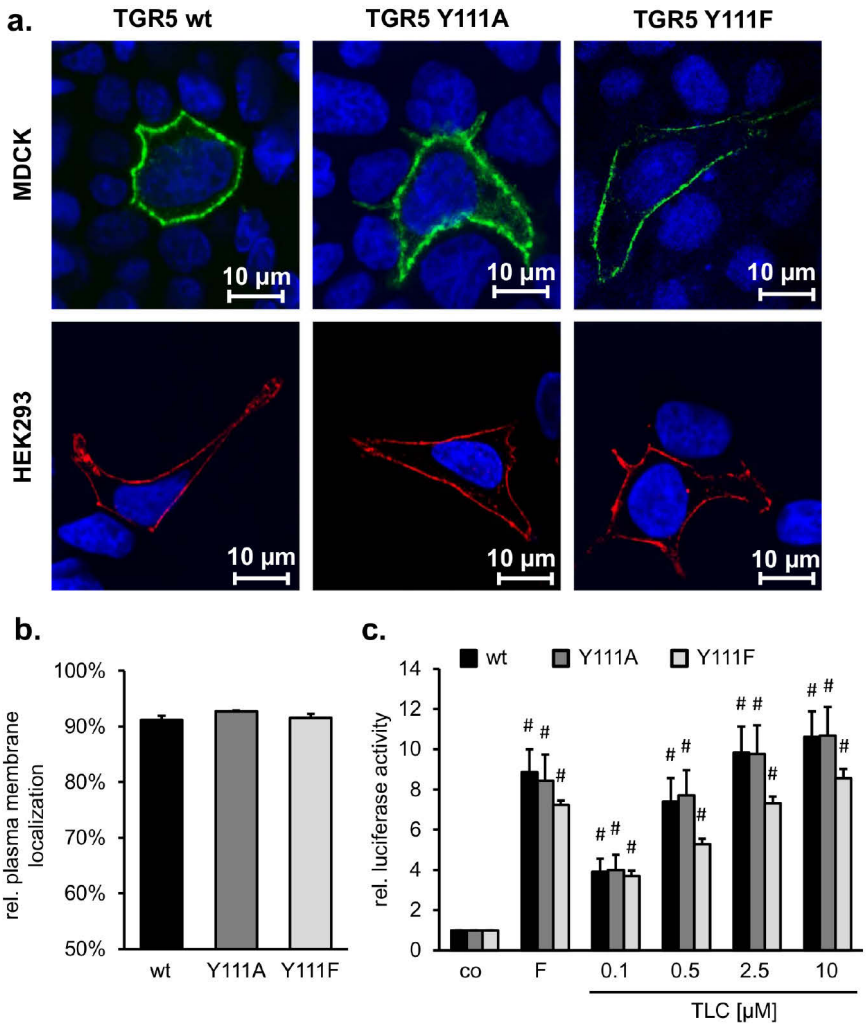


Figure 1. Localization and functional analysis of TGR5 wt and Y111 variants. (**a**) Localization of TGR5 by confocal laser scanning microscopy. MDCK cells (upper panels) were transiently transfected with FLAG-TGR5-YFP constructs. The YFP-fluorescence was detected in the plasma membrane for TGR5 wt as well as for the TGR5 Y111A and TGR5 Y111F variants. HEK293 cells (lower panels) were transiently transfected with TGR5-pcDNA constructs. TGR5 was stained using the RVLR2 antibody (in red). TGR5 as well as the TGR5 Y111A and TGR5 Y111F variants were present in the plasma membrane. Nuclei were stained with Hoechst (blue). Bars = 10 µm. (**b**) Relative quantification of TGR5 plasma membrane localization using flow cytometry. The amount of FLAG-TGR5-YFP within the plasma membrane corresponds to the amount of positive FLAG-tag labelling (=extracellular labelling) divided by the total amount of YFP-fluorescence. TGR5 Y111A and TGR5 Y111F were detected on the cell surface in 92.7% and 91.5% of the transfected cells, which was similar to the TGR5 wt with 91.2% (n = 3 independent transfection experiments). (**c**) TGR5 receptor activity was determined using a cAMP responsive luciferase assay. HEK293 cells were co-transfected with TGR5 (pcDNA3.1+), a cAMP responsive luciferase reporter construct, and a Renilla expression vector. Luciferase activity served as a measure of the rise in intracellular cAMP following activation of TGR5. Forskolin (F, 10 µM) was used as TGR5-independent positive control. TGR5 Y111A and TGR5 Y111F did not affect receptor responsiveness to the bile acid taurolithocholate (TLC). Results (wT n = 8; TGR5 Y111A n = 9; TGR5 Y111F n = 6) are expressed as mean + SEM. #Significantly different ($p \leq 0.01$) from DMSO (co = control).

their properties, and selectively integrated to reduce noise (see SI methods). For a direct proof of FRET, it is necessary to show that the observed signal changes are due to differences in FRET efficiency E and not due to local

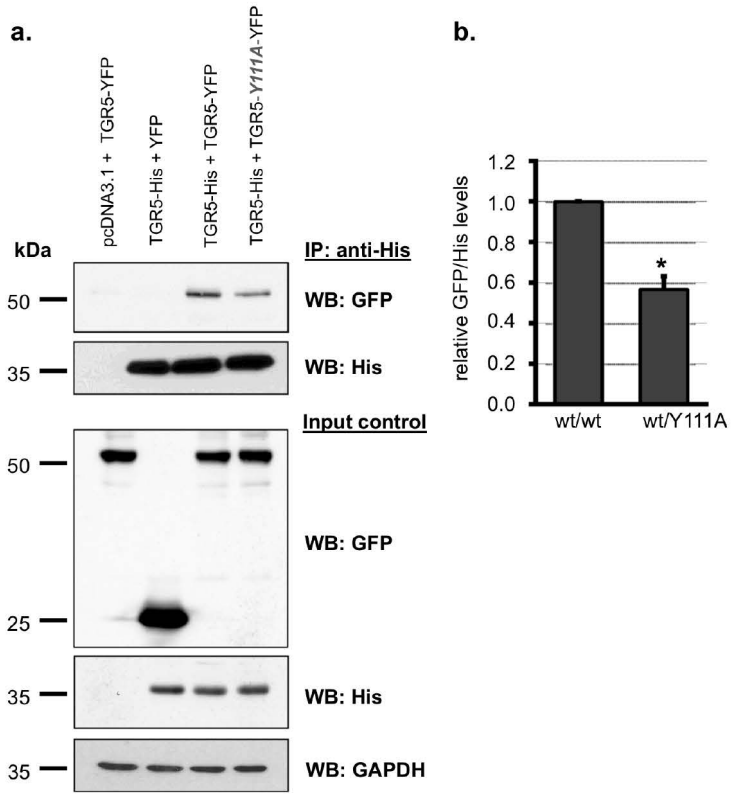


Figure 2. Detection of TGR5 multimerization by co-immunoprecipitation. (**a**) HEK293 cells were transiently transfected with pcDNA3.1 and TGR5-YFP, TGR5-His and pEYFP, TGR5-His and TGR5-YFP, or TGR5-His and TGR5 Y111A-YFP. Immunoprecipitation (IP) was carried out using an anti-His antibody. Equal volumes of the precipitate were deglycosylated with N-glycosidase-F, separated by SDS-PAGE, and blotted onto PVDF membranes. For Western blotting (WB) horseradish-peroxidase-coupled primary antibodies against His and GFP were used. TGR5-YFP was co-precipitated with TGR5-His. Mutation of tyrosine 111 to alanine in TGR5-YFP reduced the amount of co-precipitated variant receptor. Cell lysates (50 μg total protein lysates served as input controls and were separated by SDS-PAGE and proteins were blotted onto PVDF membranes. WB was carried out with horseradish-peroxidase-coupled primary antibodies against His and GFP or an antibody against glyceraldehyde-3-phosphate dehydrogenase (GAPDH). (**b**) Densitometric analysis of the anti-GFP and anti-His Western blots. Relative TGR5-TGR5 interaction was determined as relative GFP to His levels. Results are expressed as mean + SEM (n = 4), *Significantly different from wt-His/wt-YFP interaction, $p < 0.05$.

changes of fluorophore properties or transfection artifacts. Thus, it is mandatory to analyze both FRET indicators: (i) FRET-induced donor quenching due to the presence of acceptor and (ii) the occurrence of FRET-sensitized acceptor fluorescence[33].

A selection of these relevant FRET indicators is displayed in images of the TGR5 wt donor-only reference sample (Fig. 3b) and the FRET sample (Fig. 3c, first row): Signal intensity S of the donor GFP in the green detection channel by donor excitation ($S_{em,ex} = S_{G,G}$; $\lambda_{ex} = 488$ nm), signal intensity of the directly excited acceptor mCherry in the yellow detection channel $S_{Y,Y}$ ($\lambda_{ex} = 559$ nm), and as a result of FRET the FRET-sensitized mCherry signal $S_{Y,G}$. Moreover, the quenching of the donor by FRET is judged by comparing the fluorescence-weighted average lifetimes of the donor in absence $\langle \tau_{D(0)} \rangle_f$ and presence of acceptor $\langle \tau_{D(A)} \rangle_f$, respectively. If no FRET occurs, we only expect signals in the green channel. This is indeed observed for the reference measurement TGR5-GFP (Fig. 3b). Furthermore, $\langle \tau_{D(0)} \rangle_f$ does not change, as expected.

Compared to cells transfected with the donor-only reference TGR5-GFP (Fig. 3b), the MFIS-FRET measurements of the FRET sample suggest the presence of FRET, as the FRET-sensitized acceptor signal was detectable (Fig. 3c, $S_{Y,G}$ image), and $\langle \tau_{D(A)} \rangle_f$ (Fig. 3c, lifetime image) is clearly reduced compared to $\langle \tau_{D(0)} \rangle_f$. The same observations were made also for TGR5 variants Y111A and Y111F (Fig. 3c and SI Fig. 1b).

The correlated FRET-specific change of both FRET-indicators is best visualized in a 2D-histogram plotting the ratio of the corrected fluorescence intensities of donor and acceptor (F_D/F_A) (SI Table S1) versus donor

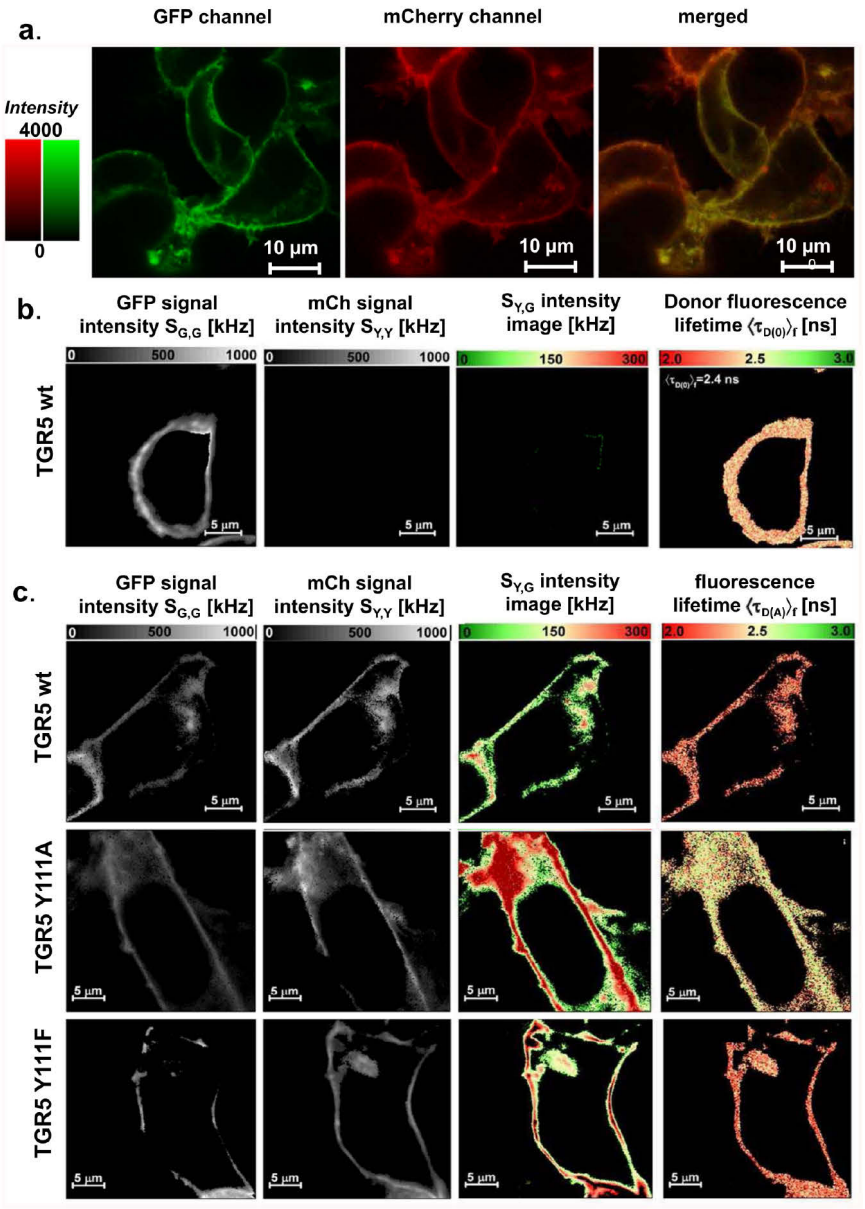


Figure 3. Detection of TGR5 multimerization by pixel-wise MFIS-FRET analysis. (**a**) HEK293 cells, transiently transfected with TGR5-GFP and TGR5-mCherry (transfection ratio 1:10), were imaged for co-localization of GFP and mCherry using sequential scanning and a scanning resolution of 1024 × 1024 pixels. Each TGR5-GFP and TGR5-mCherry picture is shown in a false color saturation mode and then overlaid by using green and yellow intensity colors. TGR5 wt-GFP and TGR5 wt-mCherry are clearly co-localized at the cell membrane. Scale bar 10µm. The TGR5 Y111 variants are shown in SI Fig. 1. (**b**) MFIS analysis of TGR5 wt-GFP transfected HEK293 cells by comparing (from left to right) the signal intensity of the donor GFP ($S_{G,G}$), signal intensity of the acceptor mCherry ($S_{Y,Y}$), the detection of yellow mCherry photons after excitation of GFP ($S_{Y,G}$, *S*: signal, *Y*: yellow emission, *G*: green excitation) as a result of FRET, and changes in the donor fluorescence lifetime $\langle\tau_{D(0)}\rangle_f$. For TGR5 wt-GFP only the donor signal but no acceptor signal is detected. The MFIS analysis of TGR5 Y111 variants is shown in SI Fig. 1. (**c**) The same parameters were used for TGR5 GFP/mCherry samples. The MFIS measurements show FRET ($S_{Y,G}$ and changes in $\langle\tau_{D(A)}\rangle_f$) in all TGR5 variants, which indicates at least homo-dimerization.

fluorescence lifetime ($\langle\tau_D\rangle_f$), where the color scale corresponds to the pixel frequency with black being highest (Fig. 4a). The correlated shift of both indicators proves the molecular proximity of TGR5 wt and TGR5 Y111A/F monomers suggesting the presence of at least homo-dimers. To study whether also higher order oligomers form, we performed acceptor titration experiments with varying donor to acceptor transfection levels resulting in an anticipated 40-fold higher acceptor concentration in the last titration step. Here, the FRET-indicators (F_D/F_A) and $\langle\tau_{D(A)}\rangle_f$ allow for a qualitative interpretation of the measurements without applying a specific model. FRET senses the local proximity of binding partners within ~80 Å. Hence, if small oligomers exist $\langle\tau_{D(A)}\rangle_f$ the fluorescence intensity ratio (F_D/F_A) will decrease with increasing acceptor concentration, whereas they do not change if only dimers exist. For TGR5 wt and TGR5 Y111F, $\langle\tau_{D(A)}\rangle_f$ reduced significantly by 17% and 14%, respectively, whereas for TGR5 Y111A $\langle\tau_{D(A)}\rangle_f$ reduced only by 7%. This behavior is also found in the fluorescence intensity ratios F_D/F_A. Here, significant transfection-level dependent FRET-changes are found for TGR5 wt and TGR5 Y111F, while only minor changes are found for TGR5 Y111A (Fig. 4a). The correlated shift of both FRET-indicators confirms that changes in FRET are indeed due to different concentrations. This suggests a significant formation of TGR5 wt and TGR5 Y111F oligomers but no or only few oligomers for TGR5 Y111A. We observed the distinct properties of TGR5 Y111A also via the fluorescence properties of the fused GFP, which was measured always as donor-only reference sample in the FRET experiments. While GFP fused to TGR5 wt and TGR5 Y111F had a fluorescence lifetime $\langle\tau_{D(0)}\rangle_f = 2.4$ ns, $\langle\tau_{D(0)}\rangle_f$ increased to 2.8 ns in the Y111A variant (SI Fig. 1b). In addition to the lifetime shift, we found a spectral red shift of 13 nm in the emission spectrum of TGR5 Y111A excited at 488 nm as compared to TGR5 wt (SI Fig. 1c).

TGR5 wt and TGR5 Y111F form higher-order oligomers, whereas TGR5 Y111A forms primarily dimers. The pixel-wise analysis of the fluorescence data by the fluorescence-averaged lifetime $\langle\tau_D\rangle_f$ and the fluorescence intensity ratios (F_D/F_A) does not allow us to resolve multiple species because the information contained in the recorded fluorescence decays is reduced to two numbers. Hence, sample heterogeneities that naturally arise in imaging cannot be resolved. To overcome this limitation, the fluorescence decays are analyzed directly by pixel-integrated analysis with high precision. Here, two fluorescence decay curves $f(t)$ are compared: the decay of a FRET sample $f_{D(A)}(t)$ and that of the donor-only reference $f_{D(0)}(t)$ (Fig. 4b). This comparison is conveniently done by computing the time-resolved FRET-induced donor decay $\varepsilon(t)$, which is defined by the ratio of the two decays $f_{D(A)}(t)/f_{D(0)}(t)$ as described in eq. (1). The supporting Figure 2 shows how $\varepsilon(t)$ plots can be interpreted. The FRET-induced donor decay $\varepsilon(t)$ allows visually identifying the population of all donor species. For instance, species with no-FRET give rise to a constant offset, while FRET-species cause decay. The slope of this decay in a semi-logarithmic plot as shown in Fig. 4b provides a measure of the rate constant of FRET, which increases with decreasing donor acceptor distance. A non-exponential decay indicates a mixture of distinct FRET species in which the donor and the acceptor are separated by different distances. The donor is quenched by all acceptors in its vicinity.

In Fig. 4b, the experimental fluorescence decays of all variants are displayed as $\varepsilon(t)$ curves. Differences in the constant offset and the slope of the decays are clearly visible. For a better comparison of the slopes only the fraction of the FRET species was determined in a fit (equation (1), results see SI Table S2 and SI Figure 2) and displayed in Fig. 4b ($\varepsilon_{FRET}(t)$ curves). At a low donor to acceptor transfection level (DA 1:40), the decay has two distinct regions: a steep slope and a shallow slope region. The steep slope corresponds to a high rate constant of FRET, while the shallow slope corresponds to a low rate constant of FRET. For TGR5 wt and TGR5 Y111F, the slope depends on the transfection ratio, while no such dependency is observed for TGR5 Y111A.

To quantify these changes we formally describe the fluorescence decays by two FRET-rate constants, which are for convenience given in units of apparent distances $R_{DA,app}$ (equation (5) and SI Table S2, SI Fig. 3). For all TGR5 variants, this k_{FRET} fit resulted in a short apparent distance $R_{DA,app\text{-}1}$ with a small fraction and a long apparent distance $R_{DA,app\text{-}2}$ with a large fraction. As shown in Fig. 4c, in TGR5 wt and TGR5 Y111F both apparent distances $R_{DA,app\text{-}1}$ and $R_{DA,app\text{-}2}$ became shorter ($R_{DA,app\text{-}1} = 40$–20 Å; $R_{DA,app\text{-}2} = 75$–50 Å) with increasing acceptor concentration. Furthermore, the species fractions also changed: the short distance-fraction increased from 7% to 30% in an acceptor-dependent manner, leading at the same time to a strong reduction of the long distance-fraction from 39% to 12%. We quantified this change by computing the mean energy transfer efficiency E_{mean} (equation (7)) of the FRET active species, which markedly increased for TGR5 wt and TGR5 Y111F in contrast to TGR5 Y111A. Considering TGR5 wt and TGR5 Y111F, the FRET efficiency changes significantly with the acceptor concentration (Fig. 4c), while this is not the case for TGR5 Y111A. Hence, the concentration of oligomers is very low for TGR5 Y111A, so that these data are best suited to study the structural features of the dimer.

Of note, to rule out any overexpression artifacts, we additionally considered proximity FRET using the titration experiments. Due to the single-molecule sensitivity of our confocal microscope, we could perform FRET experiments with acceptor concentrations of ~1 μM, which corresponds to a molecule density of < ~0.002 acceptor molecules/nm^2 (see SI Notes). At these concentrations proximity FRET is negligible ($E < 0.1$)[34].

The TGR5 ligand TC has no influence on the oligomerization state of TGR5. It has been shown that activation by ligands can influence GPCR oligomerization[10]. To determine the ligand effect on TGR5, we tested whether taurocholate (TC) stimulation, a bile acid less cytotoxic than TLC in live cells, affects oligomerization of TGR5 wt, TGR5 Y111A, and TGR5 Y111F. A time series analysis was designed, where MFIS-FRET was measured in three cells before, directly after as well as 10 and 20 min after addition of 10 μM water soluble TC. We monitored FRET via the species-averaged donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_x$. As shown in Fig. 5, $\langle\tau_{D(A)}\rangle_x$ was neither changed in donor samples (TGR5 GFP) nor in FRET samples (TGR5 GFP/mCherry). A more detailed FRET analysis of the time series experiments showed that neither the distances nor the species fractions changed markedly due to addition of TC (SI Fig. 4). These results indicate that TC does not influence the oligomerization state of TGR5 variants.

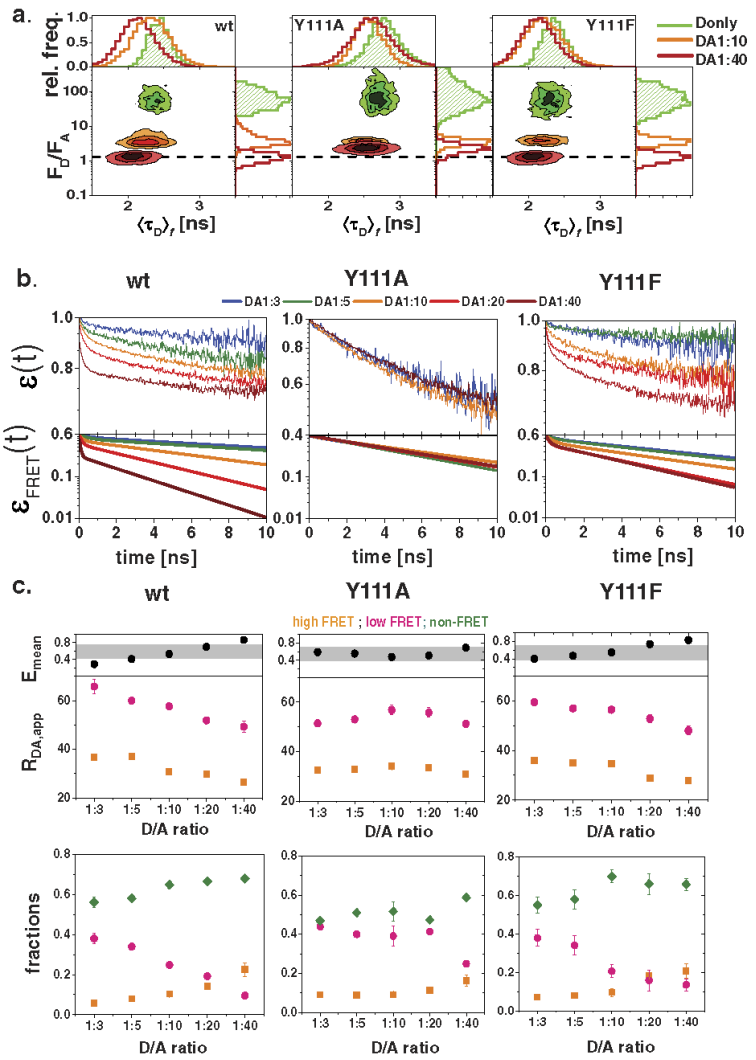


Figure 4. Pixel-integrated analyses of TGR5 FRET properties. (**a**) The MFIS-FRET 2D plots are generated with Origin 8.6 and show an overlay of two histograms of the (background, crosstalk and spectral shift) corrected fluorescence intensity ratio (F_D/F_A) vs. $\langle\tau_{D(A)}\rangle_f$. TGR5 wt and TGR5 Y111F donors (in green) showed a $\langle\tau_{D(0)}\rangle_f = 2.4$ ns and a high green to yellow signal. With increasing amounts of the acceptor mCherry (orange and red islands) both parameters were strongly reduced in TGR5 wt and TGR Y111F, but not in TGR5 Y111A. All samples were corrected for relative brightness, relative direct mCherry excitation in the green detection channel, spectral shift of the Y111A variant, and background in the green and yellow channels (see methods 5.10 eqs 2 and 3). (**b**) FRET-induced donor quenching $\varepsilon(t)$ derived from sub-ensemble fluorescence measurements on TGR5 variants at different donor-to-acceptor ratios. The time-axis measures the time between excitation and detection of donor photons. The upper row shows the experimental data. In the bottom row the offset (Non-FRET fraction) is subtracted and the result is termed $\varepsilon_{FRET}(t)$. In TGR5 wt and TGR5 Y111F, FRET clearly increased in a mCherry-dependent manner, whereas in TGR5 Y111A all $\varepsilon_{FRET}(t)$ curves behaved similar. (**c**) FRET-decays from sub-ensemble analysis at different donor-to-acceptor (D/A) ratios were fitted with a two-k_{FRET} fit to obtain two apparent distances $R_{DA,1}$ and $R_{DA,2}$ (upper row) with their corresponding FRET fractions (lower row) and to calculate the mean efficiency E_{mean}. E_{mean} increased in an acceptor-dependent manner in TGR5 wt and TGR5 Y111F, whereas E_{mean} changed only slightly in TGR5 Y111A. These changes in E_{mean} correlate with a reduction of both apparent distances $R_{DA,1}$ and $R_{DA,2}$ in TGR5 wt and TGR5 Y111F: In the lower row, the $R_{DA,1}$ fractions increase, whereas the $R_{DA,2}$ fractions decrease in an acceptor-dependent manner. Orange: $R_{DA,1}$ and $R_{DA,1}$ fraction, pink: $R_{DA,2}$ and $R_{DA,2}$ fraction, green: non-FRET fraction, the grey bar in E_{mean} represents average E_{mean} for TGR5 Y111A.

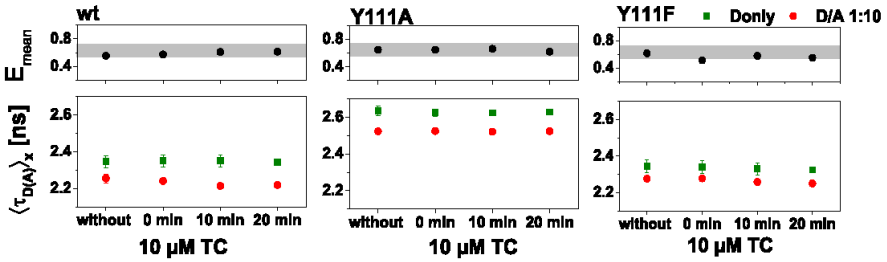


Figure 5. Influence on FRET after treatment with TGR5 ligand TC. HEK293 cells were transiently transfected with TGR5-GFP alone (Donly, green) or with TGR5-GFP and TGR5-mCherry at a ratio D/A 1:10 (DA, red). For time-series analysis three cells were selected using the Olympus time laps function, and MFIS-FRET measurements were taken before addition of 10 µM TC (without), immediately after addition of TC ($t=0$), and after 10 min and 20 min, respectively. The species-averaged donor fluorescence lifetime $\langle\tau_D\rangle_x$ was determined and plotted against time, as well as the mean efficiency E_{mean}, which was calculated from data shown in SI Fig. 4. Each point represents the average of nine cells. No lifetime changes were observed for Donly samples and DA samples in the presence of the agonist TC.

Structural arrangement of homo-di- and oligomeric TGR5. Next, we analysed which structural features of the TGR5 complexes can be extracted from the observed FRET parameters. Previous studies by Sindbert *et al.*[35] and Kalinin *et al.*[36] have shown that the extent of FRET between two flexibly linked fluorescent probes can be accurately predicted by calculating the distance distribution between all fluorophore positions that are sterically accessible (accessible volume, AV) for a given structural model. As both fused fluorescent proteins have flexible connecting amino acid residues (SI Table 3) creating a large, widely distributed structural ensemble[37], computer simulations generating probe distributions can be readily applied to study TGR5 assemblies by FRET.

Simulation of the expected FRET properties. The structural model of the TGR5 monomer required for FRET modelling was generated by performing multi-template homology modelling based on seven template structures of related class A GPCRs (see SI methods "structural models of TGR5 dimers and tetramers" and ref. 38). As shown in Fig. 6a, we generated three possible homo-dimerization models with interfaces involving TM1-TM2-H8 (for convenience abbreviated as 1/8 dimer), TM4-TM5 (4/5 dimer), or TM5-TM6 (5/6 dimer). To assure accuracy, we compared two procedures for calculating the distance distributions between fluorophore positions for the TGR5 models: (i) Explicit linker simulations based on explicit peptide linker/GFP-MD-simulations followed by calculations of conformational free energies to weight each linker-GFP configuration in the presence of a TGR5 dimer and an implicit membrane bilayer (SI Fig. 5, see also SI methods). This thermodynamic ensemble (TE)-approach is expected to be more accurate than the following procedure but the computations are time consuming. (ii) Implicit linker simulations by AV-calculations weighted by a Gaussian chain distribution, so that entropic effects and geometric factors in terms of steric exclusion effects by the TGR5 oligomer and the membrane are taken into account (SI methods). The AV approach has to be calibrated to be accurate but it has the advantage that the computation is very fast.

The TE-approach results in a hemispherical arrangement of GFP on the cytoplasmic side, which is centred on the attachment point at helix 8 of TGR5 (SI Fig. 5) and each linker/GFP configuration is Boltzmann weighted according to the conformational free energy (SI Fig. 5). Configurations of lower probability are found when GFP approaches TGR5 due to energetically unfavourable contacts. The Boltzmann-weighted distribution of distances between the linker N-terminus and the GFP fluorophore shows a peak distance of about 45 Å, while the minimal distance is about 35 Å. This is due to the fact that the fluorophore is located 20 Å away from the linker C-terminus inside the β-barrel structure of GFP and thus is inaccessible to the linker's N-terminus. The peak linker length (without considering GFP) is about 25 Å. This is about 5 Å longer than the average radius of gyration of a Gaussian chain polypeptide of the same number of residues (33 amino acids yielding 3.5 Å * $33^{0.5} = 20$ Å[39]). The deviation shows that the linker with GFP does not exactly behave like a 'perfect' Gaussian chain. The Boltzmann-weighted fluorophore position map (Fig. 6c, SI Fig. 5) was used for inter-dye distance distribution calculations.

The implicit model (Fig. 6c) was tested as an alternative to account for dye-linker diffusion. The accessible volume (AV) approach was used to estimate all possible dye positions within the linker length from the attachment point without steric clashes with the macromolecular surfaces. The fluorophores are approximated by a sphere with a defined radius, which is estimated from the physical dimensions of the molecules (left panel). The connecting linker is modelled as a flexible cylinder. To take entropic effects into account, the linker was assumed to obey Gaussian chain behaviour. Thus, the fluorophore distribution density gradually drops as the distance from the attachment point increases. For the implicit model, the 55 amino acid residues (SI methods and SI Table S3) between the structured parts of the TGR5 C-terminus and GFP were considered as a flexible sequence with unknown structure with a length of ~203.5 Å at maximal extension.

Both methods for linker simulations gave very similar results. The (1/8) dimerization model shows a distance distribution between fluorophore positions between 25–150 Å with the highest probability at 55 Å and 60 Å for the explicit and implicit linker models, respectively. The distances between fluorophores in models (4/5) and (5/6)

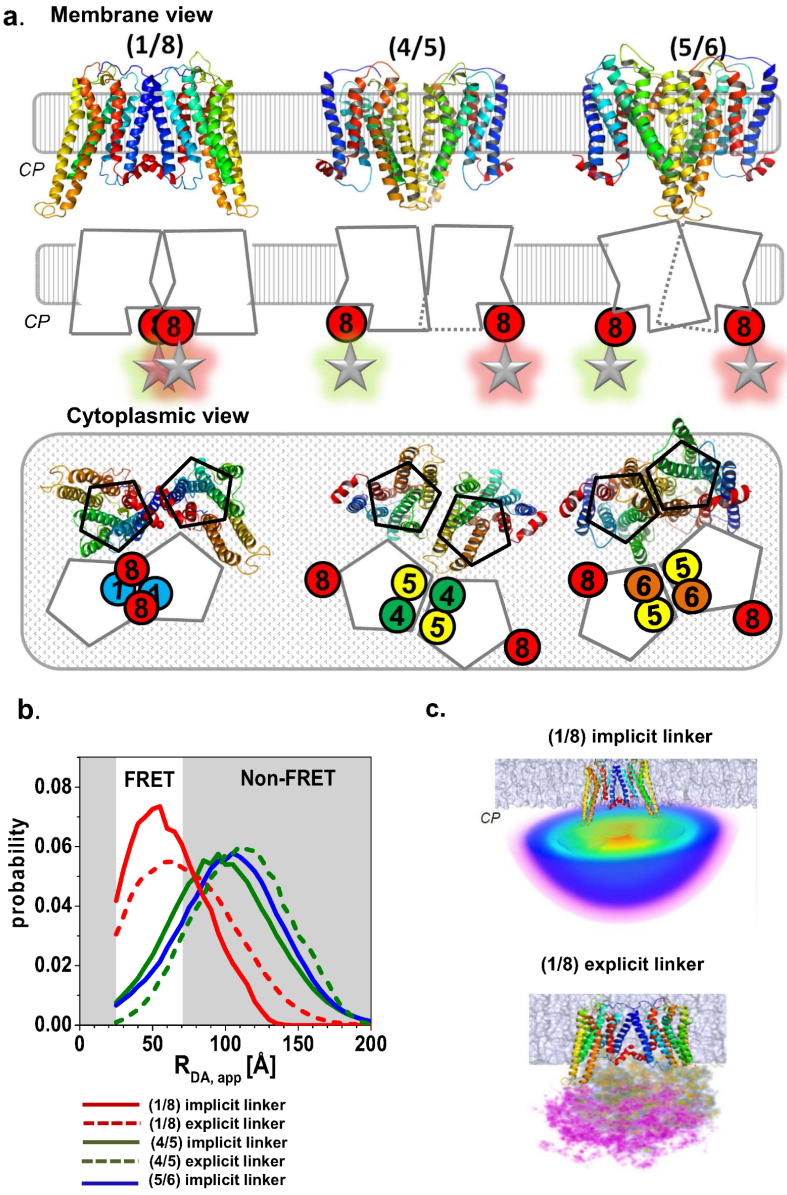


Figure 6. Homo-dimerization models and their distance distributions. (**a**) Homo-dimerization models with the following interfaces from left to right: (1/8), (4/5) and (5/6). TGR5 monomer helices are rainbow-coloured starting with TM1 in blue to H8 in red. Top row: membrane view of the interface models in cartoon and schematic representation (circles representing TMs). Bottom row: cytoplasmic view of the interface models. The fluorescent proteins, which are attached to the cytoplasmic H8, are presented as stars glowing in green for donor (eGFP) and red for acceptor (mCherry). Abbreviation: CP = cytoplasm. (**b**) Distance probability distributions calculated with an explicit (dotted line) and implicit linker (solid line) for the homo-dimerization models (1/8) (red), (4/5) (green), and (5/6) (blue). The non-FRET area is shaded in grey. (**c**) Positional distributions of the fluorescent probes for the TGR5 (1/8) interface. The implicit linker simulations yield weighted AVs for both fluorophores which overlap and create one huge sphere (top panel). The probability of the allowed fluorophore positions decrease from red, yellow over green, blue to pink. The explicit linker simulations yield a thermodynamic ensemble (bottom panel) depicted as an orange-blue and purple volume map, respectively. The ensembles also overlap to a high degree. Higher saturation represents higher fluorophore position occupancy. Both methods gave very similar results.

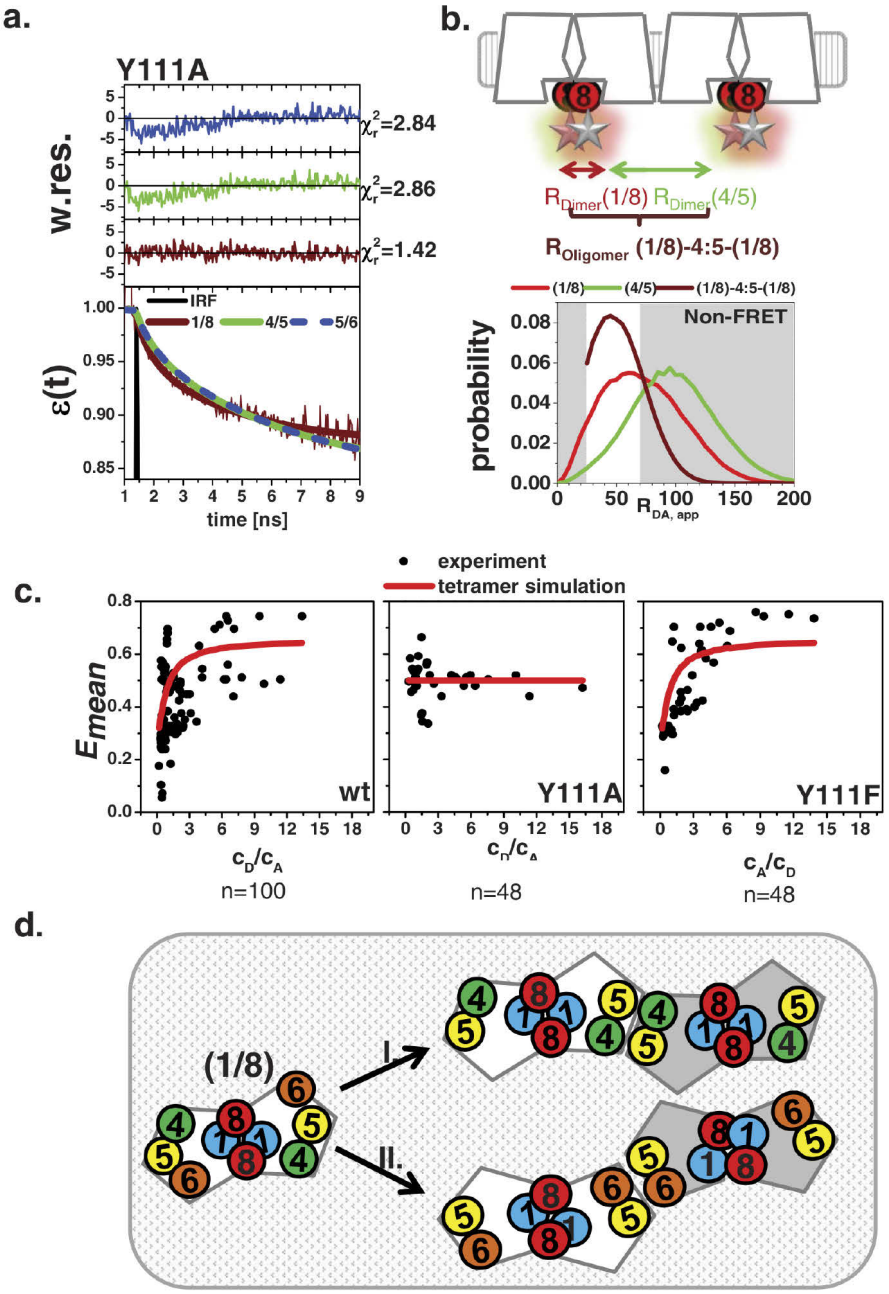


Figure 7. TGR5 oligomerization models. (**a**) Fit of the FRET-induced donor quenching curve $\varepsilon(t)$ on TGR5 Y111A with two species normalized to unity: (i) Dimer (fraction x_{Dimer}) with the complete distance distribution (FRET and Non-FRET) of the corresponding dimer models (Fig. 6b) and (ii) donor only/ FRET inactive molecules. Only the distance distribution of the 1/8 dimer model gives a satisfactory fit as judged by the weighted residuals and the reduced chi squared χ_r^2. Fit results of TGR5 Y111A for x_{Dimer}: 1/8 dimer: 0.27; 4/5 dimer: 0.59; 5/6 dimer: 0.73. (**b**) The schematic presentation shows the two individual apparent distances from the interfaces (1/8) and (4/5). Both R_{Dimer} can be converted into FRET rates. In an oligomer the two FRET rates add up and have to be convolved to calculate the new apparent distance $R_{(oligomer)}$. The resulting distance distribution is similar to the dimer (1/8). (**c**) Dependence of the TGR5 oligomerization monitored by the FRET efficiency (experiment (black) and modeled (red)) on the donor acceptor ratio c_A/c_D. In the cells the donor,

acceptor and total TGR5 concentration (including inactive mCherry (30%))) varied between 0.25–6.3 μM, 0.1–5.0 μM and 0.5–13 μM, respectively. The dimer is composed of a donor acceptor distance of 45 Å, and the tetramer is composed out of two dimers separated by 100 Å. The modeled dissociation constant of the dimer K_{D1} was fixed to 10 nM for all TGR5 variants. The values for the modeled dissociation constants of the oligomer (Tetramer) were: K_{D2}(TGR5 wt) = 70 nM, K_{D2}(TGR5 Y111F) = 200 nM, K_{D2}(TGR5 Y111A) = 2000 nM). (**d**) Two possible oligomers are reasonable I. ((1/8)-4:5-(1/8) and II. (1/8)-5:6-(1/8): TGR5 monomers form a dimer with the contact sites in TM1 (blue circle) and H8 (red circle). H8 is attached to fluorescent fusion proteins (GFP and mCherry). In a tetramer contact sites in TM4 (green circle) and TM5 (yellow) (**I**) or TM5 (yellow) and TM6 (orange) (**II**) create a second interface promoting a linear oligomer organization.

are similarly distributed with the highest probability at around 95–110 Å; i.e. the distance of most conformers is too large for significant FRET. Implicit and explicit linker models thus show very similar inter-dye distance distributions for all dimer models: The implicit model shows a 5 Å shift towards the higher length for the (1/8) dimer and a 15 Å shift towards the shorter length for the (4/5) dimer model.

Finally we can conclude that both linker simulation techniques predicted FRET and should distinguish a 1/8 dimer from 4/5 dimer and 5/6 dimer, respectively, because the FRET probe distance distributions have a characteristic peak at short distances (Fig. 6b). However, the FRET probe distance distributions of the two dimers involving TM5 are expected to be not distinguishable in our FRET experiments (Fig. 6b).

In the first step of oligomerization contact sites in TM1 and helix 8 are involved. The shape of the distance distribution (determined by our linker simulation) and the concentration-dependent change in E_{mean} (using MFIS-FRET titration experiments) should allow us to distinguish (i) oligomerization interfaces and (ii) oligomerization pattern.

The concentration-independent FRET efficiency (Fig. 4) of the TGR5 Y111A variant suggests the preferential presence of homo-dimers. Therefore, it is a perfect variant to test which of our distance probe distributions describes the FRET-induced donor quenching curve $\varepsilon(t)$ best. Figure 7a shows the fits using a model with the complete distance distribution (FRET and Non-FRET) of the corresponding dimer models (Fig. 6b, SI Table 4). Only the distance distribution of the 1/8 dimer model gives a statistically satisfactory fit as judged by the weighted residuals (*w. res.*) and the smallest χ_r^2. Hence, TM1 and helix 8 most likely form the primary oligomerization interface.

From the same titration experiments, we conclude that TGR5 wt and TGR5 Y111F are able to form higher-order oligomers because of the concentration-dependent increase in FRET efficiency (Fig. 4c). This finding implies that at least a second interface should exist for TGR5 homo-oligomer formation. As shown in Fig. 6b, the average apparent distances between fluorescent proteins attached to TGR5 helix 8 (without a coupled G-protein) were 120 Å for the (4/5) dimer model and 103 Å for (5/6) model, respectively, and the effective apparent oligomer distance for both patterns is approximately 49 Å (brown curve in Fig. 7b) due to the presence of multiple acceptors. We applied a dimer/tetramer simulation to our MFIS data to estimate the two corresponding association constants (Fig. 7c, SI Fig. 6) by analysing the dependence of the mean FRET efficiency E_{mean} on the ratio of donor to acceptor concentration (c_D/c_A). Moreover, the spread in the FRET efficiencies observed in Fig. 7c is also caused by the distinct protein concentrations in the cell and is taken into account in the simulations (SI Fig. 6a–c). For TGR5 wt and Y111F (K_D in 100 nM range), the simulations indicate that almost all dimers form tetramers, whereas TGR5 Y111A forms predominantly dimers (K_D in μM range).

Discussion

We pursued a combined strategy applying cellular biology, MFIS-FRET, molecular modelling and simulations with a focus on dimerization and higher-order oligomerization of TGR5. We studied the influence of a mutation in the TGR5 ERY motif (TGR5 Y111A and Y111F) located in the transmembrane helix 3 (TM3) on oligomerization.

For our oligomerization studies we replaced the tyrosine residue in the highly conserved "D/ERY" motif in TM 3 and belongs to one of two clusters important for structural stability in GPCRs[40]. Mutation studies in Rhodopsin showed that the tyrosine (Y) mutation alone did not or only marginally affect receptor function[41] regarding receptor expression, G-protein binding and ligand affinity in contrast to the residues D/ER. Consistent with literature results[41], the TGR5 Y111 variants, Y111A and Y111F, were normally localized at the plasma membrane and activated by both bile acid agonists TLC and TC to a level comparable to TGR5 wt. These findings implicated no obvious impaired ligand binding affinities or G-protein coupling. However, we observed significant differences in oligomer formation between Y111A and Y111F as assessed by Co-IP experiments and FRET measurements in live cells.

As the overall protein concentrations are very low (1–7 μM), we can rule out any overexpression artifacts due to proximity FRET (see SI notes). Therefore our MFIS-FRET titration data are best described with models assuming formation of the 1/8 dimer as the first step in oligomerization (Fig. 7a). In the second step, we suggest that TM5 (Fig. 7d) is involved as known from other oligomerization models of class A GPCRs[15,19,20]. According structural models were as templates for predicting the distance distributions in Fig. 7b. As shown in Fig. 7d and SI Fig. 7, oligomer array configurations[15,19,20] either could have a row or a rhomboid tetramer organization. One might assume that TGR5 oligomers most likely resemble in a one-dimensional row-like array mediated by a single oligomerization interface, because a single mutation in the ERY motif, Y111A in TM3, affects the association significantly (factor 10).

As shown in Fig. 8, the Y111 residue can interact with TM4-TM5 or TM5-TM6 dependent on the oligomerization. In both cases, the potential interaction sites involving TM5 can be affected during oligomerization. This

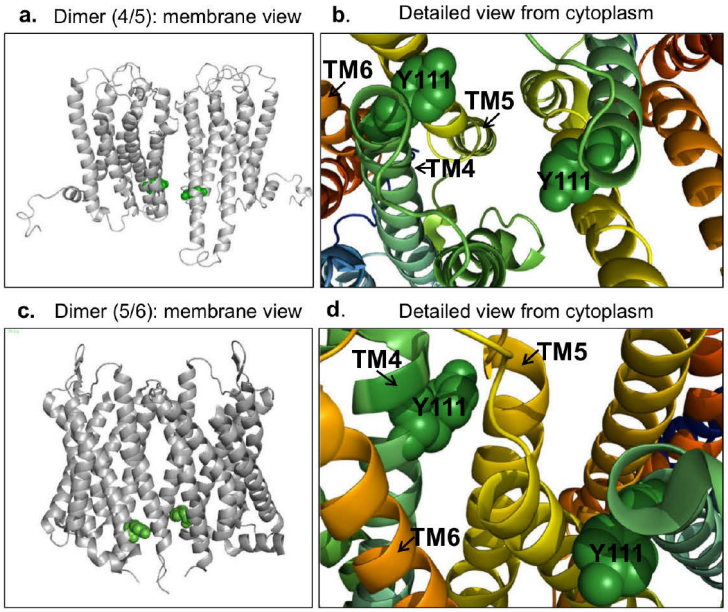


Figure 8. Influence of the Y111 residue on oligomerization. (a,c) The dimerization models of the (4/5) and (5/6) interface are displayed as a grey colored cartoon viewed from the membrane. Residue Y111 located in TM3 is depicted as a green sphere in each TGR5 monomer. (**b,d**) Blow-up of the region around residue Y111 to show possible interactions between Y111 from one TGR5 molecule with residues in TM4 (green), TM5 (yellow) and TM6 (orange) in a second TGR5 molecule.

observation is supported by two crystal structures: In the (4/5) model, as shown in CXCR4 (PDB ID: 3ODU), a charge-assisted interaction between Y111 and R146 (TM4) is possible; likewise an interaction is possible in the (5/6) model between Y111 and R280 (helix 8), as shown in the μ-opioid receptor (PDB ID: 4DK2).

It was reported that GPCR oligomerization could be affected by ligand binding[10], therefore we addressed this question in a time-series FRET analysis by ligand stimulation with TC. From simulation experiments, we expect that after G-protein binding the average apparent distances between TGR5-GFP and TGR5-mCherry get longer. Effective oligomers distributions with and without G-protein are indistinguishable, because a distance distribution difference of less than 8 Å is smaller than the anticipated accuracy of the models (see SI Table 5). In fact, the MFIS-FRET measurements showed no change in FRET properties after TC treatment, an observation that is also supported by literature[22]. As an indicator of G-protein binding, we successfully proved cAMP increase after ligand treatment in all TGR5 variants, which has also been shown recently[9,42,43]. We have no evidence that TGR5 oligomerization is affected by ligand treatment and subsequent G-protein coupling.

It is not too surprising that G protein activation does not change when reducing TGR5 higher oligomer formation, because rhodopsin and β-AR receptors in a monomeric, dimeric and oligomeric state, respectively, are capable to activate the respective G-protein[11,14,44,45]. Moreover, as described by Scarselli *et al.*[46], PALM experiments using a class A GPCR suggested that oligomerization remains unchanged by the addition of the agonist. This is in line with our findings for our class A receptor TGR5 and the bile acid ligands. While the function of higher-order oligomers for most GPCRs is still unknown, identification of dimer/oligomer interfaces will allow for targeted disruption of dimer/oligomer formation and thus elucidation of the biological relevance of these complexes. This has just been demonstrated for rhodopsin where disruption of dimerization with small peptides decreased receptor stability[44]. We recently showed that the loss of α-helicality in the TGR5 C-terminus, which constitutes the major interaction surface in the 1/8 interface, severely impairs TGR5 membrane localization and activity[9]. One can thus speculate that this influence on membrane localization and activity results from a distorted TGR5 dimerization in the ER. Additionally, the design of bivalent ligands targeting a homodimer can reduce off-target effects caused by the transactivation or inhibition of GPCRs in heterodimers[47]. Knowledge of the primary dimerization interface of a GPCR can guide the development of such bivalent ligands. The discovery that TGR5 forms higher order oligomers and that Y111 is important for this process thus is the first step for deciphering and modulating the functional relevance of TGR5 oligomerization.

To conclude, TGR5 wt forms homo-oligomers. Dimerization involves interaction contact sites in TM1 and helix 8, while its oligomerisation additionally involves TM5. Both modelled patterns, (1/8)-5:6-(1/8) and (1/8)-4:5-(1/8), are currently possible with Y111 forming charge-assisted and/or polar interactions with residues within the mentioned interfaces.

Methods

Multiparameter Fluorescence imaging spectroscopy (MFIS). All measurements in live cells were performed on an inverted confocal laser scanning microscope (FV1000 Olympus, Hamburg, Germany) additionally equipped with a single photon counting device with picosecond time-resolution (Hydra Harp 400, PicoQuant, Berlin, Germany) with home built extensions for MFD as described in[28]. Using a 60x water immersion objective (Olympus UPlanSApo NA 1.2) the sample was excited with selected wavelengths (GFP at 488 nm with 400 nW, mCherry at 559 nm with 650 nW) of a NCH white light laser with a pulse-repetition rate of 40 MHz. The emitted light was collected and separated into its parallel and perpendicular polarization and into its green and red component (beam splitter 595DCLX, AHF, Germany). GFP fluorescence was then detected by single photon avalanche detectors (PDM50-CTC, Micro Photon Devices, Bolzano, Italy) in a narrow range of its emission spectrum (bandpass filter: BS 520/35, AHF, Tübingen, Germany). mCherry fluorescence was detected by cooled hybrid detectors (HPMC-100–40, Becker&Hickl, Berlin, Germany, with custom designed cooling), of which the detection wavelength range was set by the bandpass filters HC 607/70 (AHF). MFIS images were generated via raster-scanning the sample in a continuously moving beam manner. Images were taken with 20 µs pixel dwell time and a resolution of 103 nm per pixel. With 488 nm excitation, series of 40 frames were merged into one image; with 559 nm excitation, series of 20 frames were merged together. Images were further analyzed using custom-designed software available from our homepage (http://www.mpc.hhu.de/software.html). Description of sample preparation and microscope calibration can be found in the SI methods 1 and 2.

Pixel-integrated, time-resolved ε(t) illustration. To identify appropriate pixel in the cells for further pixel-integrated analysis, we computed all fluorescence parameters for each pixel and selected the pixels in 2D-histograms of several FRET indicators (see SI methods 2 pixel-wise analysis). A pixel population with homogeneous properties was selected and then integrated for subsequent pixel-integrated sub-ensemble analysis. The time-dependent FRET parameter $\varepsilon(t)$ contains information on the underlying FRET-rate distribution and is proportional to the probability that FRET occurs at a certain time. After pixel selection, $\varepsilon(t)$ was plotted for direct visualization of molecular species with different FRET efficiencies in sub-ensemble data. $\varepsilon(t)$ is calculated as the ratio of normalized fluorescence decays of the FRET sample $f_{D(A)}(t)$ and donor-only sample, $f_{D(0)}(t)$ (see eqs 3–4).

$$\varepsilon(t) = \frac{f_{D(A)}(t)}{f_{D(0)}(t)} \tag{1}$$

$$\text{With } f_{D(A)}(t) = \varepsilon(t) \cdot f_{D(0)}(t) \tag{2}$$

$\varepsilon(t)$ is the probability density function of the occurring FRET governed by FRET rate constant(s), k_{FRET}. The decaying part of $\varepsilon(t)$ represents the features of FRET: high- or low-FRET can be directly read out from the decay slope. The amplitude of the decaying part indicates the FRET-active species fraction, x_{FRET}. Accordingly, the offset of $\varepsilon(t)$ is the FRET-inactive fraction, $(1 - x_{FRET})$.

Pixel-integrated MFIS-FRET analysis using k_{FRET} models. To determine FRET parameters from pixel-integrated, sub-ensemble data the reference samples were fitted by a multi-exponential relaxation model accounting for a multi-exponential fluorescence decay of the donor in the absence of FRET:

$$f_{D(0)}(t) = \sum_m x_D^{(m)} \cdot \exp(-t \cdot k_D^{(m)}) \tag{3}$$

in which m = 3 considers that FPs in living cells usually show at least a bi-exponential characteristic[32]. Fit parameters in donor decay include three normalized pre-exponential factors $x_D^{(m)}$ ($\sum x_D^{(m)} = 1$) and three decay rate constants $k_D^{(m)}$, which are the reciprocals of fluorescence lifetimes. The quenched donor decay $f_{D(A)}(t)$ is given by:

$$f_{D(A)}(t) = \sum_m x_D^{(m)} \cdot \exp(-t \cdot (k_D^{(m)} + k_{FRET})) \tag{4}$$

and k_{FRET} is the FRET rate constant. The fitted parameters in the $1 - k_{FRET}$ model are x_{FRET} and k_{FRET}.

From the $\varepsilon(t)$ diagrams it's clear that our data have to be fitted with $m = 2$, then we say it's a two-state model, from which we obtain two FRET rate constants and therefore two apparent distances. The quenched donor decay $f_{D(A)}(t)$ in eq. 4 is now extended:

$$f_{D(A)}(t) = \sum_m x_D^{(m)} \cdot (x_{FRET}^{(1)} \cdot \exp(-t \cdot (k_D^{(m)} + k_{FRET}^{(1)})) + x_{FRET}^{(2)} \cdot \exp(-t \cdot (k_D^{(m)} + k_{FRET}^{(2)}))) \tag{5}$$

$k_{FRET}^{(1)}, k_{FRET}^{(2)}$ are the FRET rate constants and FRET species fractions, $x_{FRET}^{(1)}, x_{FRET}^{(2)}$. In the FRET-samples molecules not performing FRET are considered as No-FRET fraction. Each FRET rate constant is converted to an apparent distance $R_{DA,app}^{(1)}$

$$R_{DA,app}^{(l)} = R_0 \cdot (k_{FRET}^{(l)} \cdot \tau_0)^{-\frac{1}{6}} \tag{6}$$

in which the unquenched GFP fluorescence lifetime is $\tau_0 = 2.4$ ns and the Förster radius between GFP and mCherry is $R_0 = 52$ Å (including static $\kappa^2 = 0.476$).

Mean energy transfer efficiency. The mean (steady-state) transfer efficiency E_{mean} is obtained using the FRET fractions and the apparent distances ($R_{DA,app}$) obtained from eq. 6.

$$E_{mean} = \frac{x_1}{1 + (R_{DA,app-1}/R_0)^6} + \frac{x_2}{1 + (R_{DA,app-2}/R_0)^6} \tag{7}$$

Effective energies of linker/GFP conformations in the presence of TGR5 dimers and an implicit membrane. Molecular dynamics simulations of GFP bound to a linker have been performed as detailed in the SI methods. Snapshots of the MD simulations of the linker/GFP construct extracted in intervals of 50 ps were stripped of water molecules and ions, and the principle axis with the lowest moment of inertia of the first residue of the linker was aligned along the z-axis. The snapshots were then rotated in steps of 90° around the z-axis to increase the sampling density and subsequently placed in proximity to residue 295 of either TGR5 monomer for any of the TGR5 dimers (1/8 interface; 4/5 interface) (SI Fig. 4). For each snapshot, the effective conformational energy $E_{\text{effective, conf}}$ (i.e., the sum of gas phase energy and solvation free energy) was computed using the FEW$^{\text{mem}}$ program[48,49], with the TGR5 dimers embedded in an implicit membrane of 34 Å width and using dielectric constants of 34, 4, and 1 for the outer to inner membrane slabs with a width of 5, 6, and 6 Å, respectively (SI Fig. 4)[50,51]; for water and protein, dielectric constants of 80 and 1 were used, respectively. The counter ion concentration for the APBS calculation[52] was set to 0.15 mM. For all other parameters, default values as set in FEW$^{\text{mem}}$ were used. All snapshots in which GFP penetrated the membrane, or in which GFP or the linker clashed with the TGR5 dimer, were omitted, leaving ~10.000 snapshots for the analysis. The distribution of the C-alpha atom of the central residue of the fluorophore from these snapshots shows that GFP essentially moves within a hemisphere on the cytosolic side of the membrane beneath the dimer (SI Fig. 4).

Thermodynamic Ensemble (TE) using explicit linker/GFP configurations. From the explicit linker/GFP configurations, the thermodynamic ensemble (TE)-distribution is computed as a weighted average of the linker distance. The weights were determined according to a Boltzmann distribution

$$P_{\text{Boltzman}} = e^{\frac{-\triangle G}{RT}} \tag{8}$$

R is the gas constant, T is 300 K, and $\triangle G$ is the difference between the Gibbs energy of the current snapshot and the energetically most favorable one. G is determined as the difference between $E_{\text{effective, conf.}}$ (see section above) and the contribution from the configurational entropy S

$$G = E_{effective,conf.} - TS \tag{9}$$

We assumed that S is dominated by the configurations of the linker, whereas configurations of GFP are assumed to provide no contribution. This seems justified given that GFP is structurally much more stable than the linker: the linker largely consists of the TGR5 C-terminus, a part of GPCRs that has either been not fully resolved in any GPCR structure due to its high flexibility[53–55] or, when resolved in small parts, shows random coil formation[56]. Thus, we considered the linker a random hetero-polymer for which low energy conformations can structurally vary largely. Therefore, a random energy model[57] was used to describe its energy landscape. According to the random energy model, the entropy of a configuration with a given $E_{\text{effective, conf.}}$ is[57]

$$S = R \ \ln \ \Omega P \tag{10}$$

with Ω being the overall number of conformational states. The probability of occurrence P for each energy state is obtained from

$$P = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(E_{effective,conf.} - \mu)^2}{2\sigma^2}\right) \tag{11}$$

with μ being the mean and σ the standard deviation of the frequency distribution of $E_{\text{effective, conf.}}$. The assumption underlying eq. 11 is that the energy is Gaussian distributed[57], which is approximately fulfilled in our case (data not shown).

MM-PBSA calculations show a range of $E_{\text{effective, conf.}}$ of several hundred $kcal\ mol^{-1}$ for proteins of sizes similar to that used in the present study[58,59]. In agreement with this, $E_{\text{effective, conf.}}$ computed for the linker/GFP configurations attached to the TGR5 dimer spans a range of ~1.000 $kcal\ mol^{-1}$. However, such an energy range would lead to unrealistically low probabilities for the higher energy configurations. We thus linearly scaled $E_{\text{effective, conf.}}$ such that the linker/GFP configuration with the highest energy has a probability of occurrence in a Boltzmann distribution of $1/\Omega$ (SI Fig. 4). Finally with the scaled energies, P (eq. 11), S (eq. 10), and G (eq. 9) were calculated, and from these the weights according to eq. 8 for the weighted average of distances between 35 and 90 Å (SI Fig. 4).

To conclude, the TEs were constructed by explicit peptide linker/GFP MD simulations followed by calculations of conformational free energies (eqs 8–11) to weight each linker-GFP configuration. In the TE approach, the weights of the points obtained from the explicit linker model were used to assign the weights of the inter-probe distances.

References

1. Kawamata, Y. *et al.* A G protein-coupled receptor responsive to bile acids. *J Biol Chem* **278,** 9435–9440, doi: 10.1074/jbc.M209706200 (2003).
2. Keitel, V. & Haeussinger, D. Perspective: TGR5 (Gpbar-1) in liver physiology and disease. *Clinics and Research in Hepatology and Gastroenterology* **36,** 412–419, doi: 10.1016/j.clinre.2012.03.008 (2012).
3. Keitel, V., Reich, M. & Haussinger, D. TGR5: Pathogenetic Role and/or Therapeutic Target in Fibrosing Cholangitis? *Clin Rev Allerg Immu* **48,** 218–225, doi: 10.1007/s12016-014-8443-x (2015).
4. Pols, T. W., Noriega, L. G., Nomura, M., Auwerx, J. & Schoonjans, K. The bile acid membrane receptor TGR5 as an emerging target in metabolism and inflammation. *J Hepatol* **54,** 1263–1272, doi: 10.1016/j.jhep.2010.12.004 (2011).
5. Macchiarulo, A. *et al.* Probing the Binding Site of Bile Acids in TGR5. *ACS Med Chem Lett* **4,** 1158–1162, doi: 10.1021/ml400247k (2013).
6. Sindhu, T. & Srinivasan, P. Exploring the binding properties of agonists interacting with human TGR5 using structural modeling, molecular docking and dynamics simulations. *RSC advances* **5,** 14202–14213, doi: 10.1039/c4ra16617e (2015).
7. Yu, D. D. *et al.* Stereoselective synthesis, biological evaluation, and modeling of novel bile acid-derived G-protein coupled Bile acid receptor 1 (GP-BAR1, TGR5) agonists. *Bioorg Med Chem* **23,** 1613–1628, doi: 10.1016/j.bmc.2015.01.048 (2015).
8. Gertzen, C. G. W. *et al.* Mutational mapping of the transmembrane binding site of the G-protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur J Med Chem* **104,** 57–72, doi: 10.1016/j.ejmech.2015.09.024 (2015).
9. Spomer, L. *et al.* A Membrane-proximal, C-terminal alpha-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5. *J Biol Chem* **289,** 3689–3702, doi: 10.1074/jbc.M113.502344 (2014).
10. Ferre, S. *et al.* G Protein-Coupled Receptor Oligomerization Revisited: Functional and Pharmacological Perspectives. *Pharmacol Rev* **66,** 413–434, doi: 10.1124/pr.113.008052 (2014).
11. Lohse, M. J. Dimerization in GPCR mobility and signaling. *Curr Opin Pharmacol* **10,** 53–58, doi: 10.1016/j.coph.2009.10.007 (2010).
12. Rivero-Muller, A., Jonas, K. C., Hanyaloglu, A. C. & Huhtaniemi, I. In Oligomerization in Health and Disease Vol. 117 *Progress in Molecular Biology and Translational Science* (eds Giraldo, J. & Ciruela, F.) 163–185 (2013).
13. Bayburt, T. H., Leitz, A. J., Xie, G., Oprian, D. D. & Sligar, S. G. Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J Biol Chem* **282,** 14875–14881, doi: 10.1074/jbc.M701433200 (2007).
14. Whorton, M. R. *et al.* Efficient coupling of transducin to monomeric rhodopsin in a phospholipid bilayer. *J Biol Chem* **283,** 4387–4394, doi: 10.1074/jbc.M703346200 (2008).
15. Kuszak, A. J. *et al.* Purification and Functional Reconstitution of Monomeric mu-Opioid Receptors allosteric modulation of agonist binding by g(i2). *J Biol Chem* **284,** 26732–26741, doi: 10.1074/jbc.M109.026922 (2009).
16. Calebiro, D. *et al.* Single-molecule analysis of fluorescently labeled G-protein-coupled receptors reveals complexes with distinct dynamics and organization. *Proc Natl Acad Sci USA* **110,** 743–748, doi: 10.1073/pnas.1205798110 (2013).
17. Fotiadis, D. *et al.* Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421,** 127–128, doi: 10.1038/421127a (2003).
18. Johnston, J. M., Wang, H., Provasi, D. & Filizola, M. Assessing the Relative Stability of Dimer Interfaces in G Protein-Coupled Receptors. *PLoS Comput Biol* **8,** doi: e1002649 10.1371/journal.pcbi.1002649 (2012).
19. Filizola, M. & Weinstein, H. The study of G-protein coupled receptor oligomerization with computational modeling and bioinformatics. *FEBS J* **272,** 2926–2938, doi: 10.1111/j.1742-4658.2005.04730.x (2005).
20. Simpson, L. M., Taddese, B., Wall, I. D. & Reynolds, C. A. Bioinformatics and molecular modelling approaches to GPCR oligomerization. *Curr Opin Pharmacol* **10,** 30–37, doi: 10.1016/j.coph.2009.11.001 (2010).
21. Huang, J., Chen, S., Zhang, J. J. & Huang, X.-Y. Crystal structure of oligomeric beta(1)-adrenergic G protein-coupled receptors in ligand-free basal state. *Nat Struct Mol Biol* **20,** 419– +, doi: 10.1038/nsmb.2504 (2013).
22. Patowary, S. *et al.* The muscarinic M-3 acetylcholine receptor exists as two differently Sized complexes at the plasma membrane. *Biochem J* **452,** 303–312, doi: 10.1042/bj20121902 (2013).
23. Salom, D. *et al.* Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc Natl Acad Sci USA* **103,** 16123–16128, doi: 10.1073/pnas.0608022103 (2006).
24. Manglik, A. *et al.* Crystal structure of the mu-opioid receptor bound to a morphinan antagonist. *Nature* **485,** 321–U170, doi: 10.1038/nature10954 (2012).
25. Kudryavtsev, V. *et al.* Monitoring dynamic systems with multiparameter fluorescence imaging. *Anal Bioanal Chem* **387,** 71–82, doi: 10.1007/s00216-006-0917-0 (2007).
26. Kudryavtsev, V. *et al.* Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements. *ChemPhysChem* **13,** 1060–1078, doi: 10.1002/cphc.201100822 (2012).
27. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. In *Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B: Super-Resolution, Particle Tracking, Multiparameter, and Force Based Methods Vol. 475 Methods in Enzymology* (ed Walter, N. G.) 455–514 (2010).
28. Weidtkamp-Peters, S. *et al.* Multiparameter fluorescence image spectroscopy to study molecular interactions. *Photochem Photobiol Sci* **8,** 470–480, doi: 10.1039/b903245m (2009).
29. Stahl, Y. *et al.* Moderation of Arabidopsis Root Sternness by CLAVATA1 and ARABIDOPSIS CRINKLY4 Receptor Kinase Complexes. *Curr Biol* **23,** 362–371, doi: 10.1016/j.cub.2013.01.045 (2013).
30. Rovati, G. E., Capra, V. & Neubig, R. R. The highly conserved DRY motif of class A G protein-coupled receptors: Beyond the ground state. *Mol Pharmacol* **71,** 959–964, doi: 10.1124/mol.106.029470 (2007).
31. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. & Brunak, S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4,** 1633–1649, doi: 10.1002/pmic.200300771 (2004).
32. Aktrap, N., Seidel, T. & Barisas, B. G. Forster distances for fluorescence resonant energy transfer between mCherry and other visible fluorescent proteins. *Anal Biochem* **402,** 105–106, doi: 10.1016/j.ab.2010.03.026 (2010).
33. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate single-molecule FRET studies using multiparamter fluorescence detection. *Methods in Enzymology* **475,** 455–514, doi: 10.1016/s0076-6879(10)75018-7 (2010).
34. King, C., Sarabipour, S., Byrne, P., Leahy, D. J. & Hristova, K. The FRET Signatures of Noninteracting Proteins in Membranes: Simulations and Experiments. *Biophys J* **106,** 1309–1317, doi: 10.1016/j.bpj.2014.01.039 (2014).
35. Sindbert, S. *et al.* Accurate distance determination of nucleic acids via Forster resonance energy transfer: implications of dye linker length and rigidity. *J Am Chem Soc* **133,** 2463–2480, doi: 10.1021/ja105725e (2011).
36. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Methods* **9,** 1218–1225, doi: 10.1038/nmeth.2222 (2012).
37. Evers, T. H., van Dongen, E. M., Faesen, A. C., Meijer, E. W. & Merkx, M. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry* **45,** doi: 10.1021/bi061288t (2006).
38. Gertzen, C. G. W. *et al.* Mutational mapping of the transmembrane binding site of the G-Protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur J Med Chem* doi: 10.1016/j.ejmech.2015.09.024 (2015).
39. Grosberg, A. Y. & Khokhlov, A. R. *Statistical physics of macromolecules.* (AIP Press, 1994).
40. Unal, H. & Karnik, S. S. Domain coupling in GPCRs: the engine for induced conformational changes. *Trends Pharmacol Sci* **33,** 79–88, doi: 10.1016/j.tips.2011.09.007 (2012).

41. Auger, G. A., Pease, J. E., Shen, X. Y., Xanthou, G. & Barker, M. D. Alanine scanning mutagenesis of CCR3 reveals that the three intracellular loops are essential for functional receptor expression. *Eur J Immunol* **32**, 1052–1058, doi: 10.1002/1521-4141(200204)32:4<1052::aid-immu1052>3.0.co;2-l (2002).
42. Duboc, H., Tache, Y. & Hofmann, A. F. The bile acid TGR5 membrane receptor: From basic research to clinical application. *Dig Liver Dis* **46**, 302–312, doi: 10.1016/j.dld.2013.10.021 (2014).
43. Jensen, D. D. *et al.* The Bile Acid Receptor TGR5 Does Not Interact with beta-Arrestins or Traffic to Endosomes but Transmits Sustained Signals from Plasma Membrane Rafts. *J Biol Chem* **288**, 22942–22960, doi: 10.1074/jbc.M113.455774 (2013).
44. Jastrzebska, B. *et al.* Disruption of Rhodopsin Dimerization with Synthetic Peptides Targeting an Interaction Interface. *J Biol Chem* **290**, 25728–25744, doi: 10.1074/jbc.M115.662684 (2015).
45. Whorton, M. R. *et al.* A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc Natl Acad Sci USA* **104**, 7682–7687, doi: 10.1073/pnas.0611448104 (2007).
46. Scarselli, M. *et al.* Revealing G-protein-coupled receptor oligomerization at the single-molecule level through a nanoscopic lens: methods, dynamics and biological function. *FEBS J* **283**, 1197–1217, doi: 10.1111/febs.13577 (2016).
47. Hiller, C., Kuhhorn, J. & Gmeiner, P. Class A G-Protein-Coupled Receptor (GPCR) Dimers and Bivalent Ligands. *J Med Chem* **56**, 6542–6559, doi: 10.1021/jm4004335 (2013).
48. Homeyer, N. & Gohlke, H. FEW: A workflow tool for free energy calculations of ligand binding. *J Comput Chem* **34**, 965–973, doi: 10.1002/jcc.23218 (2013).
49. Homeyer, N. & Gohlke, H. Extension of the free energy workflow FEW towards implicit solvent/implicit membrane MM-PBSA calculations. *Biochim Biophys Acta, Gen Subj* **1850**, 972–982, doi: 10.1016/j.bbagen.2014.10.013 (2015).
50. Stern, H. A. & Feller, S. E. Calculation of the dielectric permittivity profile for a nonuniform system: application to a lipid bilayer simulation. *J Chem Phys* **118**, 3401–3412 (2003).
51. Nymeyer, H. & Zhou, H. X. A method to determine dielectric constants in nonhomogeneous systems: application to biological membranes. *Biophys J* **94**, 1185–1193, doi: 10.1529/biophysj.107.117770 (2008).
52. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98**, 10037–10041, doi: 10.1073/pnas.181342398 (2001).
53. Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217, doi: 10.1126/science.1164772 (2008).
54. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–745 (2000).
55. Hanson, M. A. *et al.* Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855, doi: 10.1126/science.1215904 (2012).
56. Wu, B. *et al.* Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **330**, 1066–1071, doi: 10.1126/science.1194396 (2010).
57. Wales, D. J. *Energy landscapes*. (Cambridge Molecular Science, 2003).
58. Fogolari, F. & Tosatto, S. C. Application of MM/PBSA colony free energy to loop decoy discrimination: toward correlation between energy and root mean square deviation. *Protein Sci* **14**, 889–901, doi: 10.1110/ps.041004105 (2005).
59. Gohlke, H. & Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem* **25**, 238–250, doi: 10.1002/jcc.10379 (2004).

Acknowledgements

We are grateful to Dr. Nadine Homeyer and Daniel Mulnaes, Heinrich-Heine-University, Düsseldorf, for technical help with FEW^{mem} and TopModel, respectively. We are also grateful to Manuel Frohnapfel and Lucas Wäschenbach for assistance during the TGR5 MFIS measurements and to Oleg Opanasyuk for scientific discussions. This work was supported by the Deutsche Forschungsgemeinschaft through the Collaborative Research Center SFB 974 ("Communication and Systems Relevance during Liver Damage and Regeneration", Düsseldorf) and INST 208/704-1 FUGG (to H.G.) to purchase the hybrid compute cluster used in this study. We are grateful for the computing time provided by the John von Neumann Institute for Computing (NIC) to H.G. on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) (project ID: 10766; user ID: HDD15).

Author Contributions

V.K., H.G. and C.A.M.S. conceived the study and supervised the project. A.G. and Q.M. performed the MFIS-FRET measurements. S.F. and A.G. analyzed the MFIS-FRET data. L.S. and C.W. performed molecular biology experiments and provided biological material. T.O.P. performed oligomer and tetramer simulations. C.G. and M.D. performed structural modeling of TGR5 dimers and tetramers and explicit and implicit linker simulations. H.G. analyzed molecular modeling and simulation data. A.G., C.G., V.K., H.G. and C.A.M.S. wrote the paper. All authors reviewed the manuscript. D.H. approved the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Greife, A. *et al.* Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study. *Sci. Rep.* **6**, 36792; doi: 10.1038/srep36792 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Publication II – Supporting Information

Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study

Annemarie Greife, Suren Felekyan, Qijun Ma, Christoph G.W. Gertzen, Lina Spomer, Mykola Dimura, Thomas O. Peulen, Christina Wöhler, Dieter Häussinger, Holger Gohlke, Verena Keitel, Claus A.M. Seidel

Scientific Reports, (2016), 6, 36792

Supporting Information

Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study

[1]Annemarie Greife, [1]Suren Felekyan, [1]Qijun Ma, [2]Christoph G.W. Gertzen, [3]Lina Spomer, [1]Mykola Dimura, [1]Thomas O. Peulen, [3]Christina Wöhler, [3]Dieter Häussinger, [2*]Holger Gohlke, [3*]Verena Keitel, [1*]Claus A.M. Seidel

[1] Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany;

[2] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany;

[3] Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

*Corresponding authors: C.A.M. Seidel, H. Gohlke, V. Keitel

Content

Supporting Tables

Supporting Table 1: Parameters for determination of the corrected green to yellow fluorescence intensity ratio F_D/F_A necessary for the 2D histograms. The background B was determined from untransfected cells. The green to yellow fluorescence intensity ratio (F_D/F_A) was corrected for crosstalk (characterized by the crosstalk factor α), background $\langle B \rangle$, detection efficiencies of D (g_G) and A (g_Y). The acceptor fluorescence used for 2D-FRET must also be corrected for additional direct acceptor excitation DE and relative concentration dependent brightness DE_{rel}. All samples were corrected for distinct fluorescence quantum yields Φ and a spectral shift factor γ (especially for TGR5 Y111A) which is considered in the corrected green detection efficiency (g_G^*).

	α	$\langle B_G \rangle$ [kHz]	$\langle B_Y \rangle$ [kHz]	γ	Φ	DE_{rel} [kHz]
wt	0.09	0.3	1	1	1	DA1:10= +0.76 DA1:40= +3.78
Y111A	0.28	0.3	2	0.61	1.125	DA1:10= +4.28 DA1:40= +0
Y111F	0.1	0.3	1	1	1	DA1:10= +1.76 DA1:40= +3.02

Supporting Table 2: Parameters for *ε(t)* diagram in **Fig. 4** for each TGR5 variant. The parameters b0-b4 are obtained from the fit equation $f = b_0 + b_1 \cdot e^{-\frac{x}{b_2}} + b_3 \cdot e^{-\frac{x}{b_4}}$. b_0 determines the Non-FRET fraction (Donly fraction), b_1 and b_3 are the two FRET fractions and b_2 and b_4 are the corresponding decay times. **Supporting Figure 2** shows how to generate and interprete *ε(t)* diagrams.

TGR5	Fit	DA1:3	DA1:5	DA1:10	DA1:20	DA1:40
wt	b_0:	0.90	0.84	0.89	0.81	0.82
	b_1:	0.01	0.04	0.03	0.04	0.04
	b_2:	0.30	0.18	0.23	0.33	0.46
	b_3:	0.09	0.13	0.08	0.16	0.15
	b_4:	5.70	8.04	1.89	3.80	5.12
Y111A	b_0:	0.42	0.42	0.45	0.43	0.43
	b_1:	0.02	0.01	0.19	0.03	0.03
	b_2:	0.73	0.63	5.28	0.56	0.72
	b_3:	0.55	0.56	0.36	0.55	0.55
	b_4:	6.09	5.02	5.26	4.48	4.90
Y111F	b_0:	0.73	0.74	0.69	0.68	0.77
	b_1:	0.11	0.06	0.18	0.09	0.02
	b_2:	5.73	2.7	3.88	2.30	0.43
	b_3:	0.15	0.19	0.13	0.23	0.21
	b_4:	771.45	90.75	765.84	24.83	2.80

Supporting Table 3: Sequence Information for AV-simulation. Untranslated region, TGR5 coding sequence, linker and GFP (4EUL) or mCherry (2H5Q) sequence of the analyzed TGR5 variants are summarized and used for TGR5 dimer and oligomer simulations. The position of the Y111 residue in the ERY motif for mutagenesis is highlighted. Sequences with unknown secondary or tertiary structures are underlined and are kept flexible in AV simulations.

	TGR5 wt-FP
5'UTR	none
TGR5	MTPNSTGEVPSPIPKGALGLSLALASLIITANLLLALGIAWDRRLRSPPAGCFFL SLLLAGLLTGLALPTLPGLWNQSRRGYWSCLLVYLAPNFSFLSLLANLLLVH GERYMAVLRPLQPPGSIRLALLLTWAGPLLFASLPALGWNHWTPGANCSSQA IFPAPYLYLEVYGLLLPAVGAAAFLSVRVLATAHRQLQDICRLERAVCRDEPS ALARALTWRQARAQAGAMetLLFGLCWGPYVATLLLSVLAYEQRPPLGPGTL LSLLSLGSASAAAVPVAMetGLGDQRYTAPWRAAAQRCLQGLWGRASRDSP GPSIAYHPSSQSSVDLDLNY
Cloning Linker	GSTGRH
GFP (4EUL) = donor (D)	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTG KLPVPWPTLVTTFGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDD GNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMAD KQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALS KDPNEKRDHMVLLEFVTAAGITLGMDELYK*
mCherry (2H5Q) = acceptor (A)	MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLK VTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNF EDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERM YPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDIT SHNEDYTIVEQYERAEGRHSTGGMDELY*

Supporting Table 4: Comparison of results from different fit models. Donor fluorescence lifetime decay histograms in presence of acceptor in TGR5 variants are fitted based on *2k-FRET* and AV simulated distance distributions for different dimer (1/8), (4/5) and oligomer (1/8)-4:5-(1/8) interfaces with only one fit parameter – Donly fraction. Interface dimer (4/5) yields very low Donly fractions compared to all other models: it is known that ~30 % of mCherry acceptor dyes are not active in cells; hence at least ~30 % Donly fractions are expected. Based on this we concluded that interface dimer (4/5) as the primary dimer interface in TGR5 variants are less likely. wt[1] is DA1:3, all other DA ratios are 1:20.

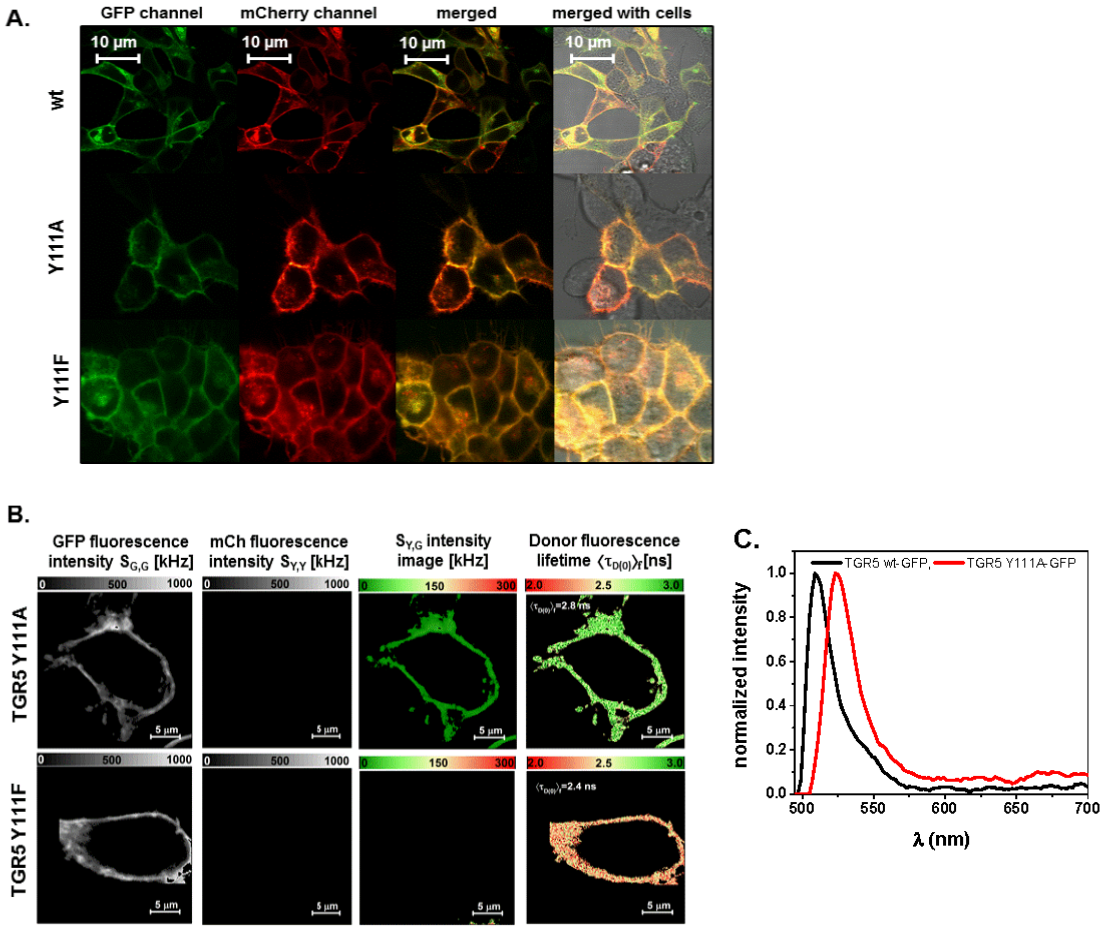
	wt[1]		**wt**		**Y111A**		**Y111F**	
	Donly fraction	X^2	Donly fraction	X^2	Donly fraction	X^2	Donly fraction	X^2
Two k_{FRET}	0.74	*1.61*	0.64	*1.59*	0.69	*1.39*	0.62	*1.61*
dimer (4/5)	0.43	*1.65*	0.00	*1.83*	0.00	*4.94*	0.14	*1.54*
dimer (1/8)	0.82	*1.91*	0.67	*2.68*	0.58	*2.61*	0.74	*1.83*
oligomer	n.d	*n.d*	0.74	*2.82*	0.65	*2.86*	0.78	*1.88*

Supporting Table 5: Overview of the mean distances $\langle R_{DA} \rangle$ calculated for the possible tetramer models with or without G-Protein. The primary interfaces for dimerization are in brackets (x/x) and secondary interfaces for oligomerization are abbreviated –x:x-. The numbers are the corresponding (transmembrane) helices involved in binding interactions. The apparent mean distances $\langle R_{DA} \rangle_{app}$ of each label pair involved in dimerization are bold. A schematic presentation of the tetramer models is shown in **Supporting Figure 4**. For example in the model (1/8)-4:5-(1/8) (with G-Protein always determined as label C, even when it is absent) label pairs A-E and B-D with the primary interfaces (1/8) show a distance 59-66 Å measured between the fluorescent proteins attached to helix 8, and the label pair A-B with the secondary interface -4:5- shows a mean distance of 133 Å. Further calculated distances in this oligomer are measured from label pairs A-D, B-E, B-D and are comparable to the distances obtained from A-B.

Label pair	$\langle R_{DA} \rangle$, [Å]					
	with G-Protein			without G-Protein		
	(1/8)-4:5-(1/8)	(1/8)-5:6-(1/8)	(5/6)-4:5-(5/6)	(1/8)-4:5-(1/8)	(1/8)-5:6-(1/8)	(5/6)-4:5-(5/6)
A-B	133	119	**128**	98	94	**103**
A-D	134	**63**	128	97	**64**	120
A-E	**66**	129	65	**58**	93	72
B-D	**59**	108	64	**57**	91	69
B-E	131	**80**	116	99	**71**	106
D-E	136	123	**104**	101	94	**103**

Supporting Figures

Supporting Figure 1

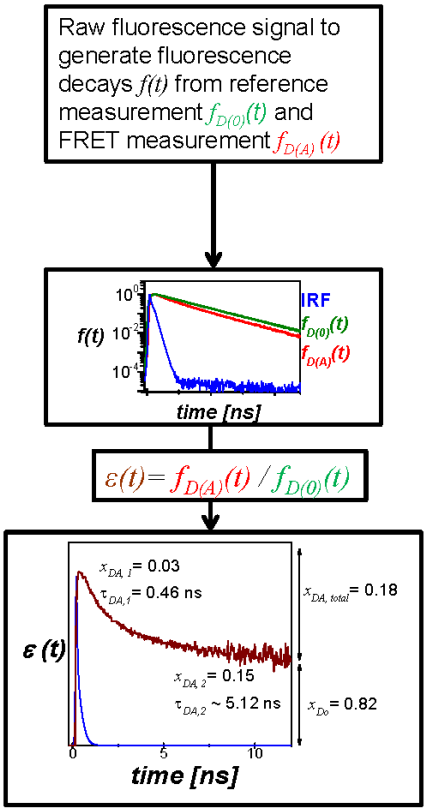


Supporting Figure 1: Live cell imaging and MFIS analysis of TGR5 donors

(a) HEK293 cells, transiently transfected with TGR5-GFP and TGR5-mCherry (transfection ratio 1:10), were imaged for co-localization of GFP and mCherry using sequential scanning and a scanning resolution of 1024 x 1024 pixels. Each TGR5-GFP and TGR5-mCherry picture is shown in a false color saturation mode and then overlaid by using green and yellow intensity colours. TGR5-GFP and TGR5-mCherry wt, Y111A and Y111F (from top to bottom) are clearly co-localized at the cell membrane. Scale bar 10 µm. **(b)** MFIS analysis of TGR5 transfected

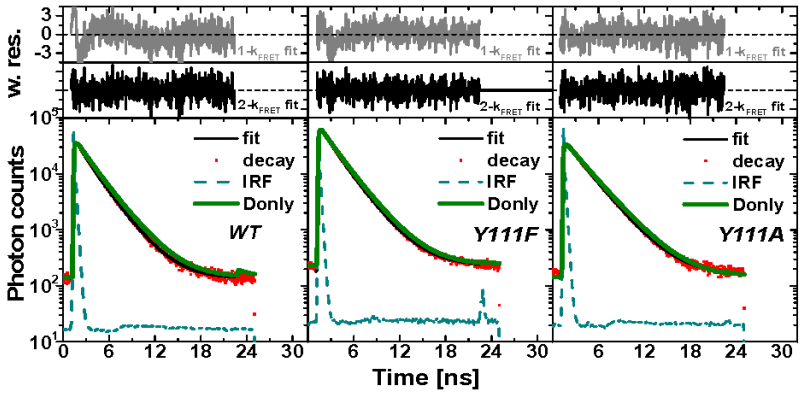
HEK293 cells by comparing (from left to right and top to bottom row) the GFP fluorescence intensity, mCherry fluorescence intensity, the donor fluorescence lifetime $\langle\tau_{D(0)}\rangle_f$, and mCherry photons after excitation of GFP ($S_{Y,G}$). The fluorescence-averaged donor lifetime in the absence of an acceptor $\langle\tau_{D(0)}\rangle_f$ in the Y111A variant is 2.8 ns compared to 2.4 ns for Y111F. The presence of green photons in the yellow channel is due to a higher crosstalk, background and red shift in the Y111A variant. **(c)** GFP was excited at 488 nm and emission spectrum was measured from 495 nm to 700 nm in a 2 nm step size and a 2 nm spectral band width at Olympus FluoView1000 microscope. TGR5 wt-GFP shows the typical emission maximum at 510 nm, whereas TGR5 Y111A-GFP shows a 13 nm red shift towards 523 nm. Three cells for each curve were measured, the background was subtracted and the average intensity normalized to the maximum. The Y111A MFIS data were corrected for the spectral shift.

Supporting Figure 2



Supporting Figure 2. Guideline for presentation and interpretation of *ε(t)* diagrams: In the first step, the raw fluorescence signal decays *f(t)* from the reference measurement $f_{D(0)}(t)$ (green) and from the FRET measurement $f_{D(A)}(t)$ (red) are selected and corrected with the instrument response function curve (*IRF*, blue) for a time shift. In the second step, the $f_{D(A)}(t)$ decay is divide through $f_{D(0)}(t)$ decay (**eqation (1) in main text**). The resulting decay $\varepsilon(t)$ is normalized to 1 and plotted versus time in ns. As example the TGR5 wt DA1:40 experiment is used. In this case the non-FRET fraction x_D (b0 in Table S2) is 0.82. The corresponding total FRET fraction $x_{DA,total}$ is 0.18. As the decay is clearly bi-exponential, two FRET fractions $x_{DA,1}$, $x_{DA,2}$ (b1, b3) and the corresponding fluorescence lifetimes $\tau_{DA,1}$, $\tau_{DA,2}$ (b2, b4) are resolvable by the equation used in Table S2.

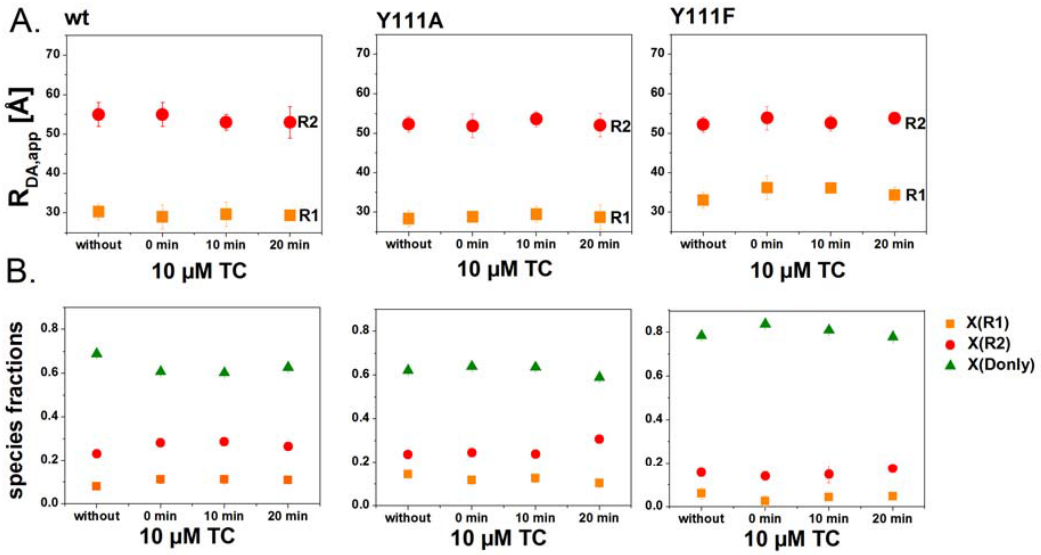
Supporting Figure 3



Supporting Figure 3: Fit fluorescence decays with different models for TGR5 variants.

Fitting the sub-ensemble fluorescence decays of the FRET samples (DA1:10) with k_{FRET} models showed that two FRET rates are necessary to fit these data accurately. The decays of Donly (TGR5-GFP) and FRET samples are in olive and red, respectively. The fitted decay with the 2-k_{FRET} model and the fitting residuals are plotted in black. The fitting residuals with 1-k_{FRET} model are plotted in grey.

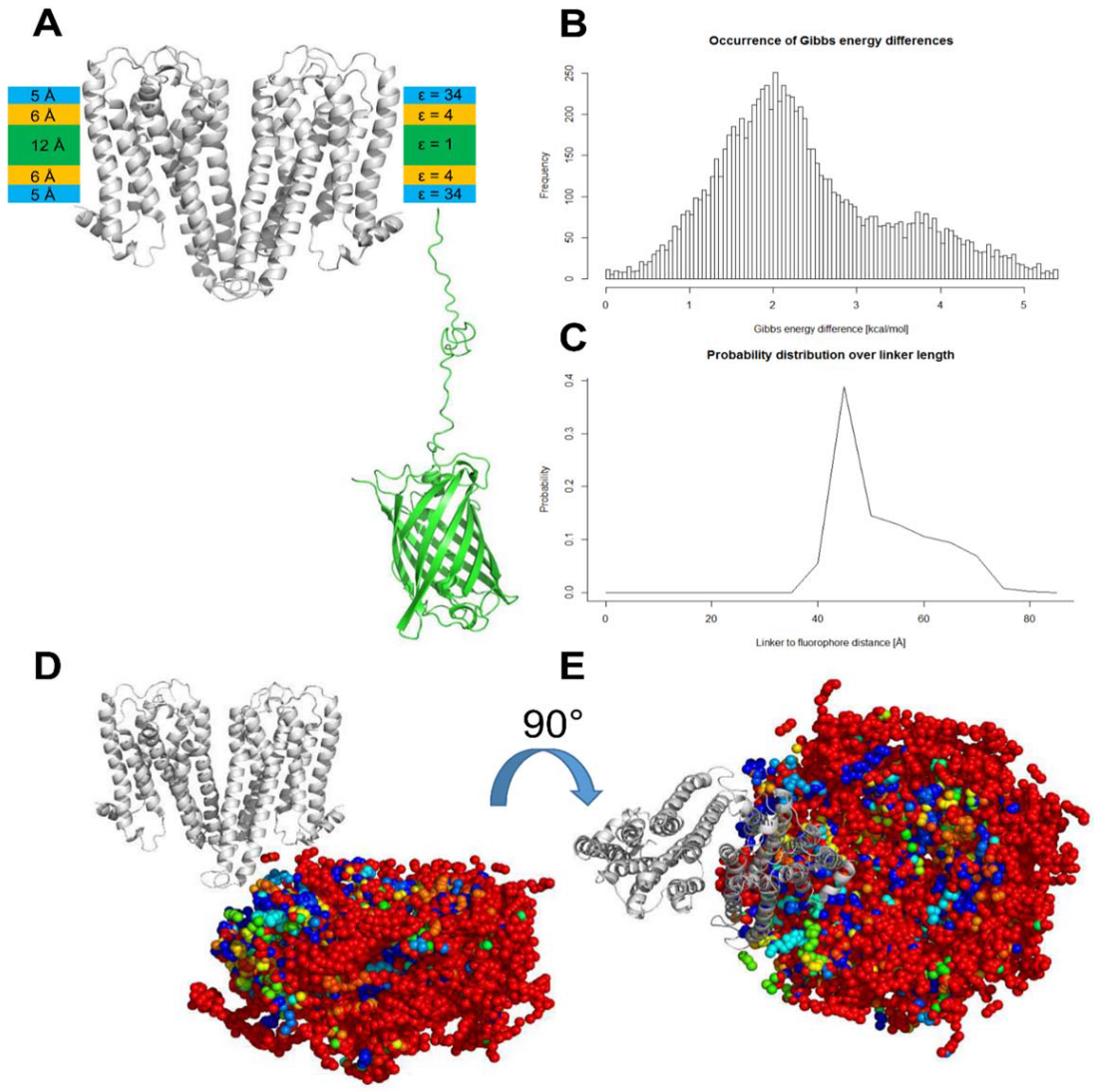
Supporting Figure 4



Supporting Figure 4: Time-series analysis after TC ligand stimulation.

HEK293 cells were transiently transfected with TGR5-GFP alone (Donly) or with TGR5-GFP and TGR5-mCherry with the D/A ratio of 1:10. To study changes in FRET after ligand addition, three cells were selected using the Olympus Time laps function and measured at different time points, including before adding 10 µM TC (without), immediately after TC addition (t=0), 10 min after and 20 min after. The apparent distances R_{DA} species fractions were fitted using self-made software. **(a)** The apparent distances are plotted versus time. Each point represents the average of nine cells (three measurements with three cells). **(b)** The species fractions X(R1), X(R2) and the Non-FRET fraction X(Donly) at four time points (representing without TC, t=0, t=10 min and t=20 min) are plotted, but no substantial change due to ligand addition could be detected. Orange=R1=high FRET distance, red= R2= low FRET distance, green=Donly fraction.

Supporting Figure 5

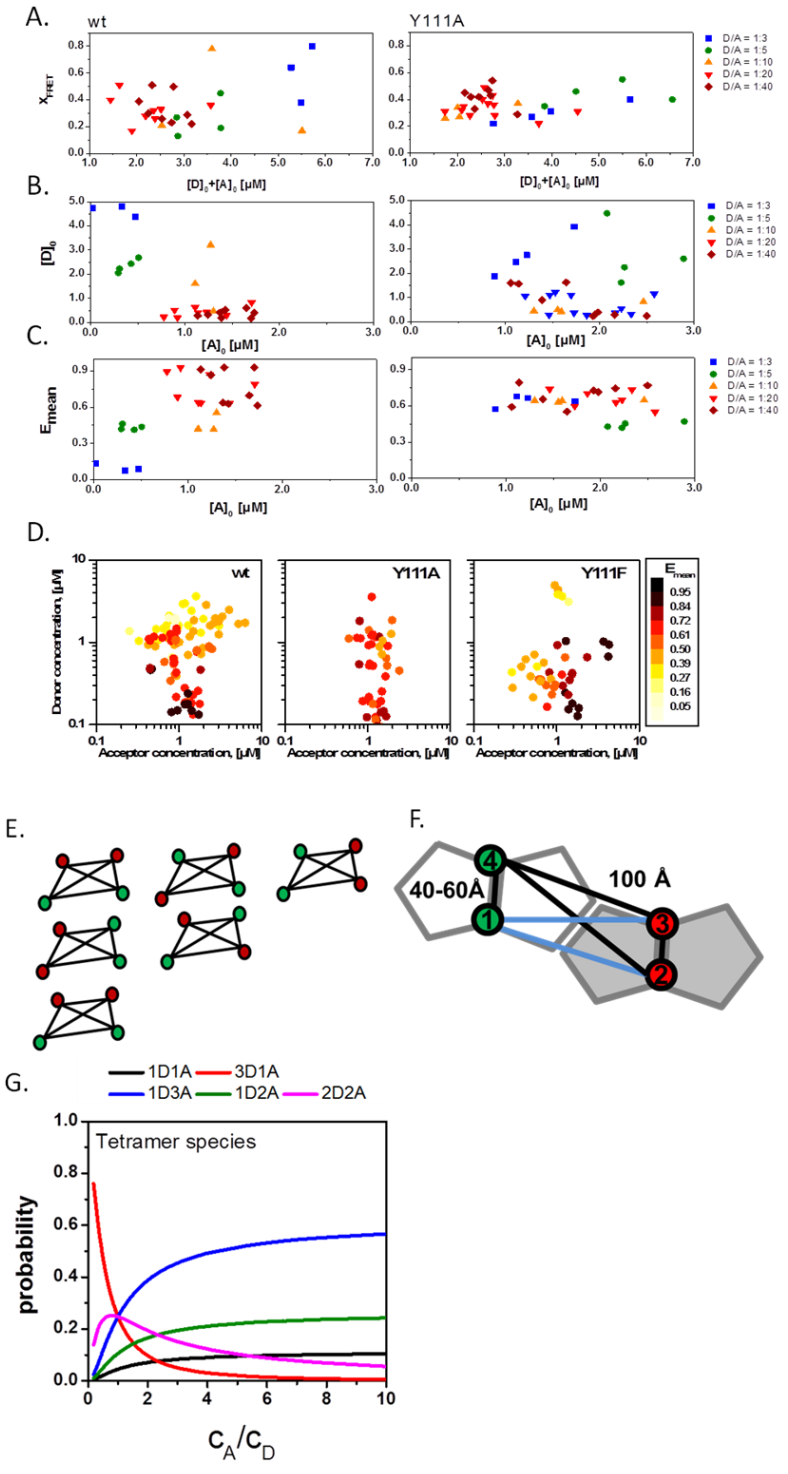


Supporting Figure 5: Explicit linker/GFP simulation and probability distribution of linker/GFP configurations.

(a) Starting structure of the TGR5 4/5 dimer (grey) and of the linker and GFP after the initial minimization (green). The linker and GFP were simulated separately from the TGR5 dimer; the structure shown here illustrates one of the composite models used for the MM-PBSA

calculations. At the 'wad' in the middle of the linker, several proline residues are present. The positioning of the implicit membrane slabs is shown in colored bars next to the TGR5 dimer. The bars on the left show the thickness of each membrane layer used in the FEW^{mem} calculations, while the bars on the right show the respective electric permittivity. **(b)** Frequency distribution of Gibbs energies (**equation (9)** in the main text) relative to the energetically most favorable snapshot after linear scaling (see main text). **(c)** Probability distribution of the Boltzmann-weighted distance between the fluorophore and the N-terminus of the linker. **(d)** + **(e)**. Ensemble of linker/GFP configurations represented in terms of the C-alpha atom of the central residue of the fluorophore generated by MD simulations with added rotations in relation to the TGR5 4/5 dimer (grey) in side (**d**) and exoplasmic view (**e**). The coloring of the C-alpha atoms corresponds to their probability ranging from lowest (blue) to highest (red). Conformations with a low probability are more frequently found in close contact to the dimer.

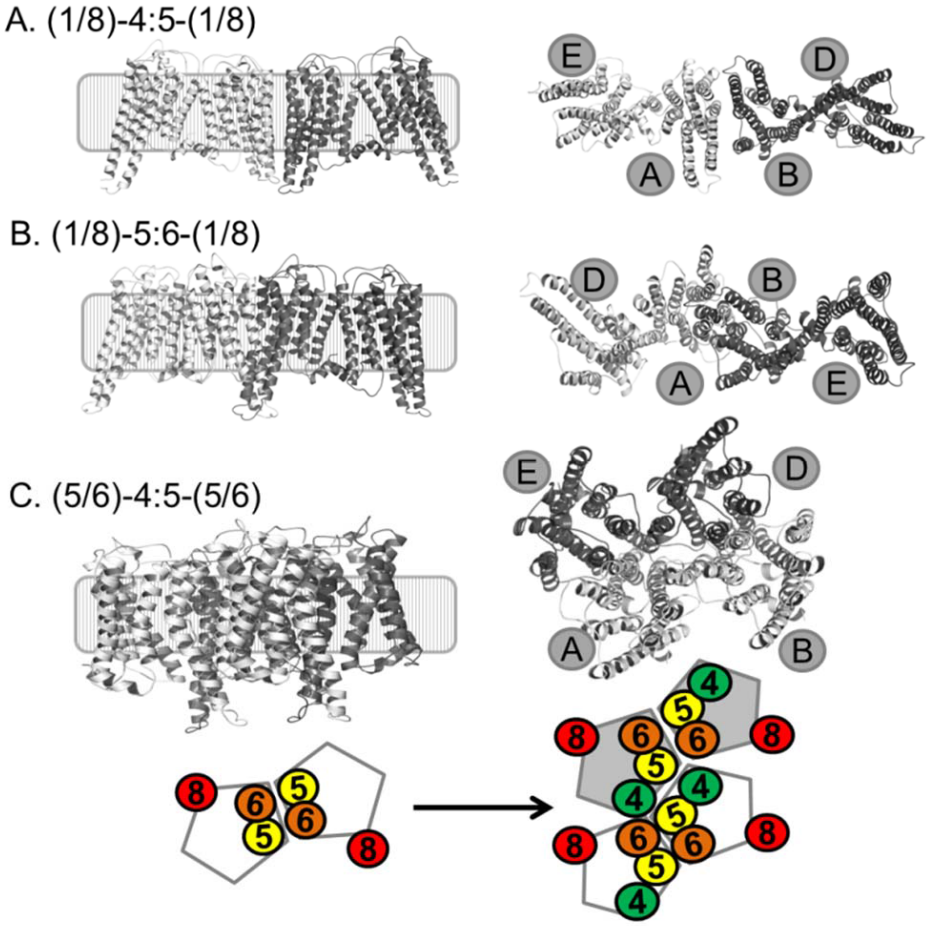
Supporting Figure 6



Supporting Figure 6: **Characterization and estimation of the association constants with a dimer/tetramer fit model**.

HEK293 cells were transiently transfected with TGR5 wt (left plots) or Y111A (right plot) donor to acceptor ratios varying from 1:3 to 1:40. **(a)** The total protein concentration $[D]_0+[A]_0$ (**eq.(4-6)**) and the FRET species fractions x_{FRET} were obtained from MFIS measurements and plotted to calculate the dissociation constant K_D. The FRET species fractions calculated from different D/A ratios were distributed equally in a concentration range of 1-7 µM. From these data K_D cannot be directly determined. The upper limit for K_D should be less than 1 µM. **(b)** The real donor $[D]_0$ and acceptor $[A]_0$ concentrations from the D/A transfection experiments were plotted for wt and Y111A to estimate differences in experimental and real concentration ratios between donor and acceptor. **(c)** E_{mean} increases in an $[A]_0$ dependent manner in wt but not in Y111A transfected cells. **(d)** Overview on the concentration ranges of donor and acceptor and its influence on E_{mean}, whose size is depicted in color. Variant specific interaction patterns are readily visible. **(e)** Description of our data by a minimal dimer/tetramer model to. In this model we assume that a tetramer is constituted of a dimer of dimers. In a tetramer the sum of donor, acceptor and unlabeled molecules is constant (**eqs. (7-9)**). Six tetramer configurations for a case of two acceptor (red) and two donor molecules (green) are possible. **(f)** Composition of a simplified rectangular tetramer molecule with random arrangement of two donors and two acceptors according to a linear organization of the GPCR. The positions of the green and red circles in the pentagram represent the fluorescent proteins attached to helix 8. **(g)** Probabilities of all tetramer species composed of a certain number of donor and acceptors (1D1A, 3D1A, 1D3A, 1D2A, 2D2A) in dependence of the acceptor to donor ratio. In our case the most probable scenario is the 2D2A case which describes our data best.

Supporting Figure 7



Supporting Figure 7: GPCR tetramer organization and AV simulations.

(a)-(c) Cartoon presentation made with the free software PyMol [1,2] from the membrane view (right) and cytoplasmic view (left) for three possible tetramer organizations. The labels A, B, D and E refer to the TGR5 monomers and are used for distance distribution calculations (**Supporting Table 5**). The corresponding dimers are colored in light grey or dark grey.

Supporting Notes

Proximity FRET

Pixel-wise analysis of the fluorescence data in TGR5 Y111A compared to wt and Y111F showed strong differences in the FRET properties, which were only detectable in an acceptor concentration-dependent manner (**Figure 4,** main text). Thus, we tested whether the observed FRET could simply be caused by a very high local concentration of acceptor proteins in the membrane, so that donor and acceptor are in proximity even though they do not interact. This phenomenon is called "*proximity FRET*".

Due to the single-molecule sensitivity of our confocal microscope, we could perform FRET experiments with acceptor concentrations of ~1-6 µM in 1.23 fl, which corresponds to a molecule density of less than ~0.02 acceptor molecules/nm^2. According to King *et al.*[3], proximity FRET is negligible ($E < 0.1$) at these concentrations.

The pixel-integrated, time-resolved FRET analysis $\varepsilon(t)$ supported the pixel-wise analysis and clearly demonstrated the presence of different FRET species in TGR5 wt and Y111F and therefore the formation of higher-order oligomers as compared to Y111A.

Supporting Methods

1. Molecular biology

Cell culture reagents

Cell culture media were from PAA (Coelbe, Germany). Foetal calf serum (FCS) was from Biochrom (Berlin, Germany). Taurolithocholic acid (TLC), Taurocholic acid (TC) and Forskolin (F) were purchased from Sigma-Aldrich (Taufkirchen, Germany) and Calbiochem (San Diego, CA, USA), respectively.

Cloning of TGR5

Human TGR5 was cloned as previously described [4]. Constructs were cloned into the pcDNA3.1+ (TGR5-pcDNA: complete CDS; TGR5-His: stop codon in complete CDS replaced by C-terminal 8xHis-tag), pGFP-N1, and pmcherry-N1 (stop codon in the complete CDS replaced by a restriction site) vectors. The FLAG-TGR5-YFP construct with an N-terminal FLAG-tag and a C-terminal YFP-tag was cloned into the pEYFP-N1 vector. All vectors were from Clontech, Palo Alto, CA. The Y111A and Y111F mutations were introduced into different TGR5 cDNA constructs using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, USA) [4]. All cloning strategies and mutagenesis primer sequences can be obtained upon request. Successful cloning and mutagenesis was verified by sequencing (GenBank accession numbers: TGR5:NM_001077191.1).

Immunofluorescence and confocal laser scanning microscopy

Human embryonic kidney 293 (HEK293) cells and Madin Darbin canine kidney cells (MDCK), grown on glass coverslips or transparent filter wells, were transiently transfected with TGR5 wt, Y111A or Y111F in pcDNA3.1+ and pEYFP-N1 vectors using Lipofectamine2000 (Invitrogen)

for 48 h according to the manufacturer's recommendations. After fixation with -20°C cold methanol for 30 sec, cells were incubated with RVLR2 [5] antibody against TGR5 and Cyanine-3 (1:500) conjugated secondary antibodies, which were purchased from Dianova (Hamburg, Germany). Nuclei were stained with Hoechst 34580 (1:20000; Invitrogen). Images were analysed on a Zeiss LSM510META confocal microscope using a multi-tracking modus. A 63 x objective and a scanning resolution of 1024 x 1024 pixels was used for all samples.

Flow cytometry

TGR5 plasma membrane protein amount was quantified by flow cytometry (FACS) using a FACS-CANTO-II (BD Biosciences; Heidelberg, Germany) as previously described [4,6]. HEK293 cells were transiently transfected with the FLAG-TGR5-YFP constructs using Lipofectamine2000. The N-terminal FLAG-tag was detected with the anti-FLAG M2-antibody (Sigma-Aldrich) using the Zenon PacificBlue Label-Kit (Invitrogen) according to manufacturer's instructions. TGR5 plasma membrane expression was calculated by the amount of FLAG-tag positive cells divided by the total amount of TGR5 positive cells as determined by YFP-fluorescence.

Reporter gene assay

HEK293 cells were transiently transfected with TGR5 wt and TGR5 Y111A and TGR5 Y111F variants in the pcDNA3.1+ construct (0.5 ug), pEYFP-N1-empty vector (1.1 ug), reporter PlasmLuc (1.6 ug; Bayer AG; Leverkusen, Germany), and Renilla expression vector (0.1 µg; Promega; Madison, WI, USA) using Lipofectamine2000. The PlasmLuc-reporter gene construct contains 5 cAMP-responsive elements (CREs) upstream of the luciferase gene. Luciferase activity was normalized to transfection efficacy, which was monitored by cotransfection with the Renilla expression vector, and served as measure for the rise in intracellular cAMP. Luciferase

activity was determined 16 hours after stimulation with DMSO, TLC or Forskolin [4,6]. The increase in TLC- and Forskolin-dependent luciferase activity is relative to the DMSO stimulation, which was set equal to 1.0.

Co-immunoprecipitation

HEK293 cells were cotransfected with TGR5-YFP and TGR5-His. Cells transfected with the empty vector (pcDNA or pEYFP-N1) and only one of the TGR5 cDNAs (TGR5-His or TGR5-YFP) served as controls. Cells were lysed with a buffer containing 50 mM Tris-HCl pH 8.0, 1% Nonidet® P40 (AppliChem) and complete protease inhibitor cocktail tablets (Roche). Protein concentration was determined by Bradford assay, and 0.05 mg protein from each sample was set aside as input control. 1.6 mg protein from each sample was used for immunoprecipitation with the µMACS His-tagged protein isolation kit (Miltenyi Biotec, Bergisch-Gladbach, Germany). His-tagged TGR5 was labelled with the anti-His microbeads and loaded onto the MACS columns. His-tagged proteins were eluted from the columns with 60 µl elution buffer and divided into two equal samples of 25 µl each. These as well as the input control samples were subjected to deglycosylation using the N-glycosidase-F Kit (Roche Diagnostics, Mannheim, Germany) for 10 min at 37°C. The deglycosylation reaction was stopped with 10% Laemmli buffer, heated to 95°C for 3 min. IP samples and input controls were separated by SDS page and proteins were transferred to PVDF membranes. His-tagged proteins were detected with the HRP-coupled anti-His antibody (dilution 1:5000, Miltenyi Biotec). YFP-coupled proteins were detected using the HRP-coupled anti-GFP antibody (dilution 1:5000, Miltenyi Biotec). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was detected with an antibody from GeneTex (dilution 1:10000) and a secondary HRP-coupled anti-mouse antibody (dilution 1:10000, Dako). Densitometry was performed using the Totallab-100 software (Nonlinear Dynamics, Durham,

NC). The relative amount of TGR5 oligomerization was calculated by dividing the amount of TGR5-YFP protein through the amount of TGR5-His protein. Wildtype TGR5-YFP/TGR5-His was set to 1.0.

2. MFIS-FRET: microscopy and analysis

Sample preparation for MFIS-FRET experiments

For live cell experiments HEK293 were seeded on 8 well chambered glass slides (Labtek, Nunc, USA) one day before transfection. Cells were transiently transfected with 0.5 µg DNA at a density of about 80% using FuGene6 (Promega) according to the manufacturer's protocol 24 to 48 h before analysis. Cell vitality and successful transfection was visually inspected before MFIS measurements.

Microscope calibration

Calibration measurements with Rhodamine 110 delivered the G-factor $G = Sg_{\perp}/Sg_{\parallel}$ for the GFP emission wavelength range (green channels). The G-factor accounts for the detection efficiency difference between detectors of both polarizations ($g_{\perp}$ and $g_{\parallel}$). The instrument response function (IRF) was measured with the back-reflection of the laser beam using a mirror and was used for iterative re-convolution in the fitting process. Furthermore, untransfected cells and water were measured at 488 nm and 559 nm for background determination.

Time series experiments of TGR5 stimulation by Taurocholic acid (TC)

To study the effect of bile acid agonists on the FRET parameters we used the water-soluble ligand TC, because addition of DMSO (necessary to dissolve TLC) affects the fluorescence signal significantly. For the time series experiments the time laps viewer function supplied by Olympus LSM was used. The motorized table was calibrated, and three cells were selected and

monitored over a 40 minutes time period. FRET measurements were taken every 10 minutes: before the addition of TC immediately after addition (t = 0 min), and after ten and twenty minutes (t = 10 min; t = 20 min). Cells were excited with 488 nm and 559 nm laser light as described above. Where necessary, changes in focus and system drift were corrected.

Pixel-wise analysis

To determine fluorescence-weighted lifetimes in a pixel-wise analysis, the histograms presenting the decay of fluorescence intensity after the excitation pulse were built for each pixel with 128 ps per bin. We used a maximum likelihood estimator (MLE) to determine the fluorescence-weighted averaged lifetime of donor molecules $\langle \tau_{D(A)} \rangle_f$ in a single pixel using a model function containing only two variables, $\langle \tau_{D(A)} \rangle_f$ and the scatter contribution fraction.

MFIS-FRET 2D histograms

For oligomerization analysis, we plotted the 2D histograms of donor lifetime $\langle \tau_{D(A)} \rangle_f$ vs the green to yellow fluorescence intensity ratio (F_D/F_A) (see equations (2) and (3)) corrected for crosstalk (characterized by the crosstalk factor α), background $\langle B \rangle$, detection efficiencies of D (g_G) and A (g_Y). The acceptor fluorescence used for 2D-FRET must also be corrected for additional direct acceptor excitation DE and relative concentration dependent brightness DE_{rel}. Furthermore all samples were corrected for distinct fluorescence quantum yields Φ and a spectral shift factor γ (especially for TGR5 Y111A) which is considered in the corrected green detection efficiency (g_G*).

$$F_D = \frac{S_G - \langle B_G \rangle}{g_G^*} \qquad (1)$$

$$F_A = \frac{S_Y - (\langle B_Y \rangle + DE_{rel}) - \alpha(S_G - \langle B_G \rangle)}{g_Y} \quad (2)$$

The crosstalk factor α is determined as the ratio between donor photons detected in the yellow channels and those detected in the green channels for the Donor only (Donly) labeled sample. The corrected detection efficiency g_G* is determined as the ratio of the spectral shift influenced by green detection (0.69) and expected green detection (1.12) multiplied with the quantum yield Φ_{Y111A} obtained from a self-made detection efficiency software. The F_D/F_A parameters for each variant are provided in **Supporting Table 1**

The simultaneous reduction in both FRET indicators $\langle \tau_{D(A)} \rangle_f$ and (F_D/F_A) indicate FRET due to proteins interaction. For a given sub-population selection of the donor fluorescence decay histogram with 32 ps time resolution was constructed for further pixel-integrated sub-ensemble analysis, and the species-averaged fluorescence lifetime of the donor $\langle \tau_{D(A)} \rangle_x$ was calculated based on fit results (species fractions x_i and lifetimes $\tau_{D(A)}$)

$$\langle \tau_{D(A)} \rangle_x = \sum_{i=1}^{n} x_i \cdot \tau_{D(A),i} \quad (3)$$

n is the number of exponents used in donor fluorescence lifetime fitting.

Determination of acceptor and donor concentration from MFIS experiments

TGR5 monomers were either labelled with donor or acceptor fluorescent proteins and transiently transfected into cells with different donor-to-acceptor concentration ratios. The fractions of active donor (denoted as D) and active acceptor (denoted as A) is f_D and f_A respectively. The rest are inactive FPs, which we considered as dark (i.e. no fluorescence emission) and dysfunctional (i.e. FRET-negative). To calculate the protein concentrations from fluorescence intensity, the detection volume of our microscope and GFP and mCherry brightness are required. The

detection volume was determined as $1.23*10^{-15}$ l from FCS measurements of Cyanine 3B (Cy3B). The fitting model applied to the obtained FCS curve assumes a 3-dimensional Gaussian-shaped volume, and a single diffusing species including transitions to a triplet state as described in [7]. The brightness of enhanced GFP and mCherry *in vivo* were individually characterized from FCS measurements of freely diffusing FPs in cytoplasm. We found that with 0.6 µW of 559 nm laser excitation at the objective, mCherry brightness is 0.68 kcpm in cytoplasm. With 0.4 µW of 485 nm laser excitation at the objective, GFP brightness is 0.56 kcpm in cytoplasm.

The average mCherry fluorescence intensity of an image with mCherry excitation ($S_{Y,Y}$) was first corrected for detector dead time, and then used to calculate the total concentration of mCherry, $[A]_0$, with the determined detection volume and the mCherry brightness:

$$[A]_0 = \frac{S^m_{Y,Y}}{brightness\ [kcpm] * confocal\ volume\ [fl]} \tag{4}$$

$$= \frac{S^m_{Y,Y}}{0.68\ kcpm * 0.8\ fl}$$

The average GFP fluorescence intensity of an image with GFP excitation was also corrected for detector dead time, and then the obtained intensity ($S^m_{G,G}$) was further corrected for the quenching effect due to FRET:

$$S_{G,G} = \frac{S^m_{G,G}}{(1 - x_{FRET}) + x_{FRET} \cdot (1 - E)} \tag{5}$$

$S_{G,G}$ is the unquenched GFP fluorescence intensity in the absence of FRET, the energy transfer efficiency E and fraction of FRET-active population, x_{FRET}, were calculated as described in the main method sections. $S_{G,G}$ was then used to calculate the total concentration of GFP, $[D]_0$. The wavelength dependent confocal volume is 0.5 fl.

Assuming the concentration of the FPs reflects the concentration of their host proteins, the TGR5 concentration (without non-fluorescent molecules) in µM was determined as:

$$\text{protein concentration} = [A]_0 + [D]_0 = c_A + c_D \qquad (6)$$

Estimation of the association constants for oligomerization

The total protein concentration and the protein association constants have to be considered to determine the oligomerization state or the chemical speciation. To calculate the transfer-efficiency for a given oligomerization the spatial organization of the molecules within the oligomers and the concentration of donor, acceptor and non-fluorescent molecules has to be considered. The total protein concentration (**equation (6)**) is given by the sum of the acceptor, the donor and the unlabeled protein concentration:

$$c_T = c_A + c_D + c_U \qquad (7)$$

Here the unlabeled protein concentration c_u equals the concentration of immature mCherry. The protein concentrations were calculated using the brightness of free GFP and free mCherry as reference Even though higher-order oligomerization is anticipated we used a simple dimer/tetramer model to describe our data as this allows for a quantitative description. In this model we assume that a tetramer is constituted of a dimer of dimers (**Supporting Fig. 6**). Hence, starting from a monomer two equilibriums have to be treated:

$$o + o \rightarrow oo \quad K_1 = \frac{[oo]}{[o][o]} \qquad (8)$$

$$oo + oo \rightarrow (oo)(oo) \quad K_2 = \frac{[(oo)(oo)]}{[oo][oo]}$$

Here o is a monomer, oo a dimer and (oo)(oo) is a tetramer. We use the monomer o as a master species. Then the total protein concentration is given by:

$$c_T = [o] + 2 \cdot [oo] + 4 \cdot [(oo)(oo)] \tag{9}$$

Now, the concentrations of the three species o, oo and (oo)(oo) for any given the total protein concentration is obtained by solving the three equations above.

To calculate the transfer efficiency we assume that donor, acceptor and unlabeled molecules behave biochemically identical. Hence, the probability of an oligomer composition is given by the probability of finding a donor, acceptor or unlabeled molecule and the counting statistics. The probabilities of finding a donor, acceptor or unlabeled molecule depend on their respective concentrations. For instance the probability of an acceptor molecule is given by the respective species and total protein concentration:

$$p_A = \frac{c_A}{c_T} \tag{10}$$

In a tetramer the sum of donor, acceptor and unlabeled molecules is constant. The probability of a certain tetramer composition is obtained by the multinomial distribution:

$$\begin{aligned} p(n_D, n_A, n_U) &= N \cdot p_D^{n_D} p_A^{n_A} p_U^{n_U} \\ &= \frac{(n_D + n_A + n_U)!}{n_D! n_A! n_U!} \cdot p_D^{n_D} p_A^{n_A} p_U^{n_U} \end{aligned} \tag{11}$$

N is the number of combinations for a given composition. Each combination might have a different FRET-rate distribution. Hence, in case of two donors and two acceptors 6 combinations as shown in **Supporting Fig. 6** contribute to the signal. If only species with at least one donor and one acceptor are considered the FRET-rate constants of overall 38 distinct species and their respective probabilities and FRET-rate constant distributions have to be calculated. The species

probabilities summarized by their donor and acceptor composition in dependence of the acceptor to donor ratio c_A/c_D are illustrated in **Supporting Fig. 6**.

FRET-rate constants are additive. Therefore in case of multiple acceptors the total FRET-rate constant experienced by a donor (i) is given by the sum of all FRET-rate constants of all acceptors (j):

$$k_{RET}^{(i)} = \frac{1}{\tau_0} \cdot \sum_j \left(\frac{R_{DA}^{(ij)}}{R_0} \right)^6 \tag{12}$$

Here $R_{DA}^{(ij)}$ is the donor acceptor distance between the donor (i) and the acceptor (j) which is determined by the spatial arrangement of the oligomer. For instance, in the case as illustrated in **Supporting Fig. 6** the two FRET-rates experienced by the donor at position 1 and the donor at position 4 are given by:

$$k_{RET}^{(1)} = \frac{1}{\tau_0} \cdot \left(\left(\frac{R_{DA}^{(13)}}{R_0} \right)^6 + \left(\frac{R_{DA}^{(12)}}{R_0} \right)^6 \right) \tag{13}$$

$$k_{RET}^{(4)} = \frac{1}{\tau_0} \cdot \left(\left(\frac{R_{DA}^{(42)}}{R_0} \right)^6 + \left(\frac{R_{DA}^{(43)}}{R_0} \right)^6 \right)$$

These FRET-rates result in first approximation in bi-exponential fluorescence decay, if the coupling between the two donors is not considered.

For a given structural arrangement all FRET-rate constants for all possible compositions (one donor one acceptor, two donors one acceptor, etc.) were calculated (**Supporting Fig. 6**). Later the average transfer-efficiencies of the tetramer compositions containing at least one donor and one acceptor were calculated.

It has to be considered that the contribution to the fluorescence signal depends on the number of donor molecules. For instance a tetramer constituted out of three donors and one acceptor molecule contributes three times more to the total signal as compared to a tetramer only constituted out of one donor, one acceptor and two unlabeled molecules.

The predicted transfer efficiency for each data point depends now only on the equilibrium association constants K_1, K_2 and the spatial arrangement of the fluorophores in the dimer and the tetramer. To reduce the number of free parameters we assumed that the tetramer can be described by a rectangular geometry where one edge is approximately 100 Å long while the second edge is between 40-50 Å long (**Supporting Fig. 6**). This assumption is in line with the homology models (**Supporting Table 5** and **Supporting Fig. 7**). Furthermore, only FRET molecules have been selected. Therefore, the first equilibrium from monomer to dimer is not monitored and only the equilibrium constant of the tetramer formation is probed. Thus, only K_2 and the dimer distance in the range of 40-60 Å is reflected in the data. For the measurements we find that a short distance of approximately 45 Å describes the data best. For the TGR5 wt and Y111F variant we find predominately a tetrameric or higher-order oligomer configuration while in case of the Y111A mutant the molecules are predominately in a dimeric configuration.

Statistical analysis

Experiments were performed independently at least three times. For MFIS-FRET at least nine cells per series in three independent experiments were measured. Results are expressed as mean ± standard error of the mean (SEM) and analysed using the two-sided student t-test. A $p \leq 0.01$ was considered statistically significant.

3. Molecular modelling and simulation

Structural models of TGR5 dimers and tetramers

Dimer models with the interface TM1 and H8 (1/8) were generated by structurally aligning two homology models of TGR5 [8] onto the dimeric crystal structure of the human κ-opioid receptor (PDB ID: 4DJH [9]) via the 'cealign' command in Pymol [2]. For dimer models with the 4/5 interface and the 5/6 interface the same procedure was applied using the human CXCR4 receptor (PDB ID: 3ODU [10]) and the murine μ-opioid receptor (PDB ID: 4DKL [11]) as alignment templates, respectively.

Tetramer models were built in a similar fashion. Here, two TGR5 dimers with the same dimer interface, e.g. (1/8), were aligned on another TGR5 dimer with a different interface, e.g. (4/5). With this procedure six tetramers were generated: (1/8) and (5/6) dimers with an oligomeric interface of (4/5); (1/8) and (4/5) dimers with an oligomeric interface of (5/6); (4/5) and (5/6) dimers with an oligomeric interface of (1/8). Subsequently, the interface residues of the respective dimer and tetramer models were energy minimized in Maestro [12,13] using the VSGB 2.0 solvation model [14]. Finally, either dimer and tetramer model were submitted to the OPM server [15] to compute its orientation in a membrane.

Explicit linker simulations: Molecular dynamics simulations of GFP bound to a linker

For computing a thermodynamic ensemble (TE) of GFP positions with an explicit linker/GFP construct, initially, the structure of the TGR5 C-terminal residues 296-330, for which no experimental structural information is available, and the nine residues that connect the C-terminus to GFP (total sequence: QRCLQGLWGRASRDS PGPSIAYHPSSQSSVDLDLN YGSTGRHVS) was generated with the 'Protein building' approach in Maestro. Phi and psi angles of zero were chosen, resulting in a straight peptide conformation and, hence, a structurally

unbiased starting structure for the molecular dynamics (MD) simulations. This linker was subsequently fused to enhanced GFP (PDB ID: 4EUL [16]), and the resulting structure was capped with acetyl and *N*-methyl amide groups at the N- and C-termini, respectively, and protonated with PROPKA[17] according to pH 7.4. We assumed the thermodynamic ensemble (TE) of mCherry to be identical to that of GFP.

Then, the linker/GFP construct was neutralized by adding counter ions and solvated in an octahedral box of TIP3P water [18] with a minimal water shell of 12 Å around the solute. The Amber14 package of molecular simulation software [19,20] and the ff14SB and GAFF [21] force fields were used to perform an all-atom MD simulations. To cope with long-range interactions, the "Particle Mesh Ewald" method [22] was used, and the SHAKE algorithm [23] was applied to bonds involving hydrogen atoms. The time step for all MD simulations was 2 fs with a direct-space, non-bonded cut-off of 8 Å. The first linker residue was fixed with positional harmonic restraints with a force constant of 100 kcal mol^{-1} $Å^{-2}$ throughout the simulations to emulate that this residue would be bound to TGR5 embedded in a membrane. At the beginning, 17500 steps of steepest decent and conjugate gradient minimization were performed; during 2500, 10000, and 5000 steps positional harmonic restraints with force constants of 25 kcal mol^{-1} $Å^{-2}$, 5 kcal mol^{-1} $Å^{-2}$, and zero, respectively, were applied to the solute atoms. Thereafter, 50 ps of NVT-MD (MD simulations with a constant number of particles, volume, and temperature) were conducted to heat up the system to 100 K, followed by 300 ps of NPT-MD (MD simulations with a constant number of particles, pressure, and temperature) to adjust the density of the simulation box to a pressure of 1 atm and to heat the system to 300 K. During these steps, a harmonic potential with a force constant of 10 kcal mol^{-1} $Å^{-2}$ was applied to the solute atoms. As the final step in thermalization, 300 ps of NVT-MD simulations were performed while gradually reducing the restraint forces on the solute atoms to zero within the first 100 ps of this step. Afterwards, six

independent production runs of NVT-MD simulations with 150 ns length each were performed. For this, the starting temperatures of the simulations at the beginning of the thermalization were varied by a fraction of a Kelvin. The conformations obtained in these simulations were pooled for further analyses.

Implicit linker simulations

Inter-dye distance distributions for all TGR5 dimer and tetramer models were calculated using an modified Accessible Volume (AV) approach [24]. Firstly, the different protein models (see 5.14) were embedded in an explicit membrane via the CHARMM-GUI membrane builder [25]. Here, a membrane with 5500 lipids of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) per layer was created employing default settings of the CHARMM-GUI. This resulted in a membrane bilayer of about 1.5 million atoms and a side length of about 620 Å to prevent the linker/GFP construct (which has an extended length of about ~229 Å) from wrapping around the edge of the membrane. As neither ions nor water were needed for AV calculations, the steps of ion and water addition were omitted during the creation of the membrane.

For the AV simulations the fluorescent probe was attached to the C-terminal amino acid of the TGR5 via a flexible linker of 203.5 Å corresponding to 55 amino acids (36 amino acids of the TGR5 C-terminus, a 6 amino acid cloning linker, and the first 13 amino acids of the GFP's (PDB ID: 4EUL) N-terminus, **see Supporting Table 2**) with a length of 3.7 Å each [26]. A dye radius of 25 Å was used as an approximation for the GFP size, resulting in a total length of ~229 Å for the linker/GFP construct. The distance between linker attachment points in most of the screened oligomer models was shorter than the effective size of the AVs resulting in AV overlap. The AVs were constructed considering geometric factors in terms of steric exclusion effects caused by the TGR5 oligomer and the membrane. To account for clashes between the dyes, which are

not addressed in the AV simulations, the inter-dye distance probability was set to zero for all distances below 25 Å. To account also for entropic effects, we introduced position weights for the implicitly modelled linker according to the Gaussian chain model, so that the non-uniform dye position probability distribution in the AV was scaled (**Supporting Fig. 7**) [27]. In the Gaussian chain model a segment length of 7.4 Å was used, as obtained from the calibration aimed to reproduce the accurate end-to-end distance probability distribution from coarse-grained Monte-Carlo simulations of the peptide linker, similar to previously published results for the flexibly linked GFP dimer [28]. The obtained AV positional distributions were used to determine the inter-probe distance distribution by measuring all distances from positions in one AV distribution with respect to all positions in the second AV distribution. Considering oligomerization (tetramer) where two acceptors may be present in the vicinity of one donor, we computed the apparent distance distribution shifts towards shorter distances by convolution of the two inter-probe distance distributions ((1/8) and (4/5)) (**Supporting Fig. 5**).

Glossary

F_D/F_A	fluorescence intensity ratio
B	background
α	crosstalk factor
g	detection efficiencies
DE	direct acceptor excitation
DE_{rel}	relative concentration dependent brightness
Φ	distinct fluorescence quantum yields
γ	spectral shift factor
g_G^*	corrected green detection efficiency
$\varepsilon(t)$	pixel-integrated, time-resolved FRET analysis
DA	donor acceptor FRET pair
$f(t)$	fluorescence signal decay
D(0)	unquenched donor
D(A)	Donor quenched by acceptor
IRF	instrument response function
x xD xA	species fraction donor (Donly) or Non-FRET fraction acceptor fraction
AV	Accessible Volume
$\langle R_{DA} \rangle_{app}$	Apparent mean distance between Donor and Acceptor
$\langle \tau_{D(0)} \rangle_f$	Fluorescence-averaged unquenched donor fluorescence lifetime
$S_{em,ex}$ $S_{G,G}$ $S_{Y,G}$	$\text{Signal}_{\text{emission, excitation}}$ Signal of green photons emitted after excitation of GFP Signal of mCherry photons emitted after excitation of GFP
c $[D]_0, [A]_0$	Concentration Real donor and acceptor concentration

E_{mean}	Mean Transfer efficiency
wt	wildtype
MFIS-FRET	Multiparameter Fluorescence Imaging Spectroscopy-Förster Resonance Energy Transfer

Supporting References

1 DeLano, W. L. & Lam, J. W. PyMOL: A communications tool for computational models. *Abstr Pap Am Chem S* **230**, U1371-U1372 (2005).

2 Schrödinger, L. The PyMOL molecular graphics system, version 1.3 r1. Py-MOL, The PyMOL Molecular Graphics System, Version 2010,. (2010).

3 King, C., Sarabipour, S., Byrne, P., Leahy, D. J. & Hristova, K. The FRET Signatures of Noninteracting Proteins in Membranes: Simulations and Experiments. *Biophys J* **106**, 1309-1317, doi:10.1016/j.bpj.2014.01.039 (2014).

4 Spomer, L. *et al.* A Membrane-proximal, C-terminal alpha-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-coupled Receptor (GPCR) TGR5. *J Biol Chem* **289**, 3689-3702, doi:10.1074/jbc.M113.502344 (2014).

5 Keitel, V. *et al.* The Membrane-Bound Bile Acid Receptor TGR5 Is Localized in the Epithelium of Human Gallbladders. *Hepatology* **50**, 861-870, doi:10.1002/hep.23032 (2009).

6 Hov, J. R. *et al.* Mutational Characterization of the Bile Acid Receptor TGR5 in Primary Sclerosing Cholangitis. *Plos One* **5**, doi:e1240310.1371/journal.pone.0012403 (2010).

7 Weidtkamp-Peters, S. *et al.* Multiparameter fluorescence image spectroscopy to study molecular interactions. *Photochem Photobiol Sci* **8**, 470-480, doi:10.1039/b903245m (2009).

8 Gertzen, C. G. W. *et al.* Mutational mapping of the transmembrane binding site of the G-Protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur J Med Chem* doi:10.1016/j.ejmech.2015.09.024 (2015).

9 Wu, H. *et al.* Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **485**, 327-U369, doi:10.1038/nature10939 (2012).

10 Wu, B. *et al.* Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **330**, 1066-1071, doi:10.1126/science.1194396 (2010).

11 Manglik, A. *et al.* Crystal structure of the mu-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321-U170, doi:10.1038/nature10954 (2012).

12 Maestro (Schrödinger, LLC., New York, NY, 2014).

13 Version 9.9.013 (Schrödinger LLC, New York, NY, 2014).

14 Li, J. *et al.* The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins-Structure Function and Bioinformatics* **79**, 2794-2812, doi:10.1002/prot.23106 (2011).

15 Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. OPM: Orientations of proteins in membranes database. *Bioinformatics* **22**, 623-625, doi:10.1093/bioinformatics/btk023 (2006).

16 Arpino, J. A. J., Rizkallah, P. J. & Jones, D. D. Crystal Structure of Enhanced Green Fluorescent Protein to 1.35 angstrom Resolution Reveals Alternative Conformations for Glu222. *Plos One* **7**, doi:e4713210.1371/journal.pone.0047132 (2012).

17 Bas, D. C., Rogers, D. M. & Jensen, J. H. Very fast prediction and rationalization of pK(a) values for protein-ligand complexes. *Proteins-Structure Function and Bioinformatics* **73**, 765-783, doi:10.1002/prot.22102 (2008).

18 Jorgensen W, C. J., Madura JD, Impey R and Klein ML. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, doi:10.1063/1.445869 (1983).

19 Case, D. A. *et al.* The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668-1688, doi:10.1002/jcc.20290 (2005).

20 D.A. Case, V. B., J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F.

Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman AMBER 14. (2014).

21 Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174, doi:10.1002/jcc.20035 (2004).

22 Darden, T., York, D. & Pedersen, L. PARTICLE MESH EWALD - AN N.LOG(N) METHOD FOR EWALD SUMS IN LARGE SYSTEMS. *J Chem Phys* **98**, 10089-10092, doi:10.1063/1.464397 (1993).

23 Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. NUMERICAL-INTEGRATION OF CARTESIAN EQUATIONS OF MOTION OF A SYSTEM WITH CONSTRAINTS - MOLECULAR-DYNAMICS OF N-ALKANES. *J Comput Phys* **23**, 327-341, doi:10.1016/0021-9991(77)90098-5 (1977).

24 Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Methods* **9**, 1218-1225, doi:10.1038/nmeth.2222 (2012).

25 Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* **97**, 50-58, doi:10.1016/j.bpj.2009.04.013 (2009).

26 Evers, T. H., van Dongen, E. M., Faesen, A. C., Meijer, E. W. & Merkx, M. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry* **45**, 13183-13192, doi:10.1021/bi061288t (2006).

27 Chiang, J., Li, I., Pham, E. & Truong, K. FPMOD: a modeling tool for sampling the conformational space of fusion proteins. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* **1**, 4111-4114, doi:10.1109/IEMBS.2006.259224 (2006).

28 Evers, T. H., van Dongen, E. M., Faesen, A. C., Meijer, E. W. & Merkx, M. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry* **45**, doi:10.1021/bi061288t (2006).

Publication III

Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists

Christoph G.W. Gertzen[§], Lina Spomer[§], Sander H.J. Smits, Dieter Häussinger, Verena Keitel, Holger Gohlke

[§] Both authors contributed equally to this work.

Eur J Med Chem (2015), 104, 57-72.

European Journal of Medicinal Chemistry 104 (2015) 57–72

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Research paper

Mutational mapping of the transmembrane binding site of the G-protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists

Christoph G.W. Gertzen [a, 1], Lina Spomer [b, 1], Sander H.J. Smits [c], Dieter Häussinger [b], Verena Keitel [b, **], Holger Gohlke [a, *]

[a] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[b] Clinic for Gastroenterology, Hepatology, and Infectious Diseases, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[c] Institute for Biochemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany

ARTICLE INFO

Article history:
Received 12 April 2015
Received in revised form
6 September 2015
Accepted 15 September 2015
Available online 24 September 2015

Keywords:
7-Helix receptor
Cyclic AMP
Quantitative structure activity relationship
Structure model
Molecular dynamics simulations
Molecular docking

ABSTRACT

TGR5 (Gpbar-1, M-Bar) is a class A G-protein coupled bile acid-sensing receptor predominately expressed in brain, liver and gastrointestinal tract, and a promising drug target for the treatment of metabolic disorders. Due to the lack of a crystal structure of TGR5, the development of TGR5 agonists has been guided by ligand-based approaches so far. Three binding mode models of bile acid derivatives have been presented recently. However, they differ from one another in terms of overall orientation or with respect to the location and interactions of the cholane scaffold, or cannot explain all results from mutagenesis experiments. Here, we present an extended binding mode model based on an iterative and integrated computational and biological approach. An alignment of 68 TGR5 agonists based on this binding mode leads to a significant and good structure-based 3D QSAR model, which constitutes the most comprehensive structure-based 3D-QSAR study of TGR5 agonists undertaken so far and suggests that the binding mode model is a close representation of the "true" binding mode. The binding mode model is further substantiated in that effects predicted for eight mutations in the binding site agree with experimental analyses on the impact of these TGR5 variants on receptor activity. In the binding mode, the hydrophobic cholane scaffold of taurolithocholate orients towards the interior of the orthosteric binding site such that rings A and B are in contact with TM5 and TM6, the taurine side chain orients towards the extracellular opening of the binding site and forms a salt bridge with $R79^{EL1}$, and the 3-hydroxyl group forms hydrogen bonds with $E169^{5.44}$ and $Y240^{6.51}$. The binding mode thus differs in important aspects from the ones recently presented. These results are highly relevant for the development of novel, more potent agonists of TGR5 and should be a valuable starting point for the development of TGR5 antagonists, which could show antiproliferative effects in tumor cells.

1. Introduction

TGR5 (Gpbar-1, M-Bar) is a class A G-protein coupled receptor (GPCR) signaling via a stimulatory G protein and is activated by both unconjugated and conjugated bile acids and various steroid hormones including neurosteroids [1–3]. TGR5 is widely expressed in humans and rodents; organs with high amounts of TGR5 mRNA expression include the brain, the liver, and the gastrointestinal tract [1,2,4,5]. In liver, TGR5 modulates hepatic microcirculation, exerts anti-inflammatory, anti-apoptotic and choleretic effects, and promotes gallbladder filling [6–9]; in the intestine, TGR5 activation in L-cells has been linked to increased secretion of the insulin response-modulating glucagon-like peptide-1. Administration of TGR5 agonists reduced liver inflammation and steatosis and improved glucose tolerance in animal models [10]; furthermore, a reduction of atherosclerotic plaque formation was observed [11]. This makes TGR5 a promising drug target for the treatment of metabolic disorders, such as non-alcoholic steatohepatitis, type II

* Corresponding author.
** Corresponding author.
E-mail addresses: verena.keitel@med.uni-duesseldorf.de (V. Keitel), gohlke@uni-duesseldorf.de (H. Gohlke).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.ejmech.2015.09.024
0223-5234/

diabetes, obesity, and atherosclerosis [11–13]. Accordingly, much effort has been devoted to the development of potent and selective agonists of TGR5 [14–18]. Due to the lack of a crystal structure of TGR5, the development has been guided by ligand-based approaches [3,14–24] so far.

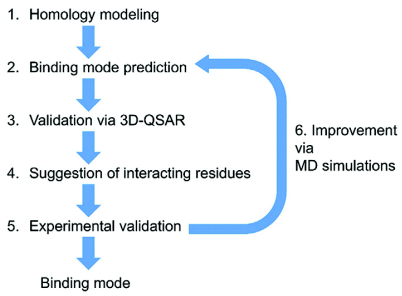
Only very recently, an integrated computational, biological, and chemical approach was presented by Macchiarulo et al. with the aim to probe the transmembrane binding site of TGR5 by mutational analysis and to predict the binding mode of agonistic bile acids and derivatives [19]. The computational part was based on a homology model of human TGR5 derived from a template structure of rhodopsin in the inactive state, with refinement of some of the binding site residues by energy minimization [19]. This resulted in the identification of "binding mode 3" [19] compliant to most of the mutagenesis data. In "binding mode 3" [19], bile acids are oriented in a head-to-tail fashion with respect to transmembrane helix (TM) 3, with the 3-hydroxyl group being involved in hydrogen-bonding interactions with N93 (position 3.33, Ballesteros–Weinstein nomenclature [25] according to the GPCR database [26], hereafter abbreviated as $N93^{3.33}$) and $Y89^{3.29}$. However, this binding mode does not explain why $E169^{5.44}$ (note that in the work of Macchiarulo et al. this residue is referred to as $E169^{5.53}$) [19], implicated to be a key residue from the degree of conservation in a TGR5 sequence alignment [19], led to a reduced TGR5 activation upon Glu169Ala mutation. Agonists in "binding mode 3" are more than 12 Å away from $E169^{5.44}$ so that it is difficult to envisage how an agonist would sense this mutation. Comparison of an active state of an agonist-bound β_2 -adrenergic receptor (β_2AR) with an inactive, antagonist-bound β_2AR state may provide an explanation for this: It revealed an inward bulge of TM5 centered around position 5.46 as the greatest structural difference in the binding pocket of the active state with a position shift of the C_α atom by 2.1 Å [27]. In addition, smaller inward movements of TM6 and TM7 were observed [27]. Such movements may influence the success of docking to a rigid TGR5 model as used by Macchiarulo et al. for agonist placement [28–30]. Furthermore, a homology model of TGR5 generated by us (see below) showed that $N93^{3.33}$ favors a conformation pointing away from the binding site such that hydrogen bond formation with the 3-hydroxyl group of bile acid derivatives, as postulated by Macchiarulo et al. [19], appears less likely. Finally, "binding mode 3" does not interact with TM6 [19]. However, interactions between agonists and TM6 are considered essential for GPCR activation [31–35]. Another two binding mode models of bile acid derivatives in a structural model of TGR5 have been presented by D'Amore et al. [36] and Yu et al. [37], both being based on combinations of molecular docking and molecular dynamics (MD) simulations applying *a priori* restraints to guide the ligand placement. Both binding modes differ from the one proposed by Macchiarulo et al., either with respect to the overall orientation of the bile acid derivative [37] or with respect to the location and interactions of the cholane scaffold [36]. In contrast to the study by Macchiarulo et al. [19] and the present study, no mutagenesis studies were performed by D'Amore et al. [36] and Yu et al. [37] to confirm the proposed binding modes.

These circumstances prompted us to predict a binding mode for natural and synthetic bile acids and neurosteroids starting from a structural model of TGR5 generated from multiple GPCR template structures and to perform mutational mapping of the transmembrane binding site of TGR5 to validate these predictions. In this process, we relaxed the TGR5 model in the presence of an agonist by all-atom MD simulations in an explicit membrane environment. Finally, we derived a protein-based 3D-QSAR model for 68 TGR5 agonists, including both bile acids and neurosteroids, with good predictive power based on the binding mode. With respect to the studies of Macchiarulo et al. [19], D'Amore et al. [36], and Yu et al. [37] our binding mode model differs in one or more of the following five aspects: I) The ligands in our binding mode are oriented parallel to the membrane, rather than perpendicular to it as in the binding mode of Yu et al. [37]; II) the cholane scaffold of bile acids binds in the vicinity of TM5 and TM6 and is rotated by 180° around the long axis compared to Macchiarulo's "binding mode 3" [19]; III) the 3-hydroxyl group of bile acids interacts with the conserved $E169^{5.44}$ [19], but neither with $N93^{3.33}$ nor with $W237^{6.48}$, which is in contrast to "binding mode 3" and the binding mode by Yu et al. [37]; IV) through this, our binding mode provides an explanation for the observed selectivity towards epimers for bile acids with a 7-hydroxyl group; V) the side chains of bile acids orient towards EL1 and, hence, are distant from $S270^{7.43}$, in contrast to "binding mode 3" [19] and the binding mode of D'Amore et al. [36].

2. Results

We pursued an iterative and integrated computational and biological approach to elucidate the binding mode of agonistic TGR5 bile acids and neurosteroids (Scheme 1); similar approaches have been successfully applied to other GPCRs [38–41]. After generating multiple structural models of TGR5 by homology modeling (step 1), initial binding modes of these models were predicted by molecular docking (step 2). The binding modes were evaluated using the predictive power of structure-based 3D-QSAR analyses as a quality criterion (step 3). Based on the best binding mode, potentially interacting residues were predicted (step 4). For experimental validation, variants of TGR5 with single-point mutations of these residues were generated, and the influence of the mutations was investigated with respect to plasma membrane localization and function using immunofluorescence staining, flow cytometry, and a cAMP responsive luciferase assay (step 5). To further improve the binding mode, the TGR5/agonist complex was relaxed by MD simulations (step 6), whereupon steps 2 to 5 were repeated to reach the final binding mode. These steps will be described in detail in the following.

Step 1 — Homology modeling of TGR5. In order to generate a structural model of TGR5, we applied a multi-template homology modeling approach. All antagonist-bound class A GPCR crystal structures with a resolution <3 Å available at the beginning of this study and identified by a PSI-BLAST [42] search with the TGR5 sequence as a query served as templates. If more than one structure matched these criteria for a GPCR, the structure with the best resolution was chosen. This resulted in seven templates, the turkey β_1-adrenergic receptor (PDB code 2VT4), the human β_2-adrenergic receptor (3D4S), the human adenosine-A_{2A} receptor (3EML), the human CXCR4 receptor (3ODU), the human dopamine-D_3 receptor (3PBL), the human muscarinic-M_2 receptor (3UON), and the human



Scheme 1. Integrated computational and biological workflow for the prediction of a binding mode of TGR5 agonists.

sphingosine 1-phosphate receptor 1 (3V2Y) with overall sequence identities with respect to the human TGR5 sequence ranging from 11% to 18%. The sequence identities and similarities of the transmembrane regions are ~20% and ~50%, respectively.

We considered two alternatives for TM7 of TGR5 when generating the multiple sequence alignment with the templates: In the first alternative, priority is given to the alignment of the conserved (D/E)X(K/R) motif (positions 8.48 to 8.50) (see Fig. S1 in the Supplemental Material (SM)); in the second alternative, conserved Pro residues in TM7 are preferentially aligned (Fig. S2 in the SM). Preliminary models of TGR5 generated by Modeller [43] from either alternative revealed that TM7 has the same length in both cases but that residues at position 7.n in the first alternative are located at position 7.$(n + 2)$ in the second one. As GPCRs possess multiple conserved Pro residues in their TMs [44], the second alternative should be more reliable than the first one. Still, favoring the second alternative based on this one criterion only seemed insufficient. Thus, we generated 50 homology models (i.e., 10 base models for which 5 loop models each were generated) for each alignment to approximately account for conformational variability of the binding site [45] and/or reduce the influence of potential modeling errors on the outcome of subsequent steps [46]. All seven templates were used for generating the homology models in a multi-template approach. Assessment of the stereochemical quality of the models by PROCHECK [47] showed that at least 85.5% of the residues are in the core region (Tables S1 and S2 in the SM). This number compares favorably to the stereochemical quality of crystal structures with a resolution <2.4 Å [48]. Using our homology modeling software Topmodel [49] to evaluate the models, normalized DOPE scores [43] also indicated a good structural quality of the models, with the second alignment alternative resulting in more favorable scores (mean score: 0.73) than the first alignment alternative (mean score 0.93). This may reflect more unfavorable intramolecular interactions caused by the altered positioning of side chains in TM7 as explained above. Ten models with the largest variations in the conformations of the transmembrane binding site were selected by visual inspection from each alignment. Their binding sites differ by a C_α atom RMSD of 0.5 ± 0.2 Å (mean ± standard error of the mean). Thereby, we ensured that the conformational space of the binding site region is spanned similarly as by the full ensembles of 50 models; yet, the number of models to deal with in the subsequent steps 2 and 3 is kept at a manageable value.

Step 2 – Prediction of an initial binding mode for TLC. In each homology model, taurolithocholate (TLC), the most active natural TGR5 agonist currently known [3], was docked using AutoDock/DrugScore. This combination of docking engine and scoring function has proven reliable in a "re-docking" evaluation [50] and when docking ligands to difficult protein binding epitopes [51]. A converged docking solution with at least 20% of all TLC configurations in the biggest cluster was found in each case (Table S3 in the SM). Next, of the TLC configurations with the lowest energy in the biggest clusters those were selected by visual inspection that "approached" E169$^{5.44}$ in TM5 and/or a residue of TM6. E169$^{5.44}$ has been implicated to be a key residue from the degree of conservation in a TGR5 sequence alignment [19], and interactions to TM6 are essential for GPCR activation [31–35]. As we aimed at the prediction of an initial binding mode in this step, for which a subsequent refinement was planned, "approached" was defined rather loosely as a distance between any pair of atoms of TLC and E169$^{5.44}$ or TM6 <8 Å. If several similar TLC configurations (i.e., with an all-atom RMSD <5 Å) met these criteria, the configuration with the lowest energy was chosen. This resulted in the selection of the docking solution from model 42 for the first ((D/E)X(K/R) motif-centered) alignment alternative (Fig. S3A in the SM) and the docking solution from model 19 for the second (proline-centered) alignment alternative (Fig. 1A). Both docking solutions have in common that TLC forms a hydrogen bond with E169$^{5.44}$ using its 3-hydroxyl group and hydrophobic interactions with Y240$^{6.51}$ and L244$^{6.55}$ in TM6. Furthermore, TLC in both configurations binds with its sulfonic acid moieties in the vicinity of TMs 1 and 7 deep inside the pocket. In model 42, however, TLC binds to S21$^{1.39}$ in TM1 (Fig. S3 in the SM) while in model 19 TLC binds to S270$^{7.43}$ in TM7 (Fig. 1A). This difference results from the shift of two residues in TM7 between the two alignment alternatives, which places S270$^{7.43}$ at the interface between TMs 6 and 7 in models of the first alignment; thus, S270$^{7.43}$ becomes unavailable as a binding partner.

Step 3 – Evaluation by structure-based 3D-QSAR analysis. Next, we tested if the two initial binding modes allow for the establishment of a quantitative structure–activity relationship for 68 TGR5 agonists (Tables S4 and S5 in the SM). We applied the structure-based 3D-QSAR approach AFMoC [52] for this. 3D-QSAR analyses have been successfully used previously in the context of binding mode exploration [53–55]. In addition, as a negative control, AFMoC analyses were performed starting from two docked TLC configurations that did not meet the "approach" criteria from Step 2 (model 26 of the second alignment alternative, in which E169$^{5.44}$ does not point into the binding pocket (Fig. S3B in the SM), and model 34 of the second alignment alternative, in which TLC does not approach TM5 at all (Fig. S3C in the SM)). Finally, to exclude that the TLC conformation alone (i.e., without considering its position and orientation in the transmembrane binding site of TGR5) can give rise to a predictive AFMoC model, the TLC configuration in model 19 was placed in the TGR5 model 42 and *vice versa*, after superpositioning the two model structures. In each case, 68 TGR5 agonists (Tables S4 and S5 in the SM) were then structurally aligned onto TLC as described in the Experimental Section.

Of all six AFMoC analyses, only the one based on the docked TLC configuration in TGR5 model 19 of the second alignment alternative (Fig. 1B) produced a significant QSAR model [56] with $q^2 = 0.37$ for six components (Table 1). Seven runs of leave-multiple-out AFMoC analysis (LMO analysis), in each of which ten agonists from the data set were randomly left out, resulted in an average $q^2 = 0.44$ for six components, showing a robust performance of the AFMoC analyses. Recently, Kramer et al. [57] showed to what extent the goodness of a QSAR model, measured in terms of the R^2 in the Kramer et al. study, depends on the uncertainty in the dependent variables (in our case, the pEC_{50} values) and the range (standard deviation) of these variables. In particular, Kramer et al. derived an equation to calculate the highest achievable R^2 ($R^2_{max.}$) given these two parameters. The pEC_{50} values of our dataset have a standard deviation of 0.81 (Tables S4 and S5 in the SM). Determining the uncertainty in the experimental pEC_{50} values is not straightforward, however, as to the best of our knowledge no standard deviations or standard errors in the mean have been reported by Sato et al. [2] from which our dataset was taken. Estimating standard deviations of pEC_{50} from confidence intervals reported by Pelliciari et al. [3] yields a value of 0.20. According to the equation by Kramer et al., this yields $R^2_{max.} = 0.94$. We consider this a favorable estimate as only experimental uncertainties of five compounds overlapping with our dataset have been reported by Pelliciari et al. Likely, a more realistic estimate is obtained if the experimental error determined by Kramer et al. over 7667 entries in the ChEMBL database is used (0.54). The equation by Kramer et al. then yields $R^2_{max.} = 0.56$, which compares favorably to $R^2 = 0.81$ obtained for our model (Table 1). Random scrambling of the biological data and recalculation of the QSAR model resulted in $q^2 = -0.01$, indicating that, with the original biological data, no chance correlation is given. In contrast, all other combinations of TLC configurations in TGR5 models resulted in QSAR models with $q^2 < 0.03$ (Table S6 in the SM). In all, these results reveal a strong sensitivity of the predictive

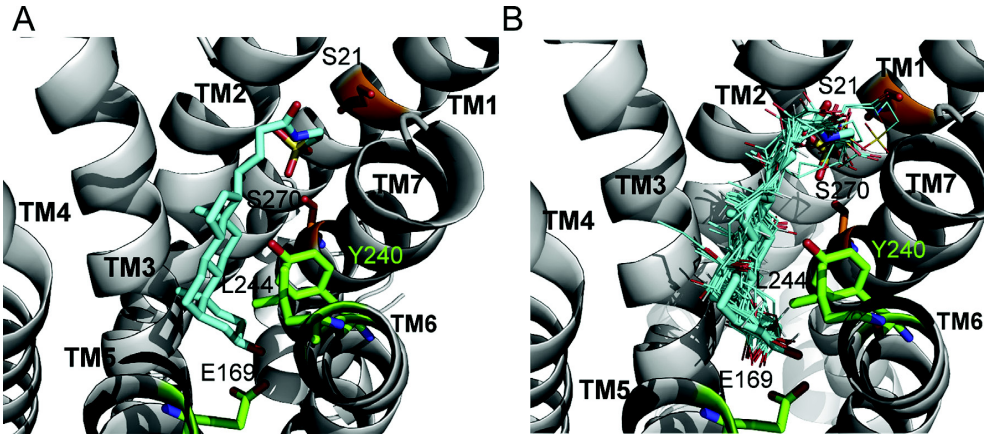


Fig. 1. Binding mode of TLC predicted by molecular docking into the initial homology model of TGR5 (**A**) and corresponding alignment of TGR5 agonists used in the initial AFMoC analysis (**B**). TLC is shown with cyan sticks, all other TGR5 agonists with cyan lines, and TGR5 in gray cartoon representation. Amino acids subjected to mutational analysis are shown in sticks representation; they are colored according to having a negative effect (green) or no effect (orange) on receptor activity upon stimulation with TLC when mutated to alanine. In the docked binding mode (panel A), the 3-hydroxyl group of TLC forms a hydrogen bond with E169$^{5.44}$ and the sulfonic acid moiety of TLC forms a hydrogen bond with S270$^{7.43}$, whereas TLC does not interact with S21$^{1.39}$. Furthermore, TLC makes hydrophobic interactions with Y240$^{6.51}$ and L244$^{6.55}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Results of the AFMoC analyses of the initial and the final binding mode models.

	Initial binding mode	Final binding mode
Spacing[a]	1.0	1.0
σ[b]	0.55	0.85
q^2[c,d]	0.37 (0.37)	0.50 (0.50)
s_{press}[e,f]	0.37	0.53
r^2[c,g]	0.81 (0.81)	0.60 (0.60)
s[e,h]	0.35	0.52
F[c,i]	44.36 (44.36)	95.31 (95.31)
Components[j]	6	1
Fractions[k,l]		
C.3	0.57	0.79
C.2	0.16	0.02
C.ar	0.01	0.01
O.3	0.16	0.15
O.2	0.01	0.00
O.co2	0.08	0.02
N.am	0.03	0.00
S.3	0.01	0.00
F	0.00	0.00

[a] In Å.
[b] Half-width of the convolution of the receptor potential fields with the Gaussian function representing the ligand atoms.
[c] Values are given considering only the part of pEC_{50} used in the PLS analysis (pEC_{50}^{PLS}) or considering the total pEC_{50} (values in parentheses).
[d] $q^2 = 1 - PRESS/SSD$ as obtained by "leave-one-out" cross-validation. PRESS equals the sum of squared differences between predicted and experimentally determined binding affinities, SSD is the sum of the squared differences between experimentally determined binding affinities and the mean of the training set binding affinities.
[e] In logarithmic units.
[f] $s_{press} = \sqrt{(PRESS/(n-h-1))}$ as obtained by "leave-one-out" cross-validation. n equals the number of data points, h is the number of components.
[g] Correlation coefficient.
[h] $s = \sqrt{(RSS/(n-h-1))}$. RSS equals the sum of squared differences between fitted and experimentally determined binding affinities.
[i] Fisher's F-value.
[j] Interaction energy components correlated to the binding affinities via PLS analysis.
[k] Fraction of the importance of the positions of the respective atom type.
[l] Sybyl atom types: C.3: sp^3-hybridized carbon; C.2: sp^2-hybridized carbon; C.ar: aromatic carbon; O.3: sp^3-hybridized oxygen; O.2: sp^2-hybridized oxygen; O.co2: oxygen on carboxylate or phosphate groups; N.am: nitrogen amide; S.3: sp^3-hybridized sulfur; F: fluorine.

power of an AFMoC model with respect to the TLC binding mode and suggest the TLC binding mode found in TGR5 model 19 to be the most plausible one. In turn, these results disfavor TGR5 model 42, which is based on the first alignment alternative.

Finally, we tested the influence of using either the α- or β-Me-LCA diastereomer in the alignments for all TLC configurations in TGR5 models. This was done as the configuration of the methyl group in position seven of ξ-Me-LCA is not known [3]. In all cases, α-Me-LCA resulted in QSAR models with $q^2 < 0$ (data not shown). This is in line with findings that synthetic derivatives of bile acids in our data set that possess alkyl chains in the 7β-position have a higher activity than their natural counterparts [3]. Because of this, further analyses only considered the 7β-diastereomer of ξ-Me-LCA.

Step 4 – Prediction of potentially interacting residues. In order to consolidate the docked TLC configuration in TGR5 model 19 (Fig. 1A), residues interacting with TLC were identified by visual inspection for subsequent mutation analysis. First, the E169A$^{5.44}$ mutation was suggested for two reasons: (I) Agonist-receptor interactions in GPCR crystal structures often feature hydrogen bonds with amino acids in the same or nearby transmembrane positions [58–60]; (II) the 3-hydroxyl group of TLC also forms a hydrogen bond with E169 in this binding mode. Second, the Y240A$^{6.51}$ and L244A$^{6.55}$ mutations were suggested because they form hydrophobic contacts with TLC's cholane scaffold (Fig. 1A) and are located in TM6, which is considered important for receptor activation [31–35]. Finally, the S21A$^{1.39}$ mutation was suggested as a negative control in order to exclude the binding mode found in model 42, in which TLC would form a hydrogen bond to S21$^{1.39}$ with its sulfonic acid moiety.

Step 5 – Influence of mutations on plasma membrane localization and function of TGR5. The impact of the suggested TGR5 variants S21A$^{1.39}$, E169A$^{5.44}$, Y240A$^{6.51}$, and L244A$^{6.55}$ on receptor activity and subcellular localization were analyzed after introduction of the respective mutations into the TGR5 cDNA plasmids and transient transfection into HEK293 cells [61,62]. Responsiveness of the mutated TGR5 proteins towards TLC was analyzed using a cAMP responsive luciferase assay as described [61,62]. Thus, luciferase activity served as a measure for the rise in intracellular cAMP concentration following either activation of

TGR5 by TLC or stimulation of the cells with forskolin, which elevates cAMP independent of TGR5 and was used as positive control [61,62]. As expected, the TLC-dependent luciferase activity of S21A$^{1.39}$ was comparable to that of wildtype TGR5 (Fig. 2A). In contrast, E169A$^{5.44}$ showed a significantly reduced activity at TLC concentrations between 0.1 and 2.5 μM, while activation by 10 μM

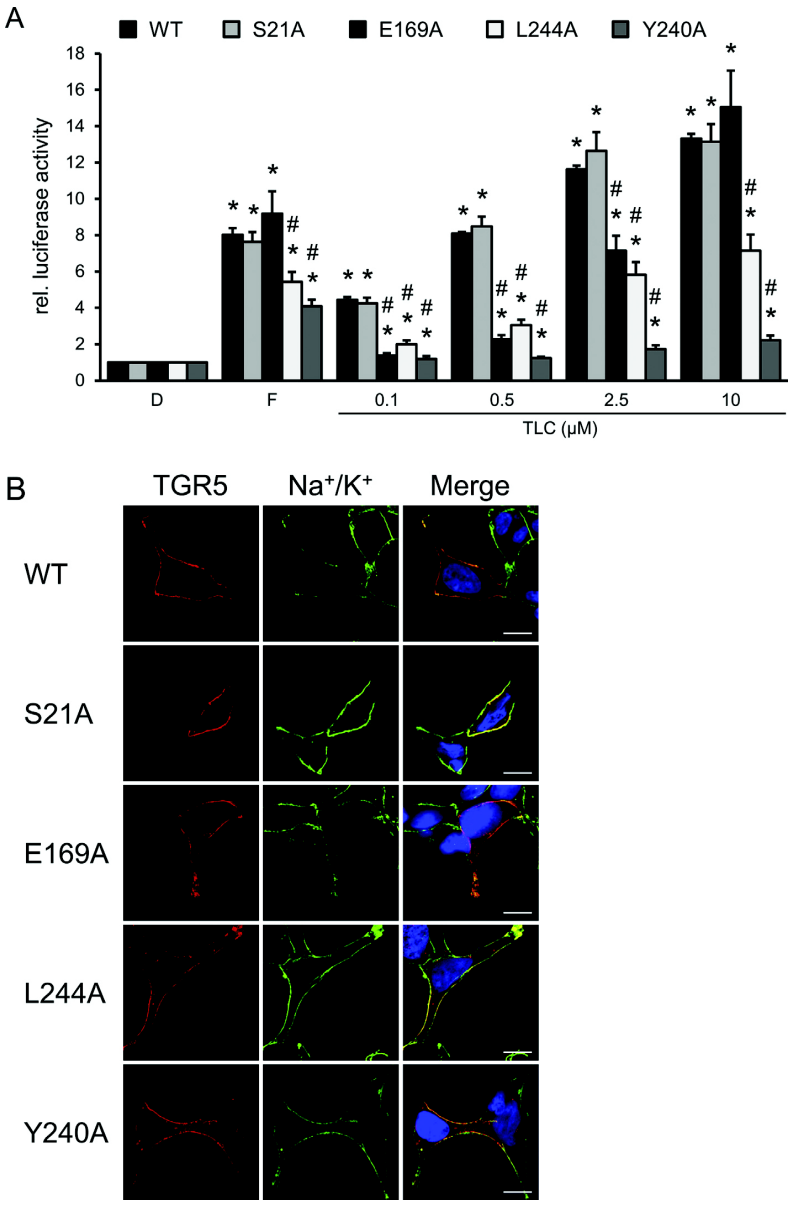


Fig. 2. Analysis of TGR5 variants in transfected HEK293 cells. HEK293 cells were transiently transfected with TGR5 wildtype (WT) and the S21A$^{1.39}$, E169A$^{5.44}$, L244A$^{6.55}$, and Y240A$^{6.51}$ variants in the pcDNA3.1+ vector. **A:** Receptor activity towards taurolithocholate (TLC) was measured using a cAMP responsive luciferase construct, and luciferase activity served as a measure of the rise in intracellular cAMP following activation of TGR5. Forskolin (F, 10 μM) was used as TGR5 independent positive control. Dimethylsulfoxide (D) was used as a negative control. The variant S21A$^{1.39}$ did not affect receptor responsiveness. E169A$^{5.44}$ and L244A$^{6.55}$ showed reduced luciferase activity at lower TLC concentrations (0.1–2.5 μM), while retaining activity similar to the WT at 10 μM TLC. The variant Y240A$^{6.51}$ almost abolished TLC-dependent luciferase activity at all concentrations tested and also significantly reduced forskolin-mediated rise in cAMP. Results (WT n = 21; S21A$^{1.39}$, E169A$^{5.44}$ n = 8; L244A$^{6.55}$, Y240A$^{6.51}$ n = 7) are expressed as mean ± SEM. *, # = significantly different ($p \leq 0.01$) from DMSO and TGR5 WT, respectively. **B:** Receptor localization was investigated by confocal laser scanning microscopy. The TGR5 protein was made visible using an anti-TGR5 antibody (RVLR2, red). WT TGR5 was almost completely targeted to the plasma membrane (PM) as demonstrated by the yellow coloring in the overlay image with the PM marker protein sodium/potassium (Na^+/K^+) ATPase (green). The variants S21A$^{1.39}$, E169A$^{5.44}$, L244A$^{6.55}$, and Y240A$^{6.51}$ were also predominantly detected in the PM, where the mutant TGR5 proteins colocalized with the PM marker protein. Nuclei were stained with Hoechst (blue). Bars = 10 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

TLC and forskolin were increased to comparable levels as found for wildtype TGR5 (Fig. 2A). Activation of L244A$^{6.55}$ by TLC led to a dose-dependent increase in luciferase activity; however, the rise in cAMP for each TLC concentration was significantly reduced as compared to wildtype (WT) TGR5 (Fig. 2A). In contrast to the E169A$^{5.44}$ variant, the L244A$^{6.55}$ variant is not fully rescuable at the highest TLC concentration tested (10 μM). This fact may reflect that the former mutation predominantly impacts the function of TGR5 whereas the latter predominantly influences TLC binding. Furthermore, the forskolin induced luciferase activity was also significantly lower than for the wildtype. The Y240A$^{6.51}$ variant showed only a slight increase in luciferase activity in response to TLC (by a factor of 2.2 ± 0.3 at 10 μM TLC) and also reduced forskolin-dependent luciferase activity (Fig. 2A).

We have previously shown that TGR5 responsiveness towards its ligand TLC is dependent on the amount of functionally active receptor protein in the plasma membrane [62]. Therefore, we investigated the localization of the TGR5 variants in transfected HEK293 cells by confocal laser scanning microscopy as well as by flow cytometry (FACS). As shown in Fig. 2B, S21A$^{1.39}$, E169A$^{5.44}$, L244A$^{6.55}$, and Y240A$^{6.51}$ reached the plasma membrane, where the mutant TGR5 proteins colocalized with the plasma membrane marker protein Na^+/K^+ ATPase. Since E169A$^{5.44}$, L244A$^{6.55}$, and Y240A$^{6.51}$ were also detected in intracellular vesicular structures, the amount of mutant receptor in the plasma membrane was quantified by FACS as described in the experimental section [61,62]. In transfected HEK293 cells 92.7 ± 0.8% of WT TGR5 was present in the plasma membrane (n = 8) (Table 2). In comparison, 88.7 ± 1.5% of S21A$^{1.39}$, 80.9 ± 2.6% of E169A$^{5.44}$, 92.1 ± 2.1% of L244A$^{6.55}$, and 82.0 ± 2.9% of Y240A$^{6.51}$ were detected in the plasma membrane, which was not significantly different from WT TGR5. Therefore, reduced membrane localization may not account for the impaired functional activity of E169A$^{5.44}$, L244A$^{6.55}$, and Y240A$^{6.51}$ (Table 2).

Step 6 – Relaxation of the TGR5/TLC complex structure by MD simulations. In order to relax the complex structure of the docked TLC configuration in the TGR5 model 19 and to test the structural stability of the binding mode, MD simulations of the complex structure in an explicit membrane environment of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine lipids were performed. For this, the Amber suite of molecular simulation programs [63] was applied together with the ff12SB, GAFF [64], and lipid14 [65] force fields for TLC, TGR5, and the lipids, respectively. Nine independent MD simulations of 100 ns length each were performed, of which the last 80 ns were evaluated. In all simulations, the TGR5 model structures remained stable, as indicated by RMSD values <6.2 Å (<3.4 Å) considering all C_α atoms (only C_α atoms of the TMs) (Fig. S4 in the SM); as expected the largest movements were observed in IL3. In general, also the TLC configuration remained stable with respect to the TGR5 model as indicated by RMSD values <3.2 Å. However, two local but important structural changes were observed. First, throughout the MD simulations, TLC showed a tendency to form an H-bond with its 3-hydroxyl group to the hydroxyl group of Y240$^{6.51}$, in addition to a hydrogen bond with E169$^{5.44}$ (Fig. S4 in the SM). Second, the hydrogen bond between TLC's sulfonic acid moiety and S270$^{7.43}$ broke apart after at most 12 ns in each MD simulation (Fig. S4 in the SM). Instead, the taurine tail reoriented towards the extracellular opening of the binding site and started to interact with R79^{EL1} of the EL1 (Fig. S4 in the SM).

Second round of steps 2 to 5 – Derivation of the final binding mode. The MD trajectories were clustered with respect to the RMSD of the TLC configurations after superimposing the TGR5 structures. Using an RMSD cutoff of 1 Å, nine clusters were identified. From these, three binding modes were chosen by visual inspection, in which TLC forms hydrogen bonds with two of the three amino acids R79^{EL1}, E169$^{5.44}$, and Y240$^{6.51}$, respectively. A fourth binding mode was identified in which TLC addresses all three of the aforementioned residues. AFMoC analyses on the 68 TGR5 agonists superimposed onto the respective TLC configurations of the first three binding modes yielded QSAR models that were insignificant ($q^2 < 0.20$) [56]. In contrast, the QSAR model based on the fourth binding mode was significant and good [56] ($q^2 = 0.50$ with one component; Table 1, Fig. S5 in the SM). Seven runs of LMO analysis, in each of which ten agonists were randomly left out from the data set, resulted in an average $q^2 = 0.54$ for one component, showing a robust performance of the AFMoC analyses. Random scrambling of the biological data and recalculation of the QSAR model resulted in $q^2 = -0.05$, indicating that with the original biological data no chance correlation is given. Thus, the MD simulations-derived binding mode (Fig. 3A) results in an AFMoC model that shows a markedly better predictive power than the AFMoC model derived from the initial, docked binding mode (Fig. 1), as judged from an increase of q^2 by 0.13 in connection with a reduction of the number of components from six to one (Table 1).

During the alignment of the TGR5 agonists onto the MD simulations-derived binding mode (Fig. 3B), we observed that agonists with a 7-hydroxyl group preferred a configuration shifted by about 3 Å towards helix 3 compared to TLC. For agonists such as TCDC with a 7-hydroxyl group in α-position, this occurred due to

Table 2
Potential interaction partner and type, mutation and its predicted effect, function and protein localization of TGR5 variants for the initial and final binding mode models.

Ligand	Potential interaction partner	Predicted interaction type	Mutation	Predicted effect upon mutation	Influence on function	Function[a]	Membrane localization[b]
Initial binding mode							
TLC	S21$^{1.39}$	No interaction	S21A	No effect	No loss	104 ± 8	89 ± 2
TLC	E169$^{5.44}$	Hydrogen bond	E169A	Binding	Rescuable	119 ± 16	81 ± 3
TLC	Y240$^{6.51}$	Hydrophobic	Y240A	Activation	Not rescuable	18 ± 2*	82 ± 3
TLC	L244$^{6.55}$	Hydrophobic	L244A	Activation	Not rescuable	57 ± 7*	92 ± 2
Final binding mode							
TLC	R79^{EL1}	Salt bridge	R79A	Binding	Rescuable	65 ± 7*	91 ± 1
TLC	Y89$^{3.29}$	Hydrophobic	Y89A	Binding	Rescuable	117 ± 17	94 ± 1
TCDC	Y89$^{3.29}$	Hydrogen bond	Y89A	Binding	Rescuable	92 ± 8[c]	94 ± 1
TUDC	Y89$^{3.29}$	Hydrophobic	Y89A	Binding	Rescuable	39 ± 7[d,*]	94 ± 1
TLC	N93$^{3.33}$	No interaction	N93A	No effect	Not rescuable	12 ± 1*	33 ± 2*
TLC	Y240$^{6.51}$	Hydrogen bond	Y240F	Activation	Not rescuable	17 ± 1*	86 ± 2
TLC	S270$^{7.43}$	No interaction	S270A	No effect	No loss	81 ± 8	91 ± 1

*Significantly different ($p < 0.01$) from TGR5 wildtype.
[a] Function at 10 μM TLC as percent of wildtype ± SEM.
[b] Percentage of cell membrane localization determined by FACS analysis ± SEM.
[c] Function at 50 μM TCDC as percent of wildtype ± SEM.
[d] Function at 100 μM TUDC as percent of wildtype ± SEM.

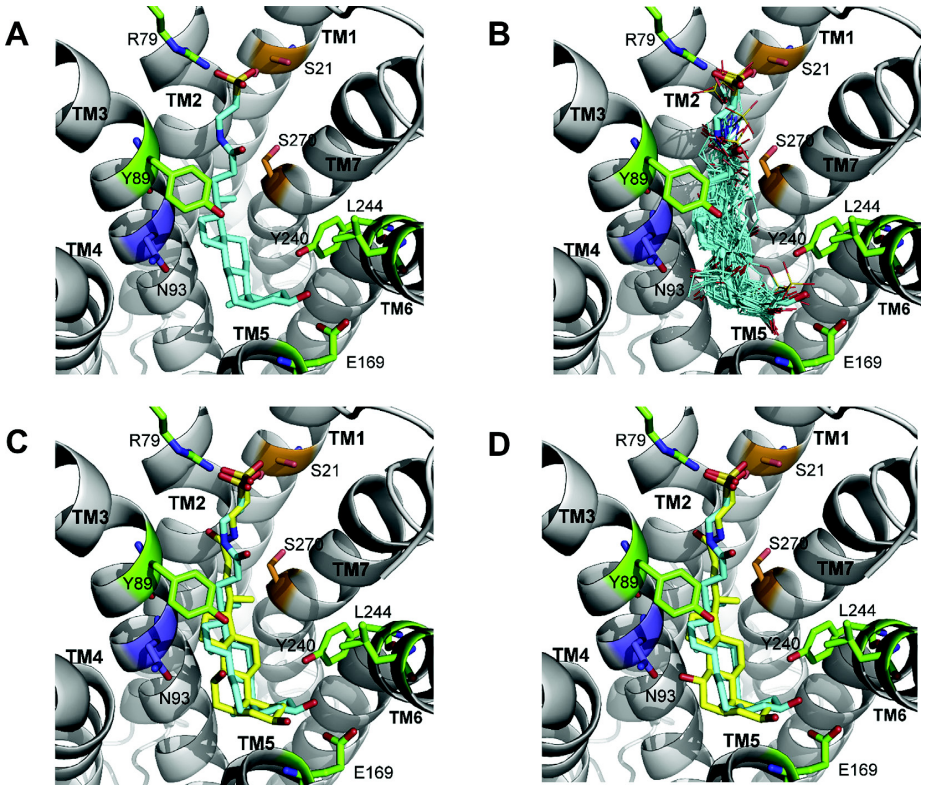


Fig. 3. Binding mode of TLC after refinement of the TLC/TGR5 complex by MD simulations (**A**), corresponding alignment of TGR5 agonists used for the second AFMoC analysis (**B**), binding mode of TCDC (**C**) and TUDC (**D**) as compared to TLC. TLC is shown with cyan sticks, all other TGR5 agonists in panel B with cyan lines, and TGR5 in gray cartoon representation. Amino acids subjected to mutational analysis are shown in sticks representation; they are colored according to having a negative effect (green) or no effect (orange) on receptor activity upon stimulation with TLC when mutated to alanine. The effect of the N93A$^{3.33}$ mutation (navy) could not be determined as the mutant was mainly retained intracellularly (see Fig. 4). In the refined binding mode (panel A), TLC forms a salt bridge to R79 in the EL1 with its sulfonic acid moiety. Furthermore, it forms hydrogen bonds with E169$^{5.44}$ and Y240$^{6.51}$ with its 3-hydroxyl group. Hydrophobic interactions are formed between TLC and Y89$^{3.29}$ and L244$^{6.55}$. Binding mode of TCDC (sticks representation, yellow) with its 7α-hydroxyl group forming a hydrogen bond to Y89$^{3.29}$, moved towards TM3 relative to TLC (**C**), and binding mode of TUDC (sticks representation, yellow) with its 7β-hydroxyl group forming a hydrogen bond to N93$^{3.33}$, moved towards TM3 relative to TLC (**D**). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

hydrogen bond formation with Y89$^{3.29}$ in TM3 (Fig. 3C). Agonists such as TUDC with a 7-hydroxyl group in β-position formed a hydrogen bond with N93$^{3.33}$ instead (Fig. 3D). In addition, TCDC and TUDC form hydrophobic interactions to Y89 with the D-ring of the cholane scaffold, as does TLC.

Consequently, mutations of the following residues should yield experimental support to the MD simulations-derived binding mode, including the shifted configurations for agonists with a 7-hydroxyl group: R79A^{EL1} and Y240F$^{6.51}$, as TLC forms hydrogen bonds with these residues; S270A$^{7.43}$ as a negative control, as TLC does not form hydrogen bonds with this residue; Y89A$^{3.29}$ and N93A$^{3.33}$, as these residues are predicted to have an influence on the activation of TGR5 by TCDC and TUDC, respectively, but not TLC.

Introduction of N93A$^{3.33}$ completely abolished TLC induced luciferase activity, while the forskolin-mediated rise in cAMP was unaffected (Fig. 4A). Immunolocalization of N93A$^{3.33}$ transfected HEK293 and CHO cells revealed that the mutated receptor was mainly retained intracellularly (Fig. 4C, see also Fig. S4 in the SM): only 33.0 ± 2.2% of N93A$^{3.33}$ reached the plasma membrane of transfected HEK293 cells as measured by FACS analysis (Table 2), thus explaining the complete loss of functional activity of this mutant (Fig. 4A). The activity of S270A$^{7.43}$ was not significantly different to that of wildtype TGR5 at high concentrations of TLC (Fig. 4A). Even though R79A^{EL1} showed a dose-dependent increase in TLC-dependent luciferase activity, it was significantly reduced compared to wildtype TGR5 (Fig. 4A). Furthermore, R79A^{EL1} also affected the forskolin-mediated rise in cAMP. Y240F$^{6.51}$ almost completely abolished TLC induced luciferase activity and significantly diminished the effect of forskolin (Fig. 4A). Y89A$^{3.29}$ reduced luciferase activity at TLC concentrations of 0.1 and 0.5 μM (Fig. 4B) and resulted in an EC_{50} value of 8.90 μM compared to 0.68 μM for TGR5 WT (Table 3) as obtained from dose/response curves (Fig. S4 in the SM). Y89A$^{3.29}$ significantly decreased TCDC-dependent cAMP production at 2.5 and 10 μM (Fig. 4B), paralleled by an increase in the EC_{50} (WT: 2.30 μM; Y89A$^{3.29}$: 29.57 μM; Table 3, Fig. S4 in the SM). Stimulation of Y89A$^{3.29}$ with TUDC resulted in a dose-dependent increase in luciferase activity; however, compared to WT TGR5, this rise in cAMP was significantly reduced for all concentrations tested (10–500 μM, n = 3–10, Fig. 4B), paralleled by an increase in the EC_{50} (WT: 50.50 μM; Y89A$^{3.29}$: 105.15 μM; Table 3, Fig. S4 in the SM). The EC_{50} values determined for TGR5 WT in our study are in agreement with previously reported values [3] (see also Table S2 in the SM) (Table 3).

Apart from N93A$^{3.33}$ (Fig. 4C, Fig. S4 in the SM), variants

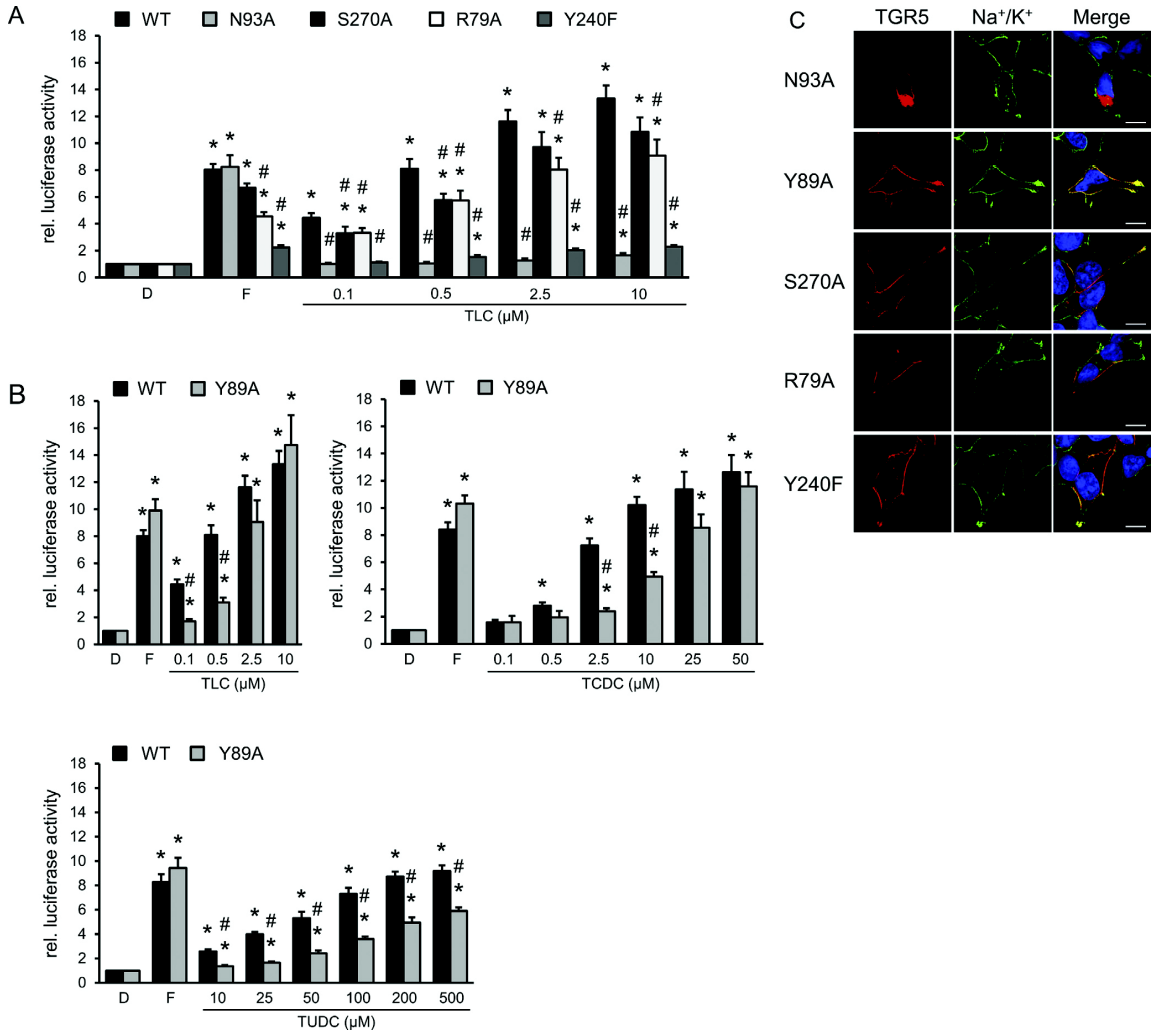


Fig. 4. Mutational analysis of residues involved in TGR5 ligand binding. HEK293 cells were transiently transfected with TGR5 WT and the TGR5 variants N93A$^{3.33}$, S270A$^{7.43}$, R79A^{EL1}, Y240F$^{6.51}$, and Y89A$^{3.29}$. **A:** Receptor responsiveness towards TLC was measured using a cAMP responsive luciferase construct. The variant N93A$^{3.33}$ completely abolished receptor activity towards TLC stimulation, however, it did not affect the forskolin-mediated rise in cAMP. The variants S270A$^{7.43}$ and R79A^{EL1} showed a dose-dependent rise in luciferase activity following incubation with TLC; however, compared to the WT the increase in cAMP was significantly lower for S270A$^{7.43}$ at lower TLC concentrations (0.1 and 0.5 μM) and for R79A^{EL1} at all TLC doses used. Y240F$^{6.51}$ completely abolished TLC-induced luciferase activity and also significantly diminished the effect of forskolin (WT n = 21; N93A$^{3.33}$, S270A$^{7.43}$ n = 8; R79A^{EL1}, Y240F$^{6.51}$ n = 9). **B:** Stimulation of TGR5 WT and the variant Y89A$^{3.29}$ with TLC, taurochenodeoxycholate (TCDC), and tauroursodeoxycholate (TUDC). Introduction of Y89A$^{3.29}$ significantly reduced luciferase activity after stimulation with 0.1 and 0.5 μM TLC (n = 8). When tested for TCDC, the mutation Y89A$^{3.29}$ significantly affected the induced luciferase activity at 2.5 and 10 μM of TCDC (n = 7–15). Incubation with TUDC (10–500 μM, n = 3–10) resulted in a significantly reduced rise in cAMP as measured by luciferase activity for the Y89A$^{3.29}$ variant as compared to WT. Results are expressed as mean ± SEM. *, # = significantly different ($p \leq 0.01$) from DMSO (D) and TGR5 WT, respectively. **C:** Receptor localization was investigated by confocal laser scanning microscopy. The TGR5 protein was made visible using an anti-TGR5 antibody (RVLR2, red). WT TGR5 was almost completely targeted to the plasma membrane (PM) as demonstrated by the yellow coloring in the overlay image with the PM marker protein sodium/potassium (Na^+/K^+) ATPase (green). The variants Y89A$^{3.29}$, S270A$^{7.43}$, R79A^{EL1}, and Y240F$^{6.51}$ reached predominantly the PM, while the mutant N93A$^{3.33}$ was mainly retained intracellularly. Nuclei were stained with Hoechst (blue). Bars = 10 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Y89A$^{3.29}$, S270A$^{7.43}$, R79A^{EL1}, and Y240F$^{6.51}$ were detected in the plasma membrane of transfected HEK293 cells by immunofluorescence staining and confocal laser scanning microscopy (Fig. 4C). FACS analysis demonstrated that 94.2 ± 0.9%, 91.0 ± 1.0%, 91.2 ± 0.5%, and 85.5 ± 1.8% of Y89A$^{3.29}$, S270A$^{7.43}$, R79A^{EL1}, and Y240F$^{6.51}$, respectively, were sorted to the plasma membrane of transfected cells (Table 2), indicating that the observed changes in functional activity were independent of receptor plasma membrane localization.

3. Discussion

TGR5 is a promising drug target for the treatment of metabolic disorders [5,6,15]. So far, the development of agonists of TGR5 has been guided by ligand-based approaches [14–18] due to the lack of a crystal structure of TGR5. Related to the first structural model of TGR5, Hov et al. [61] described two clinically relevant mutations, W83R$^{3.23}$ and V178M$^{5.52}$, that decrease the activity of TGR5 upon TLC stimulation. In the TGR5 model by Hov et al. [61] and the one

Table 3
EC_{50} values of TLC, TCDC, and TUDC with respect to TGR5 WT and the Y89A$^{3.29}$ variant.[a]

Ligand	EC_{50}[b]	
	TGR5 WT	TGR5 Y89A$^{3.29}$
TLC	0.68 ± 0.23 (0.29)	8.90 ± 0.20
TCDC	2.30 ± 0.84 (1.91)	29.57 ± 9.20
TUDC	50.50 ± 16.20 (30.20)	105.15 ± 17.10

[a] Receptor responsiveness towards the different ligands was measured using a cAMP responsive luciferase assay. The values were calculated using the Prism software. See Fig. S7 in the SM for dose/response curves.
[b] In μM. Values in parentheses for TGR5 WT were computed from pEC_{50} values reported in Ref. [3] (see also Table S4 in the SM).

presented here, these residues point to the outside of the GPCR into possible dimerization interfaces [66,67]. Hence, they may impact TGR5 function by interfering with TGR5 di-/multimerisation [68–70]. Macchiarulo et al. recently presented "binding mode 3" of agonist bile acids and derivatives in TGR5 [19]. However, this model could not explain all data from mutagenesis experiments, and ligands did not show any interaction with TM6 in this model (see below) [19]. Two further binding mode models presented by Yu et al. [37] and D'Amore et al. [36] differ in important aspects from the one of Macchiarulo et al. By pursuing an integrated computational and biological approach, here, we present an alternative and extended binding mode model of TGR5 agonists, including bile acids and neurosteroids, which is in agreement with experimental findings from eight mutagenesis experiments and in which the agonists interact with the conserved [19] residue E169$^{5.44}$ and residues in TM6 that are essential for GPCR activation.

Integrated approaches similar to ours have been successfully applied for deducing binding site characteristics and/or generating models of binding modes for other GPCRs [38–41]. However, the lack of an experimental high-resolution structure of TGR5 can lead to uncertainty in the structure-based prediction of the binding mode of TGR5 agonists, despite promising progress in molecular modeling and ligand docking for GPCRs [29]. We met this concern in four ways in our integrated approach.

I) We pursued the strategy to generate multiple homology models of TGR5 as well as multiple binding mode models within the TGR5 structures, and used those to screen for the one(s) that are consistent with knowledge on agonist interactions important for GPCR activation [31–35] and exhaustive experimental data on agonist potency (see III) below). This reduces the risk to commit to an inappropriate model early in the process. As such, we employed multi-template homology modeling on seven templates to foster the generation of structural models of TGR5 in which the structural uncertainty due to less well conserved regions in any single template is minimized [71] (step 1). Furthermore, we considered two alternatives for TM7 of TGR5 when generating the multiple sequence alignment with the templates (step 1) to account for the fact that – despite sequence identities and similarities of the TM regions of ~20% and ~50% – both the alignment of the (D/E)X(K/R) motif, frequently occurring in the template structures, or conserved Pro residues [72] could be meaningful. This resulted in 50 homology models generated for each alignment, respectively.
II) We accommodated structural uncertainties in the binding site region of these homology models when placing the agonist TLC by molecular docking (step 2). These uncertainties arise from the homology modeling procedure *per se* [29] as well as from using only antagonist-bound template structures for generating these models (see IV) below) [73]. We did so, first, by performing ensemble docking into subsets of ten homology models each. Docking into multiple fixed receptor conformations is a practical means to incorporate conformational rearrangements of a receptor binding pocket [74]. Second, we used DrugScore [75] as objective function, that way benefitting from the reduced steepness (i.e., increased softness) of the knowledge-based potentials, which has been recognized as an advantage [76] as such potentials are more robust to small changes in a receptor conformation.
III) For identifying a TLC binding mode (step 2), we exploited knowledge on the importance of residue E169$^{5.44}$ in TM5 derived from evolutionary information [19] and, more generally, of TM6 for GPCR activation [31–35] in the screening for docking solutions that "approach" these TGR5 regions. We did this in addition to evaluating the standard criteria of converged docking runs and energetically favorable docking solutions [77,78] so as to compensate for that the AutoDock/DrugScore combination shows success rates of generating "good" docking solutions in about 70–80% of the cases only in evaluation studies [79,80]. This resulted in two initial binding modes of TLC, one in a TGR5 model derived from the (D/E)X(K/R) motif-centered alignment and one in a TGR5 model derived from the proline-centered alignment. To further differentiate between the two, structure-based 3D-QSAR models based on the AFMoC approach [52] and 68 TGR5 agonists were generated. We expected that only an agonist alignment based on an appropriate initial binding mode will result in an AFMoC model with predictive power. Note that four negative controls were considered in this process, too, in that AFMoC models were also derived for two TLC configurations that did not meet the above "approach" criteria, as were AFMoC models for the two "approaching" TLC conformations that had been exchanged between the TGR5 models. Only the initial TLC binding mode in the proline-centered TGR5 model resulted in a good AFMoC model ($q^2 = 0.37$), whereas the four negative controls as well as the TLC binding mode in the (D/E)X(K/R) motif-centered TGR5 model resulted in AFMoC models without any predictive power.
IV) Predicting activated states and long loops of GPCRs is one of the remaining challenges when modeling GPCR structures [29]. We thus relaxed the initial binding mode of the agonist TLC in the proline-centered TGR5 model by exhaustive all-atom MD simulations in an explicit membrane environment of 0.9 μs length in total. We did so, first, to account for possible subtle changes [73] in the binding site of our TGR5 model, which was generated from antagonist-bound template structures, and, second, to enable a conformational sampling of the extracellular loop EL1 in the presence of the agonist. The overall TLC configuration remained stable; however, two important local changes were observed, leading to the involvement of Y240$^{6.51}$ in hydrogen bonding with the 3-hydroxyl group of TLC, in addition to E169$^{5.44}$, and to a reorientation (interaction) of the taurine tail of TLC towards (with) R79 of EL1. While interactions to TM6 have long been considered essential for GPCR activation [31–35], compelling evidence has also emerged in recent years revealing a critical role for ELs in many fundamental aspects of GPCR function, including ligand binding and activation [81]. Again, multiple relaxed binding modes of TLC were identified from the conformational ensemble based on them interacting with E169$^{5.44}$, Y240$^{6.51}$, and R79^{EL1}, which were subsequently evaluated by generating AFMoC models for the 68 TGR5

agonists superimposed onto them. Only the binding mode in which TLC interacts with all three of these residues resulted in a significant and good AFMoC model ($q^2 = 0.50$), the predictive power of which was also markedly improved compared to the AFMoC model for the initial binding mode. This suggests that the relaxed, final binding mode model is a better representation of the "true" binding mode.

In the final binding mode (Figs. 3A and 5), the TLC configuration resembles the location and orientation of an antagonist sphingolipid mimic co-crystallized with the human sphingosine 1-phosphate receptor 1 (S1P1) (PDB code: 3V2Y) [82] (Fig. S4 in the SM). TGR5 is phylogenetically related to members of the lipid activated family of GPCRs that include S1P receptors [1,83]; the human sphingosine 1-phosphate receptor 2 (S1P2) was shown to be activated by bile acids [84]. The hydrophobic cholane scaffold of TLC orients towards the interior of the orthosteric binding site such that rings C and D are rather parallel to TM3; this interaction is mimicked by the phenyl ring of the sphingolipid mimic in S1P1. Rings A and B of TLC are in contact with the hydrophobic side chains of TM5 and TM6; similar interactions occur between the aliphatic chain of the sphingolipid mimic and TM6 of S1P1. The taurine side chain is oriented towards the extracellular opening of the binding site, as is the phosphonic acid moiety of the sphingolipid mimic bound to S1P1. As to specific interactions of TLC, the 3-hydroxyl group forms hydrogen bonds with E169$^{5.44}$ and Y240$^{6.51}$, and the sulfonic acid moiety forms a salt bridge with R79^{EL1}. Neurosteroids, e.g., pregnandiol, adopt a similar binding mode as TLC (Fig. S8 in the SM), with rings A and B also interacting with TM5 and TM6, and rings C and D being slightly shifted towards the interior of the binding site.

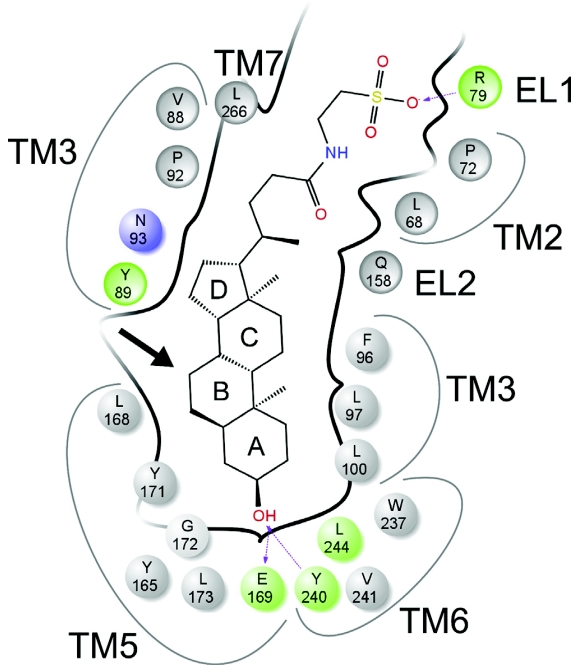


Fig. 5. 2D diagram of interactions between TLC and TGR5 in the final binding mode model. Identifiers for rings A–D of the cholane scaffold of TLC are shown. Interacting residues are shown in circles; hydrogen bonds and salt bridges are shown as purple arrows. The residues are colored according to having a negative effect (green) on receptor activity upon stimulation with TLC when mutated to alanine. The effect of the N93A$^{3.33}$ mutation (navy) could not be determined as the mutant was mainly retained intracellularly (see Fig. 4). Residues showing no effect on receptor activity upon stimulation with TLC, shown in orange in Fig. 1 and 2, are not interacting. Groups of residues in a TM are indicated by gray arcs. Position 7 of the cholane scaffold, which is hydroxylated in TCDC or TUDC, is indicated by a black arrow. In TCDC, the 7-hydroxyl group interacts with Y89$^{3.29}$ in the final binding mode model. The diagram was made with Maestro [120]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To further substantiate our binding mode model, we performed mutagenesis experiments of residues lining the binding site and compared predicted effects due to these mutations with the impact of the TGR5 variants on receptor activity (Figs. 2 and 4, Table 2). In particular, we distinguished between an effect on binding and one on activation. For the former, we expect that receptor function can be rescued by increasing the agonist concentration in the activity assay (i.e., by a right-shift in the dose response curve), whereas no such rescue is possible for the latter. Initially, we probed for interactions between the cholane scaffold and residues in TM5 and TM6 (upper part of Table 2). We expected that the E169A$^{5.44}$ mutation will influence the binding of TLC due to the loss of the hydrogen bond interaction with the 3-hydroxyl group. This prediction was confirmed by a right-shift in the dose–response curve in the E169A$^{5.44}$ variant (Fig. 2A, Table 2) and is in line with findings of other GPCR agonists interacting with the residue at position 5.44, too [59]. Macchiarulo et al. also reported an up to 70-fold increase in EC_{50} values of bile acid derivatives compared to wild type TGR5 for this mutation, but could not explain this finding by a loss of interactions to the bile acid derivatives as these are located ~12 Å away from this residue in their binding mode [19]. The binding mode of Yu et al. does not interact with E169$^{5.44}$ either [37], whereas the one of D'Amore does via the 3-hydroxyl group [36]. Macchiarulo et al. also discussed an alternative binding mode in their study in which the side chain of a bile acid points to the interior of the orthosteric binding site (referred to as "tail to head pose" in Ref. [19]). Such a binding mode is highly unlikely for TLC as it would bring two negatively charged moieties (E169$^{5.44}$ and the sulfonic acid group of TLC) in close proximity. Accordingly, the pEC_{50} value of TLC-3-sulfate is more than two log units smaller than that of TLC (Table S4 in the SM), likely a result of the repulsion between the sulfate group at position 3 of the cholane scaffold and E169$^{5.44}$ as suggested by our binding mode model (Fig. 3A).

For mutation L244A$^{6.55}$, we predicted a negative impact on activation of TGR5 as this mutation should abolish hydrophobic interactions to the A ring of the cholane scaffold. Residue 6.55 is often addressed by agonists binding to other GPCRs [31–35], and interactions with residues on TM6 are important for GPCR activation [31–35]. In agreement with this prediction, activation of L244A$^{6.55}$ by TLC showed a rise in luciferase activity for each TLC concentration that was significantly reduced compared to wildtype TGR5 (Fig. 2A). Regarding Y240$^{6.51}$, this residue was initially mutated to alanine, and we expected a negative impact on TGR5 activation for the same reasons as for L244$^{6.55}$. This prediction was confirmed in that the Y240A$^{6.51}$ variant showed only a slight increase in luciferase activity in response to TLC (Fig. 2A). As TLC forms a hydrogen bond with its 3-hydroxyl group with Y240$^{6.51}$ in the final binding mode model (Fig. 3A), we also investigated the Y240F$^{6.51}$ variant (lower part of Table 2); the loss of the hydrogen bond was expected to lead to a negative effect on TGR5 activation. Y240F$^{6.51}$ indeed almost completely abolished TLC induced luciferase activity (Fig. 4A). In none of the binding modes of Macchiarulo et al. [19], Yu et al. [37], and D'Amore et al. [36] did the bile acid derivative form a direct hydrogen bond between the 3-hydroxyl group and Y240$^{6.51}$. Note, that in both the Y240A$^{6.51}$ and Y240F$^{6.51}$ variants significantly diminished forskolin effects were found (Figs. 2A and 4A) suggesting an increased constitutive activity [61], which may be caused by the loss of intra-TGR5

interactions in the Y240A/F variants and, thus, a destabilization of the inactive state of TGR5.

Abolishing the salt bridge interaction predicted to exist between the sulfonate moiety of TLC and R79^{EL1} in the final binding mode model should have a negative impact on TLC binding. This prediction was confirmed by a R79A^{EL1} variant, which showed a significantly reduced and right-shifted luciferase activity compared to wildtype TGR5 (Fig. 4A). A similar salt bridge interaction has been observed between the phosphonic acid moiety of the sphingolipid mimic and a lysine of S1P1 [82]. In the binding mode of Yu et al. [37] the negatively charged side chain of the bile acid derivative interacts with R80^{EL1}; however, this interaction was enforced by restraint docking and did not develop in the course of unbiased MD simulations as in our case. Considering that TLC needs to adopt an extended conformation in order to interact simultaneously with E169$^{5.44}$, L244$^{6.55}$, Y240$^{6.51}$, and R79^{EL1} can explain why taurine-conjugated bile acid derivatives show in general higher pEC$_{50}$ values than glyco-conjugated or non-conjugated derivatives (Table S4 in the SM): In the latter cases, the shorter side chains can lead to weakened interactions to R79^{EL1} and/or conformational strain in EL1 to reach the side chain.

In contrast, no effect on TLC binding was expected for the S21A$^{1.39}$ mutation because this mutation served as negative control in order to exclude the possibility of a (D/E)X(K/R) motif-centered TGR5 model (Fig. S3A in the SM). As predicted, the TLC-dependent luciferase activity of S21A$^{1.39}$ was comparable to that of wildtype TGR5 (Fig. 4A). Thus, combining the results of the R79A^{EL1} mutation as a positive control as well as of the S21A$^{1.39}$ and S270A$^{7.43}$ mutations as two negative controls leads to the rejection of the binding mode found in the (D/E)X(K/R) motif-centered TGR5 model 42, in which the sulfonic acid moiety binds to S21$^{1.39}$. In the binding mode of D'Amore et al. the sulfate moiety of the bile acid derivative forms a water-mediated hydrogen bond with S21$^{1.39}$. Another mutation for which no effect on TLC binding was predicted is S270A$^{7.43}$: While S270$^{7.43}$ was involved in hydrogen bonding interactions with the sulfonate moiety of TLC in the initial binding mode (Fig. 1A), these interactions were not stable in any of the MD simulations used for refinement (Fig. S4E in the SM). This prediction was confirmed in that the activity of the S270A$^{7.43}$ variant was similar to that of wildtype TGR5 (Fig. 4A). This result differs from findings from Macchiarulo et al. according to which the S270A$^{7.43}$ variant resulted in a complete loss of activity [19]; Macchiarulo et al. did not quantify the plasma membrane localization of this variant [19]. In "binding mode 3" of Macchiarulo et al. the side chain of the bile acid derivative forms a hydrogen bond with S270$^{7.43}$. Finally, we did not expect an effect on TLC induced TGR5 activation by the N93A$^{3.33}$ mutation, as N93 does not interact with the 3-hydroxyl group in our final binding mode (Fig. 3). Introduction of N93A$^{3.33}$ led to a complete loss of TLC induced luciferase activity (Fig. 4A); however, immunolocalization of N93A$^{3.33}$ transfected HEK293 and CHO cells revealed that the mutated receptor was mainly retained intracellularly (Fig. S6 in the SM), and only 33% of the mutated protein reached the plasma membrane of transfected HEK293 cells, thus explaining the complete loss of functional activity of this mutant (Fig. 4C). Note that a complete loss of activity of a N93A$^{3.33}$ variant was also reported by Macchiarulo et al. [19]. However, while membrane localization in that study was assessed by immunofluorescence staining of the transfected CHO cells, no quantification of plasma membrane localization was performed. The authors thus interpreted the complete loss of activity as a confirmation of the non-existent hydrogen bonding with the 3-hydroxyl group of bile acid derivatives in their "binding mode 3" [19]. This data by Macchiarulo et al. [19] was also used by D'Amore et al. [36] and Yu et al. [37] to substantiate their binding modes.

Finally, we validated the prediction that Y89$^{3.29}$ is involved in hydrogen bond formation with the 7α-hydroxyl group of TCDC, which should result in a negative effect on the binding of that bile acid (Fig. 3C). Indeed, the Y89A$^{3.29}$ mutation significantly affected TCDC-dependent cAMP production at 2.5 and 10 μM (Fig. 4B) and leads to a ~13-fold increase in the EC$_{50}$ value compared to WT (Table 3). Macchiarulo et al. also investigated the activity of a Y89A$^{3.29}$ variant with respect to wild type TGR5 [19]; they did this in the context of hydrogen bond formation with the 3-hydroxyl group of bile acids in their "binding mode 3". However, while all six bile acid derivatives tested by these authors carried a 3-hydroxyl group, pronounced increases in EC$_{50}$ values were only found for those five derivatives that additionally carried a 7α-hydroxyl group [19]; this finding suggests that their results also reflect a loss of hydrogen bonding to the 7α- rather than the 3-hydroxyl group. The involvement of N93$^{3.33}$ in hydrogen bond formation with TUDC (Fig. 3D) could not be validated because the N93A$^{3.33}$ variant was mainly retained intracellularly (see above; Fig. 4C). However, the Y89A$^{3.29}$ variant only weakly (~2-fold) affected the EC$_{50}$ value of TUDC (Table 3), likely because the loss of hydrophobic interactions between the phenyl ring and ring D of the cholane scaffold (Fig. 3D) is compensated by the hydrogen bond formation with N93$^{3.33}$. Such a compensation is missing in the case of TLC, which can explain why a ~13-fold increase in the EC$_{50}$ value is found for the Y89A$^{3.29}$ variant compared to WT (Table 3). Taken together, our binding mode model provides an explanation for the observed epimeric selectivity for bile acids with a 7-hydroxyl group.

4. Conclusion

We presented a binding mode model of agonist bile acids and neurosteroids in TGR5 (Figs. 3A, B and 5), elucidated by an iterative and integrated computational and biological approach. An alignment of 68 TGR5 agonists based on this binding mode leads to a significant and good structure-based 3D-QSAR model, which constitutes the most comprehensive structure-based 3D-QSAR study undertaken so far for TGR5 agonists and suggests that our binding mode model is a close representation of the "true" binding mode. Our binding mode is further substantiated in that effects predicted for mutations in the binding site agree with experimental analyses on the impact of these TGR5 variants on receptor activity in all eight cases where the variants were trafficked to the plasma membrane. Our binding mode model differs in one or more of five aspects from binding mode models described recently: I) Our binding mode is oriented parallel to the membrane, rather than perpendicular to it as the binding mode of Yu et al. [37]; II) the cholane scaffold of bile acids binds in the vicinity of TM5 and TM6 and is rotated by 180° around the long axis compared to "binding mode 3" of Macchiarulo et al. [19]; III) the 3-hydroxyl group of bile acids interacts with the conserved [19] E169$^{5.44}$ but neither with N93$^{3.33}$ nor with W237$^{6.48}$, which is in contrast to both "binding mode 3" and the binding mode of Yu et al. [37]; IV) through this, our binding mode provides an explanation for the observed epimeric selectivity for bile acids with a 7-hydroxyl group; V) the side chains of bile acids orient towards EL1 and, hence, are distant to S270$^{7.43}$, in contrast to "binding mode 3" [19] and the binding mode of D'Amore et al. [36] These results are highly relevant for the development of novel, more potent agonists of TGR5. Furthermore, as only subtle changes in the binding site of active and inactive state GPCRs have been found [73], these results should also be a valuable starting point for the development of TGR5 antagonists, which could show antiproliferative effects in various gastrointestinal tumors [9,85–87] and also alleviate cholestatic pruritus [9,88,89].

5. Materials and methods

Homology modeling of the TGR5 structure. Templates for the homology modeling of TGR5 were identified by a PSI-BLAST [42] search on the Protein Data Bank [90] using an *E*-value cutoff of 5 to account for the low sequence identity among GPCRs. Of each class A GPCR type available with resolution <3 Å, the crystal structure with the lowest resolution was included into the set of templates (PDB codes: 2VT4 [91], 3D4S [92], 3EML [93], 3ODU [66], 3PBL [94], 3UON [95], and 3V2Y [96]). Subsequent alignment of the first 295 amino acids of the target sequence (UniProt accession number: Q8TDU6) and the template sequences was carried out with TCoffee [97]. The resulting multiple sequence alignment was manually corrected with the help of the program Jalview [98] considering secondary structure information. Due to two possibilities of aligning TM7 (see Results section for details), models for both alignments were generated, and both models were subjected to further analyses. The model generation was carried out with the program Modeller [99] (v9.9) in a multi-template approach employing the dope-loopmodel algorithm, owing to the length of EL2 of TGR5 (~18 amino acids). The disulfide bridge between residues 85 (position 3.25) and 155 (EL2) was explicitly specified in the modeling script. 50 models, generated from ten base models with five loop models each, were created for each of the two alignments. The resulting models were checked with Procheck [47] for their stereochemical properties, and models with errors, e.g. chain breaks, knots, and residues in disallowed regions in the Ramachandran plot, in the transmembrane regions were discarded. For each of the two alignments, ten models with structural variations in the binding pocket conformations were visually selected for further analyses. These models were aligned in Pymol [100] with the 'cealign' command onto the respective first homology model to allow easier comparison between the subsequent results.

Docking of TLC. For the molecular docking, TLC was drawn with ChemDraw Ultra [101], converted into a 3D structure, and energy minimized with Moloc using the MAB forcefield [102]. According to physiological pH, the sulfonate moiety was modeled in the deprotonated form. TLC was then docked into each of the 2× 10 TGR5 models using AutoDock3 [103] as a docking engine and the DrugScore [104] distance-dependent pair-potentials as an objective function as described in Ref. [50]. Because of the flexible taurine moiety, a clustering RMSD cutoff of 2.0 Å, a maximum number of $3*10^6$ energy evaluations, and a maximum number of $5*10^4$ generations were chosen; for all other docking parameters default values were used. Docking solutions with more than 20% of all configurations in the largest cluster were considered sufficiently converged, and the configuration with the lowest docking energy of that cluster was used for further evaluation.

Initial AFMoC analysis. For the AFMoC analysis, all TGR5 agonists were considered for which EC_{50} values were provided by Sato et al. [3]. This ensures that EC_{50} values are comparable as they were determined with the same assay [105]. The pEC_{50} values encompass a range of 3.62 log units and are approximately uniformly distributed. Agonist structures were drawn with ChemDraw Ultra [101], converted into a 3D structure, and energy minimized with Moloc using the MAB forcefield [102]. Subsequently, all ligands were aligned with Moloc via a pair- and atom-wise alignment onto the selected docking solutions of TLC using the cholane scaffold as a template for the orientation of the ligands. Finally, each ligand was minimized in the presence of the respective homology model, which was kept frozen. The AFMoC analysis was conducted with DrugScore potential fields, which encompass all ligands by at least 8 Å, and standard settings [52]. The influence of the σ value, which determines the "local smearing" of atomic protein-to-ligand interactions over neighboring points of a grid in the binding site, was tested by modifying σ from 0.55 to 1.15 Å in increments of 0.15 Å. For ξ-Me-LCA, for which the configuration at position 7 is unknown, the respective diastereomers α-Me-LCA and β-Me-LCA were tested in separate AFMoC analyses. For all AFMoC analyses leave-one-out (LOO) and leave-multiple-out analyses (LMO) were performed. Additionally, LOO-analyses after random scrambling of the biological data ("y-scrambling") were performed.

MD simulations of TGR5 with bound TLC in an explicit membrane. For generating the starting structure, the TGR5 orientation in the membrane was first determined by performing calculations with the Orientations of Proteins in Membranes (OPM) server [106] on the TGR5 homology model that has resulted in a valid AFMoC model. This includes the complete model containing loops except the last 35 residues of the C-terminus. The oriented TGR5 structure was then used for setting up the membrane for the MD simulations with the CHARMM-GUI [107]. The TGR5 termini were capped with ACE and NME groups for the N- and C-terminus, respectively. One hundred 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) lipids were selected for each layer of the membrane, and the system was hydrated with 15 Å of TIP3P [108] water on each side of the membrane, resulting in a rectangular box. The shortest distance between the edges of the box and the closest atom of the receptor was at least 10 Å. 15 mM KCl was included in the water as per default CHARMM-GUI settings. The resulting system size was ~85.000 atoms. Partial charges for TLC were determined according to the RESP procedure [109] applying gaussian09 [110] for the calculation of electrostatic potential maps and antechamber [63]; torsion angle parameters for TLC were determined following the FEW^{mem} approach [111,112]. MD simulations were performed with Amber14 [113] using the ff12SB, GAFF [64], and lipid14 [65] force fields. The equilibration of the system was carried out using pmemd [114,115] followed by production runs using the GPU-accelerated pmemd.cuda version [116].

The Particle Mesh Ewald [114] method was used to treat long-range electrostatic interactions. Bond lengths involving bonds to hydrogen atoms were constrained using SHAKE [117]. The time step for all MD simulations was 2 fs with a direct-space, non-bonded cutoff of 8 Å. Initially, 12,000 steps of steepest descent and conjugate gradient minimization were performed. Harmonic restraints with a force constant of 500 kcal mol^{-1} $Å^{-2}$ were applied to all solute atoms; the force constant was reduced to 10 kcal mol^{-1} $Å^{-2}$ after 2000 steps. Then, NVT-MD (MD simulations with a constant number of particles, volume, and temperature) was carried out for 50 ps during which the system was heated from 100 K to 300 K. Subsequent NPT-MD (MD simulations with a constant number of particles, pressure, and temperature) was used for 100 ps to adjust the solvent density. In both steps, harmonic restraints with a force constant of 10 kcal mol^{-1} $Å^{-2}$ were applied to all solute atoms. A final unrestrained NPT-MD was performed for 200 ps. In the following 100 ns of unrestrained NPT-MD, conformations were extracted every 40 ps for analysis. In addition, eight independent MD simulations were performed to assess the statistical significance of the results, each started from a slightly different end temperature after the equilibration step.

Afterwards, the trajectories were clustered with TLC as a reference. A hierarchical agglomerative clustering algorithm was applied as implemented in cpptraj [118], with a minimum distance of 1 Å between the clusters in combination with a symmetry-corrected root mean-square deviation (RMSD) between the structures as a distance measure. Subsequently, the cluster representatives were visually inspected for interactions between TLC and TGR5 residues known to influence the binding affinity.

Second AFMoC analyses. Cluster representatives of the nine simulations addressing at least two of the residues $R79^{EL1}$, $E169^{5.44}$, and $Y240^{6.51}$ were used as templates for another round of AFMoC

analyses. For this, the TLC configuration was energy minimized in the presence of the TGR5 conformation using Moloc. All receptor atoms were kept fixed except for side chains of $Y89^{3.29}$ and $Y240^{6.51}$. The AFMoC analyses were then carried out as described above. The AFMoC model with the highest q^2 was chosen as the final binding mode model.

Materials. Cell culture media was from PAA (Coelbe, Germany). Foetal calf serum (FCS) was from Biochrom (Berlin, Germany). Taurolithocholic acid (TLC), taurochenodeoxycholic acid (TCDC), tauroursodeoxycholic acid (TUDC) and forskolin (Forsk) were obtained from Sigma–Aldrich (Taufkirchen, Germany) and Calbiochem (San Diego, CA, USA), respectively.

Cloning of TGR5 variants. Human TGR5 was cloned as previously described [7,61]. Using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, USA) mutations were introduced in two different TGR5 cDNA constructs [61,62]: One construct, containing the part of the 5′UTR as well as the complete human TGR5 sequence including the stop codon, was cloned into the pcDNA3.1+ vector (Clontech, Palo Alto, CA) and used for immunofluorescence staining and luciferase assays. The second construct, FLAG-TGR5-YFP comprised the TGR5 CDS with an N-terminal FLAG-tag and was cloned into the pEYFP-N1 vector, generating a C-terminal YFP-Tag (Clontech, Palo Alto, CA). This construct was used for FACS-analysis. All mutagenesis primer sequences and cloning strategies can be obtained upon request. Successful cloning and mutagenesis was verified by sequencing and comparison to the TGR5 reference sequence (GenBank accession numbers: TGR5: NM_001077191.1).

Immunofluorescence and confocal laser scanning microscopy. Human embryonic kidney 293 (HEK293) cells were cultured in DMEM with 10% FCS and were kept at 37 °C and 5% CO_2. Cells (grown on glass coverslips) were transiently transfected with different TGR5-pcDNA3.1+ mutation constructs using Lipofectamine2000 (Invitrogen) for 48 h according to the manufacturer's recommendations. Fixed cells (methanol, −20 °C, 30 s) were incubated with the anti-TGR5 (RVLR2) antibody (directed against amino acids 298–318 of human TGR5; 1:500) [7] and sodium/potassium (Na^+/K^+; 1:100) ATPase (Sigma–Aldrich) diluted in 2% FCS in PBS ($^{-/-}$). Diluted fluoresceine (1:100) and Cyanine-3 (1:500) conjugated secondary antibodies were purchased from Dianova (Hamburg, Germany). Nuclei were stained with Hoechst 34580 (1:20,000; Invitrogen). Images were analyzed on a Zeiss LSM510META confocal microscope using a multitracking modus. A 63× objective and a scanning resolution of 1024 × 1024 pixels was used for all samples.

Flow cytometry. Plasma membrane localization of TGR5 was determined by flow cytometry (FACS) using a FACS-CANTO-II (BD Biosciences; Heidelberg, Germany) as described [61,62]. HEK293 cells were transiently transfected with the FLAG-TGR5-YFP constructs using Lipofectamine2000. The N-terminal FLAG-tag was detected with the anti-FLAG M2-antibody (Sigma–Aldrich) using the Zenon PacificBlue Label-Kit (Invitrogen) according to manufacturer's instructions. TGR5 PM expression was quantified by the amount of FLAG-tag positive cells divided by the total amount of TGR5 positive cells as determined by YFP-fluorescence [61,62].

Reporter gene assay. HEK293 cells grown on 6-wells were transiently co-transfected with human TGR5 variants in pcDNA3.1+ (0.5 μg diluted with 1.1 μg pEYFP-N1 vector), a reporter construct containing 5 cAMP responsive elements in front of the luciferase gene (Bayer AG; Leverkusen, Germany; 1.6 μg), and Renilla expression vector (Promega; Madison, WI, USA; 0.1 μg) using the transfection reagent Polyethylenimin (PEI, Sigma–Aldrich) [119]. The transfection approach was performed in Optimem medium in a 3:1 ratio of transfection reagent to DNA. After 48 h cells were lysed and luciferase assay was carried out with the dual-luciferase kit (Promega). Luciferase activity, normalized to transfection efficacy, served as measure for the rise in intracellular cAMP following forskolin stimulation (positive control) and TGR5 activation by TLC or TCDC and TUDC. Incubation with the respective substances was carried out over 16 h [61,62]. The increase in bile acid and forskolin-dependent luciferase activity is expressed relative to the DMSO stimulation (negative control), which was set to 1.0 for each experiment. The relative luciferase activity was plotted against the logarithm of the concentration of TLC, TCDC, and TUDC, respectively, in order to calculate the EC_{50}, and the data was evaluated using the Prism program.

Statistical analysis. Experiments were performed independently at least three times. Results are expressed as mean ± standard error of the mean (SEM) and analyzed using the two-sided student t-test. A $p \leq 0.01$ was considered statistically significant.

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding sources

This work was supported by the Deutsche Forschungsgemeinschaft through the Collaborative Research Center SFB 974 ("Communication and Systems Relevance during Liver Damage and Regeneration", Düsseldorf) and the Clinical Research Group KFO 217 ("Hepatobiliary Transport in Health and Disease", Düsseldorf). The funding sources had no involvement in the conduct of the research or preparation and submission of this article.

Notes

The authors declare no competing financial interest.

Acknowledgments

We are grateful to Stefanie Lindner und Waltraud Kuβ for technical assistance, to Dr. Nadine Homeyer, Yasemin Bilgic, and Alina Völz for help with the molecular modeling, and the "Zentrum für Informations und Medientechnologie" (ZIM) at the Heinrich Heine University for computational support.

Abbreviations

3D-QSAR	three-dimensional quantitative structure–activity relationship
AFMoC	Adaption of Fields for Molecular Comparison
MD	molecular dynamics
TCDC	taurochenodeoxycholic acid/taurochenodeoxycholate
TLC	taurolithocholic acid/taurocholate
TM	transmembrane helix
TUDC	tauroursodeoxycholic acid/tauroursodeoxycholate

Appendix A. Supplemental material

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejmech.2015.09.024.

References

[1] Y. Kawamata, R. Fujii, M. Hosoya, M. Harada, H. Yoshida, M. Miwa, S. Fukusumi, Y. Habata, T. Itoh, Y. Shintani, S. Hinuma, Y. Fujisawa, M. Fujino,

A G protein-coupled receptor responsive to bile acids, J. Biol. Chem. 278 (2003) 9435–9440.
[2] T. Maruyama, Y. Miyamoto, T. Nakamura, Y. Tamai, H. Okada, E. Sugiyama, T. Nakamura, H. Itadani, K. Tanaka, Identification of membrane-type receptor for bile acids (M-BAR), Biochem. Biophys. Res. Commun. 298 (2002) 714–719.
[3] H. Sato, A. Macchiarulo, C. Thomas, A. Gioiello, M. Une, A.F. Hofmann, R. Saladin, K. Schoonjans, R. Pellicciari, J. Auwerx, Novel potent and selective bile acid derivatives as TGR5 agonists: biological screening, structure-activity relationships, and molecular modeling studies, J. Med. Chem. 51 (2008) 1831–1841.
[4] G. Vassileva, A. Golovko, L. Markowitz, S.J. Abbondanzo, M. Zeng, S. Yang, L. Hoos, G. Tetzloff, D. Levitan, N.J. Murgolo, K. Keane, H.R. Davis Jr., J. Hedrick, E.L. Gustafson, Targeted deletion of Gpbar1 protects mice from cholesterol gallstone formation, Biochem. J. 398 (2006) 423–430.
[5] T. Maruyama, K. Tanaka, J. Suzuki, H. Miyoshi, N. Harada, T. Nakamura, Y. Miyamoto, A. Kanatani, Y. Tamai, Targeted disruption of G protein-coupled bile acid receptor 1 (Gpbar1/M-Bar) in mice, J. Endocrinol. 191 (2006) 197–205.
[6] V. Keitel, R. Reinehr, P. Gatsios, C. Rupprecht, B. Görg, O. Selbach, D. Häussinger, R. Kubitz, The G-protein coupled bile salt receptor TGR5 is expressed in liver sinusoidal endothelial cells, Hepatology 45 (2007) 695–704.
[7] V. Keitel, K. Cupisti, C. Ullmer, W.T. Knoefel, R. Kubitz, D. Häussinger, The membrane-bound bile acid receptor TGR5 is localized in the epithelium of human gallbladders, Hepatology 50 (2009) 861–870.
[8] V. Keitel, D. Häussinger, Perspective: TGR5 (Gpbar-1) in liver physiology and disease, Clin. Res. Hepatol. Gastroenterol. 36 (2012) 412–419.
[9] V. Keitel, M. Reich, D. Haussinger, TGR5: pathogenetic role and/or therapeutic target in fibrosing cholangitis? Clin. Rev. Allergy Immunol. 48 (2015) 218–225.
[10] C. Thomas, A. Gioiello, L. Noriega, A. Strehle, J. Oury, G. Rizzo, A. Macchiarulo, H. Yamamoto, C. Mataki, M. Pruzanski, R. Pellicciari, J. Auwerx, K. Schoonjans, TGR5-mediated bile acid sensing controls glucose homeostasis, Cell Metab. 10 (2009) 167–177.
[11] T.W.H. Pols, M. Nomura, T. Harach, G. Lo Sasso, Maaike H. Oosterveer, C. Thomas, G. Rizzo, A. Gioiello, L. Adorini, R. Pellicciari, J. Auwerx, K. Schoonjans, TGR5 activation inhibits atherosclerosis by reducing macrophage inflammation and lipid loading, Cell Metab. 14 (2011) 747–757.
[12] T.W.H. Pols, L.G. Noriega, M. Nomura, J. Auwerx, K. Schoonjans, The bile acid membrane receptor TGR5 as an emerging target in metabolism and inflammation, J. Hepatol. 54 (2011) 1263–1272.
[13] A. Perino, T.W.H. Pols, M. Nomura, S. Stein, R. Pellicciari, K. Schoonjans, TGR5 reduces macrophage migration through mTOR-induced C/EBPβ differential translation, J. Clin. Invest. 124 (2014) 5424–5436.
[14] A. Gioiello, E. Rosatelli, R. Nuti, A. Macchiarulo, R. Pellicciari, Patented TGR5 modulators: a review (2006 – present), Expert Opin. Ther. Pat. 22 (2012) 1399–1414.
[15] R. Pellicciari, A. Gioiello, A. Macchiarulo, C. Thomas, E. Rosatelli, B. Natalini, R. Sardella, M. Pruzanski, A. Roda, E. Pastorini, K. Schoonjans, J. Auwerx, Discovery of 6α-ethyl-23(S)-methylcholic acid (S-EMCA, INT-777) as a potent and selective agonist for the TGR5 receptor, a novel target for diabesity, J. Med. Chem. 52 (2009) 7958–7961.
[16] R. Pellicciari, H. Sato, A. Gioiello, G. Costantino, A. Macchiarulo, B.M. Sadeghpour, G. Giorgi, K. Schoonjans, J. Auwerx, Nongenomic actions of bile acids. synthesis and preliminary characterization of 23- and 6,23-alkyl-substituted bile acid derivatives as selective modulators for the g-protein coupled receptor TGR5, J. Med. Chem. 50 (2007) 4265–4268.
[17] M.R. Herbert, D.L. Siegel, L. Staszewski, C. Cayanan, U. Banerjee, S. Dhamija, J. Anderson, A. Fan, L. Wang, P. Rix, A.K. Shiau, T.S. Rao, S.A. Noble, R.A. Heyman, E. Bischoff, M. Guha, A. Kabakibi, A.B. Pinkerton, Synthesis and SAR of 2-aryl-3-aminomethylquinolines as agonists of the bile acid receptor TGR5, Bioorg. Med. Chem. Lett. 20 (2010) 5718–5721.
[18] B.W. Budzik, K.A. Evans, D.D. Wisnoski, J. Jin, R.A. Rivero, G.R. Szewczyk, C. Jayawickreme, D.L. Moncol, H. Yu, Synthesis and structure–activity relationships of a series of 3-aryl-4-isoxazolecarboxamides as a new class of TGR5 agonists, Bioorg. Med. Chem. Lett. 20 (2010) 1363–1367.
[19] A. Macchiarulo, A. Gioiello, C. Thomas, T.W. Pols, R. Nuti, C. Ferrari, N. Giacchè, F. De Franco, M. Pruzanski, J. Auwerx, Probing the binding site of bile acids in TGR5, ACS Med. Chem. Lett. 4 (2013) 1158–1162.
[20] A. Macchiarulo, A. Gioiello, C. Thomas, A. Massarotti, R. Nuti, E. Rosatelli, P. Sabbatini, K. Schoonjans, J. Auwerx, R. Pellicciari, Molecular field analysis and 3D-quantitative structure–activity relationship study (MFA 3D-QSAR) unveil novel features of bile acid recognition at TGR5, J. Chem. Inf. Model. 48 (2008) 1792–1801.
[21] C. Genet, C. Schmidt, A. Strehle, K. Schoonjans, J. Auwerx, R. Saladin, A. Wagner, Redefining the TGR5 triterpenoid binding pocket at the C-3 position, ChemMedChem 5 (2010) 1983–1988.
[22] C. Genet, A. Strehle, C. Schmidt, G. Boudjelal, A. Lobstein, K. Schoonjans, M. Souchet, J. Auwerx, R. Saladin, A. Wagner, Structure–activity relationship study of betulinic acid, a novel and selective TGR5 agonist, and its synthetic derivatives: potential impact in diabetes, J. Med. Chem. 53 (2009) 178–190.
[23] R.E. Martin, C. Bissantz, O. Gavelle, C. Kuratli, H. Dehmlow, H.G. Richter, U. Obst Sander, S.D. Erickson, K. Kim, S.L. Pietranico-Cole, 2-Phenoxy-nicotinamides are potent agonists at the bile acid receptor GPBAR1 (TGR5), ChemMedChem 8 (2013) 569–576.
[24] R. Pellicciari, A. Gioiello, P. Sabbatini, F. Venturoni, R. Nuti, C. Colliva, G. Rizzo, L. Adorini, M. Pruzanski, A. Roda, Avicholic acid: a lead compound from birds on the route to potent TGR5 modulators, ACS Med. Chem. Lett. 3 (2012) 273–277.
[25] J.A. Ballesteros, H. Weinstein, [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, in: C.S. Stuart (Ed.), Methods Neurosci., Academic Press, 1995, pp. 366–428.
[26] V. Isberg, B. Vroling, R. van der Kant, K. Li, G. Vriend, D. Gloriam, GPCRDB: an information system for G protein-coupled receptors, Nucleic Acids Res. 42 (2014) D422–D425.
[27] S.G. Rasmussen, B.T. DeVree, Y. Zou, A.C. Kruse, K.Y. Chung, T.S. Kobilka, F.S. Thian, P.S. Chae, E. Pardon, D. Calinski, J.M. Mathiesen, S.T. Shah, J.A. Lyons, M. Caffrey, S.H. Gellman, J. Steyaert, G. Skiniotis, W.I. Weis, R.K. Sunahara, B.K. Kobilka, Crystal structure of the beta2 adrenergic receptor-Gs protein complex, Nature 477 (2011) 549–555.
[28] A. Schafferhans, G. Klebe, Docking ligands onto binding site representations derived from proteins built by homology modelling, J. Mol. Biol. 307 (2001) 407–427.
[29] I. Kufareva, V. Katritch, Raymond C. Stevens, R. Abagyan, Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges, Structure 22 (2014) 1120–1139.
[30] I. Kufareva, M. Rueda, V. Katritch, R.C. Stevens, R. Abagyan, G.D. Participants, Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment, Structure 19 (2011) 1108–1126.
[31] M.R. Braden, J.C. Parrish, J.C. Naylor, D.E. Nichols, Molecular interaction of serotonin 5-HT2A receptor residues Phe339(6.51) and Phe340(6.52) with superpotent n-benzyl phenethylamine agonists, Mol. Pharmacol. 70 (2006) 1956–1964.
[32] F. Xu, H. Wu, V. Katritch, G.W. Han, K.A. Jacobson, Z.-G. Gao, V. Cherezov, R.C. Stevens, Structure of an agonist-bound human A2A adenosine receptor, Science 332 (2011) 322–327.
[33] K. Jarnagin, S. Bhakta, P. Zuppan, C. Yee, T. Ho, T. Phan, R. Tahilramani, J.H.B. Pease, A. Miller, R. Freedman, Mutations in the B2 bradykinin receptor reveal a different pattern of contacts for peptidic agonists and peptidic antagonists, J. Biol. Chem. 271 (1996) 28277–28286.
[34] J. Hwa, R.M. Graham, D.M. Perez, Identification of critical determinants of α1-adrenergic receptor subtype selective agonist binding, J. Biol. Chem. 270 (1995) 23189–23195.
[35] S. Moro, D. Guo, E. Camaioni, J.L. Boyer, T.K. Harden, K.A. Jacobson, Human P2Y1 receptor: molecular modeling and site-directed mutagenesis as tools to identify agonist and antagonist recognition sites, J. Med. Chem. 41 (1998) 1456–1466.
[36] C. D'Amore, F.S. Di Leva, V. Sepe, B. Renga, C. Del Gaudio, M.V. D'Auria, A. Zampella, S. Fiorucci, V. Limongelli, Design, synthesis, and biological evaluation of potent dual agonists of nuclear and membrane bile acid receptors, J. Med. Chem. 57 (2014) 937–954.
[37] D.D. Yu, K.M. Sousa, D.L. Mattern, J. Wagner, X. Fu, N. Vaidehi, B.M. Forman, W. Huang, Stereoselective synthesis, biological evaluation, and modeling of novel bile acid-derived G-protein coupled bile acid receptor 1 (GP-BAR1, TGR5) agonists, Bioorg. Med. Chem. 23 (2015) 1613–1628.
[38] A. Ivanenkov Yan, G. Majouga Alexander, S. Veselov Mark, V. Chufarova Nina, S. Baranovsky Sergey, I. Filkov Gleb, Computational approaches to the design of novel 5-HT6 R ligands, Rev. Neurosci. (2014) 451.
[39] A. Evers, G. Klebe, Ligand-supported homology modeling of G-protein-coupled receptor sites: models sufficient for successful virtual screening, Angew. Chem. Int. Ed. 43 (2004) 248–251.
[40] J. Zylberg, D. Ecke, B. Fischer, G. Reiser, Structure and ligand-binding site characteristics of the human P2Y11 nucleotide receptor deduced from computational modelling and mutational analysis, Biochem. J. 405 (2007) 277–286.
[41] J. Kim, J. Wess, A.M. van Rhee, T. Schöneberg, K.A. Jacobson, Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor, J. Biol. Chem. 270 (1995) 13987–13997.
[42] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[43] N. Eswar, M.A. Marti-Renom, B. Webb, M.S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative protein structure modeling with MODELLER, Curr. Protoc. Bioinforma. (2006) 5.6.1–5.6.30.
[44] E.B. Van Arnam, H.A. Lester, D.A. Dougherty, Dissecting the functions of conserved prolines within transmembrane helices of the D2 dopamine receptor, ACS Chem. Biol. 6 (2011) 1063–1068.
[45] P. Cozzini, G.E. Kellogg, F. Spyrakis, D.J. Abraham, G. Costantino, A. Emerson, F. Fanelli, H. Gohlke, L.A. Kuhn, G.M. Morris, M. Orozco, T.A. Pertinhez, M. Rizzi, C.A. Sotriffer, Target flexibility: an emerging consideration in drug discovery and design, J. Med. Chem. 51 (2008) 6237–6255.
[46] E.M. Novoa, L.R.d. Pouplana, X. Barril, M. Orozco, Ensemble docking from homology models, J. Chem. Theory. Comput. 6 (2010) 2547–2557.
[47] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck – a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[48] A.L. Morris, M.W. MacArthur, E.G. Hutchinson, J.M. Thornton, Stereochemical quality of protein structure coordinates, Proteins Struct. Funct. Bioinf. 12

(1992) 345–364.
[49] N. Widderich, M. Pittelkow, A. Höppner, D. Mulnaes, W. Buckel, H. Gohlke, S.H. Smits, E. Bremer, Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase, J. Mol. Biol. 426 (2014) 586–600.
[50] C.A. Sotriffer, H. Gohlke, G. Klebe, Docking into knowledge-based potential fields: a comparative evaluation of DrugScore, J. Med. Chem. 45 (2002) 1967–1970.
[51] D.M. Krüger, G. Jessen, H. Gohlke, How good are state-of-the-art docking tools in predicting ligand binding modes in protein–protein interfaces? J. Chem. Inf. Model. 52 (2012) 2807–2811.
[52] H. Gohlke, G. Klebe, DrugScore meets CoMFA: adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein, J. Med. Chem. 45 (2002) 4153–4170.
[53] J.K. Buolamwini, H. Assefa, CoMFA and CoMSIA 3D QSAR and docking studies on conformationally-restrained cinnamoyl HIV-1 integrase inhibitors: exploration of a binding mode at the active site, J. Med. Chem. 45 (2002) 841–852.
[54] T.I. Oprea, C.L. Waller, G.R. Marshall, Three-dimensional quantitative structure-activity relationship of human immunodeficiency virus (I) protease inhibitors. 2. Predictive power using limited exploration of alternate binding modes, J. Med. Chem. 37 (1994) 2206–2215.
[55] R. Ragno, M. Artico, G. De Martino, G. La Regina, A. Coluccia, A. Di Pasquali, R. Silvestri, Docking and 3-D QSAR studies on indolyl aryl sulfones. Binding mode exploration at the HIV-1 reverse transcriptase non-nucleoside binding site and design of highly active N-(2-hydroxyethyl)carboxamide and N-(2-hydroxyethyl)carbohydrazide derivatives, J. Med. Chem. 48 (2004) 213–223.
[56] P. Hecht, The Developing Practice of Comparative Molecular Field Analysis, in: 3D QSAR in Drug Design: Volume 1: Theory Methods and Applications, 1993, p. 443.
[57] C. Kramer, T. Kalliokoski, P. Gedeck, A. Vulpetti, The experimental uncertainty of heterogeneous public K_i data, J. Med. Chem. 55 (2012) 5165–5173.
[58] D.A. Shapiro, K. Kristiansen, W.K. Kroeze, B.L. Roth, Differential modes of agonist binding to 5-hydroxytryptamine2A serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5, Mol. Pharmacol. 58 (2000) 877–886.
[59] N.J. Pollock, A.M. Manelli, C.W. Hutchins, M.E. Steffey, R.G. MacKenzie, D.E. Frail, Serine mutations in transmembrane V of the dopamine D1 receptor affect ligand interactions and receptor activation, J. Biol. Chem. 267 (1992) 17780–17786.
[60] A. Mansour, F. Meng, J.H. Meador-Woodruff, L.P. Taylor, O. Civelli, H. Akil, Site-directed mutagenesis of the human dopamine D2 receptor, Eur. J. Pharmacol. Mol. Pharmacol. Sect. 227 (1992) 205–214.
[61] J.R. Hov, V. Keitel, J.K. Laerdahl, L. Spomer, E. Ellinghaus, A. ElSharawy, E. Melum, K.M. Boberg, T. Manke, T. Balschun, Mutational characterization of the bile acid receptor TGR5 in primary sclerosing cholangitis, PLoS One 5 (2010) e12403.
[62] L. Spomer, C.G.W. Gertzen, B. Schmitz, D. Häussinger, H. Gohlke, V. Keitel, A membrane-proximal, C-terminal α-helix is required for plasma membrane localization and function of the G protein-coupled receptor (GPCR) TGR5, J. Biol. Chem. 289 (2014) 3689–3702.
[63] V.B.D.A. Case, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, P.A. Kollman, AMBER 14, University of California, San Francisco, 2014.
[64] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general Amber force field, J. Comput. Chem. 25 (2004) 1157–1174.
[65] C.J. Dickson, B.D. Madej, Å.A. Skjevik, R.M. Betz, K. Teigen, I.R. Gould, R.C. Walker, Lipid14: the Amber lipid force field, J. Chem. Theory. Comput. 10 (2014) 865–879.
[66] B. Wu, E.Y.T. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi, D.J. Hamel, P. Kuhn, T.M. Handel, V. Cherezov, R.C. Stevens, Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, Science 330 (2010) 1066–1071.
[67] H. Wu, D. Wacker, M. Mileni, V. Katritch, G.W. Han, E. Vardy, W. Liu, A.A. Thompson, X.-P. Huang, F.I. Carroll, S.W. Mascarella, R.B. Westkaemper, P.D. Mosier, B.L. Roth, V. Cherezov, R.C. Stevens, Structure of the human [kgr]-opioid receptor in complex with JDTic, Nature (2012) (advance online publication).
[68] R.J. Ward, J.D. Pediani, A.G. Godin, G. Milligan, Regulation of oligomeric organization of the serotonin 5-HT2C receptor observed by spatial intensity distribution analysis, J. Biol. Chem. 290 (2015) 12844–12857.
[69] H. Zheng, E.A. Pearsall, D.P. Hurst, Y.H. Zhang, J. Chu, Y.L. Zhou, P.H. Reggio, H.H. Loh, P.Y. Law, Palmitoylation and membrane cholesterol stabilize mu-opioid receptor homodimerization and G protein coupling, BMC Cell Biol. 13 (2012).
[70] P.E. Volynsky, A.A. Polyansky, G.N. Fakhrutdinova, E.V. Bocharov, R.G. Efremov, Role of dimerization efficiency of transmembrane domains in activation of fibroblast growth factor receptor 3, J. Am. Chem. Soc. 135 (2013) 8105–8108.
[71] P. Larsson, B. Wallner, E. Lindahl, A. Elofsson, Using multiple templates to improve quality of homology models in automated homology modeling, Protein Sci. 17 (2008) 990–1002.
[72] M.S. Sansom, H. Weinstein, Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices, Trends Pharmacol. Sci. 21 (2000) 445–451.
[73] S.G. Rasmussen, H.J. Choi, J.J. Fung, E. Pardon, P. Casarosa, P.S. Chae, B.T. Devree, D.M. Rosenbaum, F.S. Thian, T.S. Kobilka, A. Schnapp, I. Konetzki, R.K. Sunahara, S.H. Gellman, A. Pautsch, J. Steyaert, W.I. Weis, B.K. Kobilka, Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor, Nature 469 (2011) 175–180.
[74] M. Totrov, R. Abagyan, Flexible ligand docking to multiple receptor conformations: a practical alternative, Curr. Opin. Struct. Biol. 18 (2008) 178–184.
[75] H. Gohlke, M. Hendlich, G. Klebe, Knowledge-based scoring function to predict protein-ligand interactions, J. Mol. Biol. 295 (2000) 337–356.
[76] P. Ferrara, H. Gohlke, D.J. Price, G. Klebe, C.L. Brooks 3rd, Assessing scoring functions for protein-ligand interactions, J. Med. Chem. 47 (2004) 3032–3047.
[77] D.M. Kruger, J. Bergs, S. Kazemi, H. Gohlke, Target flexibility in RNA-ligand docking modeled by elastic potential grids, ACS Med. Chem. Lett. 2 (2011) 489–493.
[78] S. Kazemi, D.M. Kruger, F. Sirockin, H. Gohlke, Elastic potential grids: accurate and efficient representation of intermolecular interactions for fully flexible docking, ChemMedChem 4 (2009) 1264–1268.
[79] C.A. Sotriffer, H. Gohlke, G. Klebe, Docking into knowledge-based potential fields: a comparative evaluation of DrugScore, J. Med. Chem. 45 (2002) 1967–1970.
[80] D.M. Kruger, G. Jessen, H. Gohlke, How good are state-of-the-art docking tools in predicting ligand binding modes in protein-protein interfaces? J. Chem. Inf. Model. 52 (2012) 2807–2811.
[81] M. Wheatley, D. Wootten, M.T. Conner, J. Simms, R. Kendrick, R.T. Logan, D.R. Poyner, J. Barwell, Lifting the lid on GPCRs: the role of extracellular loops, Br. J. Pharmacol. 165 (2012) 1688–1703.
[82] M.A. Hanson, C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer, M.G. Sanna, G.W. Han, P. Kuhn, H. Rosen, R.C. Stevens, Crystal structure of a lipid G protein-coupled receptor, Science 335 (2012) 851–855.
[83] K. Hama, J. Aoki, LPA3, a unique G protein-coupled receptor for lysophosphatidic acid, Prog. Lipid Res. 49 (2010) 335–342.
[84] E. Studer, X. Zhou, R. Zhao, Y. Wang, K. Takabe, M. Nagahashi, W.M. Pandak, P. Dent, S. Spiegel, R. Shi, W. Xu, X. Liu, P. Bohdan, L. Zhang, H. Zhou, P.B. Hylemon, Conjugated bile acids activate the sphingosine-1-phosphate receptor 2 in primary rodent hepatocytes, Hepatology 55 (2012) 267–276.
[85] H. Yasuda, S. Hirata, K. Inoue, H. Mashima, H. Ohnishi, M. Yoshiba, Involvement of membrane-type bile acid receptor M-BAR/TGR5 in bile acid-induced activation of epidermal growth factor receptor and mitogen-activated protein kinases in gastric carcinoma cells, Biochem. Biophys. Res. Commun. 354 (2007) 154–159.
[86] J. Hong, J. Behar, J. Wands, M. Resnick, L.J. Wang, R.A. DeLellis, D. Lambeth, R.F. Souza, S.J. Spechler, W. Cao, Role of a novel bile acid receptor TGR5 in the development of oesophageal adenocarcinoma, Gut 59 (2010) 170–180.
[87] W. Cao, W. Tian, J. Hong, D. Li, R. Tavares, L. Noble, S.F. Moss, M.B. Resnick, Expression of bile acid receptor TGR5 in gastric adenocarcinoma, Am. J. Physiol. Gastrointest. Liver Physiol. 304 (2013) G322–G327.
[88] T. Lieu, G. Jayaweera, P. Zhao, D.P. Poole, D. Jensen, M. Grace, P. McIntyre, R. Bron, Y.M. Wilson, M. Krappitz, S. Haerteis, C. Korbmacher, M.S. Steinhoff, R. Nassini, S. Materazzi, P. Geppetti, C.U. Corvera, N.W. Bunnett, The bile acid receptor TGR5 activates the TRPA1 channel to induce itch in mice, Gastroenterology 147 (2014) 1417–1428.
[89] F. Alemi, E. Kwon, D.P. Poole, T. Lieu, V. Lyo, F. Cattaruzza, F. Cevikbas, M. Steinhoff, R. Nassini, S. Materazzi, R. Guerrero-Alba, E. Valdez-Morales, G.S. Cottrell, K. Schoonjans, P. Geppetti, S.J. Vanner, N.W. Bunnett, C.U. Corvera, The TGR5 receptor mediates bile acid-induced itch and analgesia, J. Clin. Invest. 123 (2013) 1513–1530.
[90] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[91] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G.W. Leslie, C.G. Tate, G.F.X. Schertler, Structure of a [bgr]1-adrenergic G-protein-coupled receptor, Nature 454 (2008) 486–491.
[92] M.A. Hanson, V. Cherezov, M.T. Griffith, C.B. Roth, V.-P. Jaakola, E.Y.T. Chien, J. Velasquez, P. Kuhn, R.C. Stevens, A specific cholesterol binding site is established by the 2.8 Å structure of the human β2-adrenergic receptor, Structure 16 (2008) 897–905.
[93] V.-P. Jaakola, M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y.T. Chien, J.R. Lane, A.P. IJzerman, R.C. Stevens, The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist, Science 322 (2008) 1211–1217.
[94] E.Y.T. Chien, W. Liu, Q. Zhao, V. Katritch, G. Won Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov, R.C. Stevens, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, Science 330 (2010) 1091–1095.
[95] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga, T. Kobayashi, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, Nature 482 (2012) 547–551.

[96] M.A. Hanson, C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer, M.G. Sanna, G.W. Han, P. Kuhn, H. Rosen, R.C. Stevens, Crystal structure of a lipid G protein–coupled receptor, Science 335 (2012) 851–855.
[97] C. Notredame, D.G. Higgins, J. Heringa, T-coffee: a novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 (2000) 205–217.
[98] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview version 2—a multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (2009) 1189–1191.
[99] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using modeller, in: Curr. Protoc. Bioinformatics, John Wiley & Sons, Inc., 2002.
[100] L. Schrödinger, The PyMOL Molecular Graphics System, Version 1.3 r1, PyMOL, the PyMOL Molecular Graphics System, Version, 1, 2010.
[101] N. Mills, ChemDraw ultra 10.0, J. Am. Chem. Soc. 128 (2006) 13649–13650.
[102] P.R. Gerber, K. Müller, MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry, J. Comput. Aided Mol. Des. 9 (1995) 251–268.
[103] F. Osterberg, G.M. Morris, M.F. Sanner, A.J. Olson, D.S. Goodsell, Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock, Proteins 46 (2002) 34–40.
[104] S. Radestock, M. Bohm, H. Gohlke, Improving binding mode predictions by docking into protein-specifically adapted potential fields, J. Med. Chem. 48 (2005) 5466–5479.
[105] T. Kalliokoski, C. Kramer, A. Vulpetti, P. Gedeck, Comparability of mixed IC_{50} data – a statistical analysis, PLoS One 8 (2013) e61007.
[106] M.A. Lomize, A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, OPM: orientations of proteins in membranes database, Bioinformatics 22 (2006) 623–625.
[107] S. Jo, J.B. Lim, J.B. Klauda, W. Im, CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes, Biophys. J. 97 (2009) 50–58.
[108] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. (1983) 926–935.
[109] C.I. Bayly, P. Cieplak, W. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, J. Phys. Chem. 97 (1993) 10269–10280.
[110] M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, Gaussian 09, Revision B. 01, Gaussian, Inc., Wallingford, CT, 2010.
[111] N. Homeyer, H. Gohlke, Extension of the free energy workflow FEW towards implicit solvent/implicit membrane MM–PBSA calculations, Biochim. Biophys. Acta Gen. Subj. 1850 (2015) 972–982.
[112] N. Homeyer, H. Gohlke, FEW: a workflow tool for free energy calculations of ligand binding, J. Comput. Chem. 34 (2013) 965–973.
[113] D. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, H. Gohlke, Amber 14, 2014.
[114] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log (N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[115] D.M. York, T.A. Darden, L.G. Pedersen, The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods, J. Chem. Phys. 99 (1993) 8345–8348.
[116] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with Amber on GPUs. 2. Explicit solvent particle mesh Ewald, J. Chem. Theory. Comput. 9 (2013) 3878–3888.
[117] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[118] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory. Comput. 9 (2013) 3084–3095.
[119] O. Boussif, F. Lezoualc'h, M.A. Zanta, M.D. Mergny, D. Scherman, B. Demeneix, J.-P. Behr, A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine, PNAS 92 (1995) 7297–7301.
[120] Maestro, in: Schrödinger, LLC, New York, NY, 2014.

Publication III – Supporting Information

Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists

Christoph G.W. Gertzen[§], Lina Spomer[§], Sander H.J. Smits, Dieter Häussinger, Verena Keitel, Holger Gohlke

[§] Both authors contributed equally to this work.

Eur J Med Chem (2015), 104, 57-72.

Supplementary Material

Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of Bile Acid and Neurosteroid Agonists

Christoph G. W. Gertzen[1], Lina Spomer[2], Sander H.J. Smits[3], Dieter Häussinger[2], Verena Keitel[2]*, Holger Gohlke[1]*

[1]Institute for Pharmaceutical and Medicinal Chemistry, [2]Clinic for Gastroenterology, Hepatology, and Infectious Diseases, and [3]Institute for Biochemistry, Heinrich-Heine-University, Universitätsstr. 1, 40225 Düsseldorf, Germany.

Table of Contents

Supplemental Tables

Table S1. Stereochemical quality of the homology models generated from the (D/E)X(K/R) motif-centered alignment as analyzed by Procheck[1].

Model	Residues in core regions[a]	Residues in allowed regions[a]	Residues in generous regions[a]	Residues in disallowed regions[a]
1	90.6	7.1	2.0	0.4
2	86.3	8.2	3.9	1.6
3	89.8	6.3	2.7	1.2
4	85.9	10.6	2.0	1.6
5	87.5	6.7	4.7	1.2
6	89.0	8.6	1.2	2.4
7	88.2	8.2	1.2	2.4
8	89.4	7.8	1.2	1.6
9	89.0	9.0	1.2	0.8
10	87.8	9.0	2.0	1.2
11	87.8	10.6	0.4	1.2
12	87.5	8.2	2.4	2.0
13	89.8	5.9	3.9	0.4
14	85.5	10.6	2.0	2.0
15	87.1	6.7	5.1	1.2
16	89.0	7.5	1.6	2.0
17	89.0	8.6	0.8	1.6
18	89.8	7.5	1.6	1.2
19	91.0	7.8	0.8	0.4
20	86.7	9.0	3.1	1.2
21	87.1	10.2	2.0	0.8
22	86.7	9.0	3.1	1.2
23	89.0	9.0	1.2	0.8
24	85.1	11.4	2.7	0.8
25	89.0	5.5	4.3	1.2
26	89.4	8.2	1.8	0.8
27	89.8	8.2	0.8	1.2
28	89.4	7.5	1.2	2.0
29	90.2	8.2	1.2	0.4
30	86.7	9.4	2.4	1.6
31	88.2	10.2	0.4	1.2
32	85.9	8.2	3.1	2.7
33	87.5	6.3	4.3	2.0
34	85.5	10.6	2.0	2.0
35	87.8	5.9	4.7	1.6
36	88.6	7.5	2.0	2.0
37	90.2	8.2	0.8	0.8
38	87.5	8.6	2.4	1.6
39	89.8	8.2	1.6	0.4

Table S1 continued

40	87.8	9.4	2.0	0.8
41	87.1	11.0	0.4	1.6
42	86.3	8.2	3.9	1.6
43	89.8	5.9	3.5	0.8
44	85.5	11.0	2.0	1.6
45	88.2	6.3	4.7	0.8
46	88.2	8.6	2.4	0.8
47	88.6	9.0	1.2	1.2
48	89.0	7.8	1.2	2.0
49	90.2	8.6	0.8	0.4
50	87.5	9.4	2.0	1.2

[a] The percentage of residues is given that lie within core, allowed, generous, or disallowed regions of the Ramachandran plot according to the definitions used by Procheck[1].

Table S2. Stereochemical quality of the homology models generated from the conserved proline-centered alignment as analyzed by Procheck[1].

Model	Residues in core regions[a]	Residues in allowed regions[a]	Residues in generous regions[a]	Residues in disallowed regions[a]
1	87.9	8.2	2.3	1.6
2	89.1	8.9	1.2	0.8
3	89.1	6.6	1.6	2.7
4	87.9	9.7	1.9	0.4
5	87.9	9.3	2.3	0.4
6	87.5	8.9	1.6	1.9
7	88.3	8.9	1.9	0.8
8	90.7	8.2	0.0	1.2
9	88.3	10.1	0.8	0.8
10	86.4	8.6	2.7	2.3
11	88.7	7.4	1.9	1.9
12	89.5	8.9	0.8	0.8
13	90.7	6.2	1.2	1.9
14	87.9	10.1	1.6	0.4
15	87.5	9.3	1.9	1.2
16	87.2	8.2	1.6	3.1
17	88.3	8.2	1.9	1.6
18	89.9	8.6	0.4	1.2
19	88.3	9.7	1.2	0.8
20	87.5	8.6	2.3	1.6
21	87.9	8.2	1.9	1.9
22	89.1	8.9	1.2	0.8
23	90.3	6.2	1.6	1.9
24	88.7	8.9	1.9	0.4
25	87.9	8.9	2.3	0.8
26	87.5	8.6	1.9	1.9
27	88.7	8.9	1.6	0.8
28	89.5	8.6	0.4	1.6
29	88.7	9.3	1.6	0.4
30	86.8	8.2	3.1	1.9
31	88.3	7.4	1.9	2.3
32	89.9	8.6	0.8	0.8
33	89.5	7.0	1.6	1.9
34	87.9	10.1	1.6	0.4
35	87.5	9.7	2.3	0.4
36	87.9	8.9	1.6	1.6
37	88.3	8.6	1.6	1.6
38	89.9	8.2	0.4	1.6
39	88.3	9.7	1.2	0.8

Table S2 continued

40	87.2	8.9	2.3	1.6
41	89.1	7.4	1.9	1.6
42	88.7	8.9	1.2	1.2
43	88.7	7.8	0.8	2.7
44	88.3	9.3	1.9	0.4
45	87.2	10.1	2.3	0.4
46	87.5	8.9	1.6	1.9
47	88.3	8.9	1.9	0.8
48	89.9	8.2	0.4	1.6
49	88.3	10.1	1.2	0.4
50	86.4	8.9	3.1	1.6

[a] The percentage of residues is given that lie within core, allowed, generous, or disallowed regions of the Ramachandran plot according to the definitions used by Procheck[1].

Table S3. Results of the molecular docking of TLC into selected homology models.

Alignment	Model[a]	Poses in the largest cluster[b]	Energy of the best pose in the largest cluster[c]
(D/E)X(R/K)	7	20	-11.92
	22	80	-12.75
	23	50	-13.37
	24	46	-13.70
	26	42	-13.10
	36	35	-14.25
	42	83	-12.75
	48	28	-12.56
	49	51	-12.82
	50	36	-13.12
Proline	9	50	-13.83
	11	76	-13.10
	12	34	-11.65
	17	64	-14.65
	18	60	-12.43
	19	22	-14.92
	22	34	-11.58
	25	40	-12.88
	26	31	-11.67
	34	32	-14.07

[a] The model number refers to the number assigned in the modelling process.

[b] Percentage of poses in the largest cluster.

[c] In kcal mol^{-1}.

Table S4. TGR5 agonists with a cholane scaffold used for the AFMoC analyses. Adapted from ref. [2].

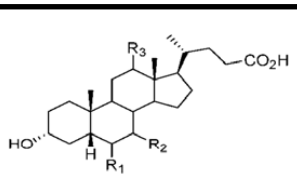
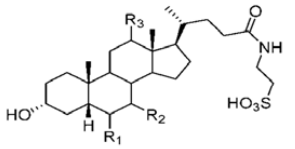
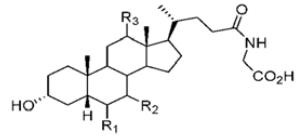
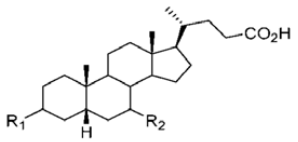
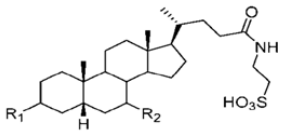
Triple substituted TGR5 agonists					
Ligand	**R_1**	**R_2**	**R_3**	**pEC_{50}**	**Structure**
Acid-form					
CDCA	—H	α-OH	—H	5.17	
CA	—H	α-OH	α-OH	4.87	
LCA	—H	—H	—H	6.24	
DCA	—H	—H	α-OH	5.90	
lago-DCA	—H	—H	β-OH	5.36	
UDCA	—H	β-OH	—H	4.44	
HDCA	α-OH	—H	—H	4.50	
muro-CA	β-OH	—H	—H	5.31	
Tauro-form					
TCDC	—H	α-OH	—H	5.72	
TCA	—H	α-OH	α-OH	5.31	
TLC	—H	—H	—H	6.54	
TDC	—H	—H	α-OH	6.10	
TUDC	—H	β-OH	—H	4.52	
THDC	α-OH	—H	—H	4.62	
Glyco-form					
GCDC	—H	α-OH	—H	5.41	
GCA	—H	α-OH	α-OH	4.87	
GLC	—H	—H	—H	6.27	
GDC	—H	—H	α-OH	5.93	
GUDC	—H	β-OH	—H	4.47	
GHDC	α-OH	—H	—H	4.44	
Double substituted TGR5 agonists					
Ligand	**R_1**	**R_2**		**pEC_{50}**	**Structure**
Acid-form					
5β-cholanic-acid	—H	—H		5.22	
LCA-S	α-OSO_3H	—H		<4.00	
LCA-Ac	α-OCOCl	—H		<3.99	
dehydro-LCA	= O	—H		6.57	
iso-LCA	β-OH	—H		5.90	
7-ξ-Me-LCA	α-OH	—CH_3		7.12	
7α-F-LCA	α-OH	α-F		6.60	
7β-F-LCA	α-OH	β-F		5.64	
3-dehydro-CDCA	= O	α-OH		5.40	
3-deoxy-CDCA	—H	α-OH		4.84	
7β-Me-CDCA	α-OH	α-OH, β-CH_3		5.21	
7β-Et-CDCA	α-OH	α-OH, β-CH_2-CH_3		5.58	
7β-Pr-CDCA	α-OH	α-OH, β-$(CH_2)_2$-CH_3		6.11	
Tauro-form					
TLC-S	α-OSO_3H	—H		4.32	
7β-Me-TCDC	α-OH	α-OH, β-CH_3		5.63	
7β-Et-TCDC	α-OH	α-OH, β-CH_2-CH_3		6.14	
7β-Pr-TCDC	α-OH	α-OH, β-$(CH_2)_2$-CH_3		6.44	

Table S4 continued

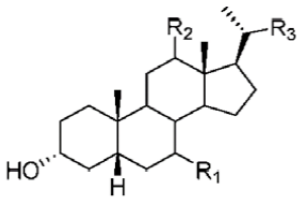
Sidechain substituted TGR5 agonists					
Ligand	**R_1**	**R_2**	**R_3**	**pEC_{50}**	**Structure**
LCA Me-Ester	—H	—H		6.23	
Nor-LCA	—H	—H		6.11	
CDCA Me-Ester	α-OH	—H		5.36	
Nor-CDCA	α-OH	—H		4.98	
Dinor-CDCA	α-OH	—H		<4.00	
CDC-OH	α-OH	—H		6.92	
CDC-Sul	α-OH	—H		6.36	
22S,23S-CCDCA	α-OH	—H		5.88	
22S,23R-CCDCA	α-OH	—H		5.54	
22R,23R-CCDCA	α-OH	—H		4.12	
22R,23S-CCDCA	α-OH	—H		<4.00	
C-OH	α-OH	α-OH		6.06	
C-Sul	α-OH	α-OH		6.00	
Nor-UDCA	β-OH	—H		4.33	
Dinor-UDCA	β-OH	—H		<3.50	
UDC-OH	β-OH	—H		5.66	
UDC-Sul	β-OH	—H		5.30	

Table S5. TGR5 agonists with different scaffolds used for the AFMoC analyses. Adapted from ref. [2].

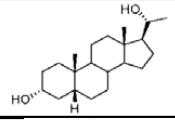
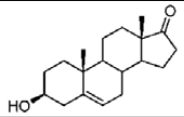
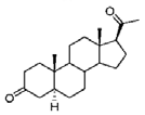
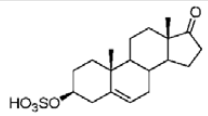
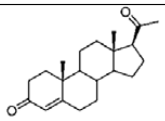
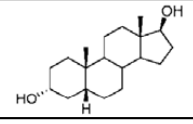
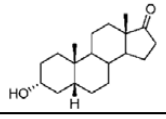
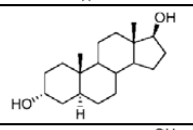
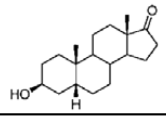
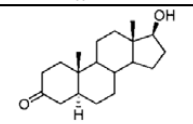
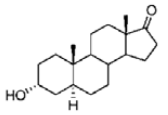
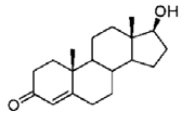
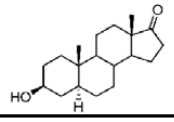
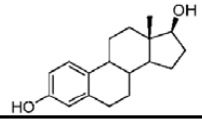
Ligand	Structure	pEC_{50}	Ligand	Structure	pEC_{50}
Pregnandiol		6.07	Dehydroepi-androsterone		5.48
5α-Pregnandione		6.54	Dehydroepi-androsterone-Sul		3.90
Progesterone		5.56	Etiocholandiol		5.61
Etiocholanolone		5.42	Androstandiol		5.15
Epieticholanolone		5.58	Androstanolone		5.35
Androsterone		5.21	Testosterone		4.74
Epiandrosterone		5.49	17β-Estradiol		4.42

Table S6. q^2 values of AFMoC analyses for model/ligand combinations of different docking solutions.

Model[a]	Alignment[b]	Template ligand[c]	q^2 [d]
19	Proline	19	0.37 (0.37)
19	Proline	42	-0.09 (-0.09)
26	Proline	26	0.03 (0.03)
34	Proline	34	-0.1 (-0.1)
42	(D/E)X(K/R)	19	-0.01 (-0.01)
42	(D/E)X(K/R)	42	-0.13 (-0.13)

[a] The model number refers to the number assigned in the modelling process.

[b] The center the sequence alignment of the homology models is based on.

[c] The template ligand numbers refers to the number of the homology model it was found in.

[d] q^2 = 1 - PRESS/SSD as obtained by "leave-one-out" cross-validation. PRESS equals the sum of squared differences between predicted and experimentally determined binding affinities, SSD is the sum of the squared differences between experimentally determined binding affinities and the mean of the training set binding affinities. Values are given considering only the part of the binding affinity (pEC_{50}^{PLS}) used in the PLS analysis or considering the total binding affinity (values in parentheses).

Supplemental Figures

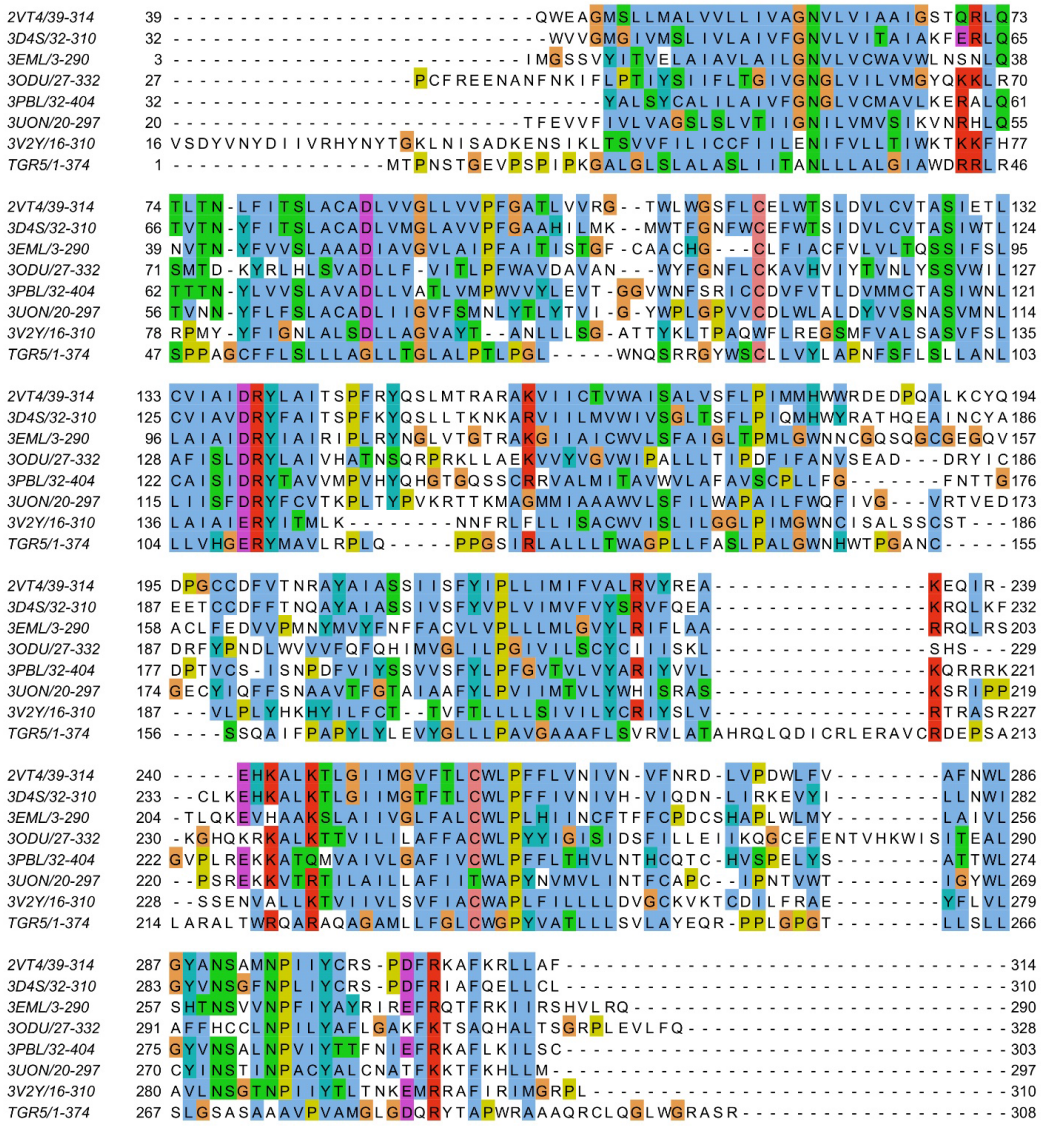


Figure S1. Sequence alignment for the homology modelling for the first alignment alternative ((D/E)X(K/R)-centered).

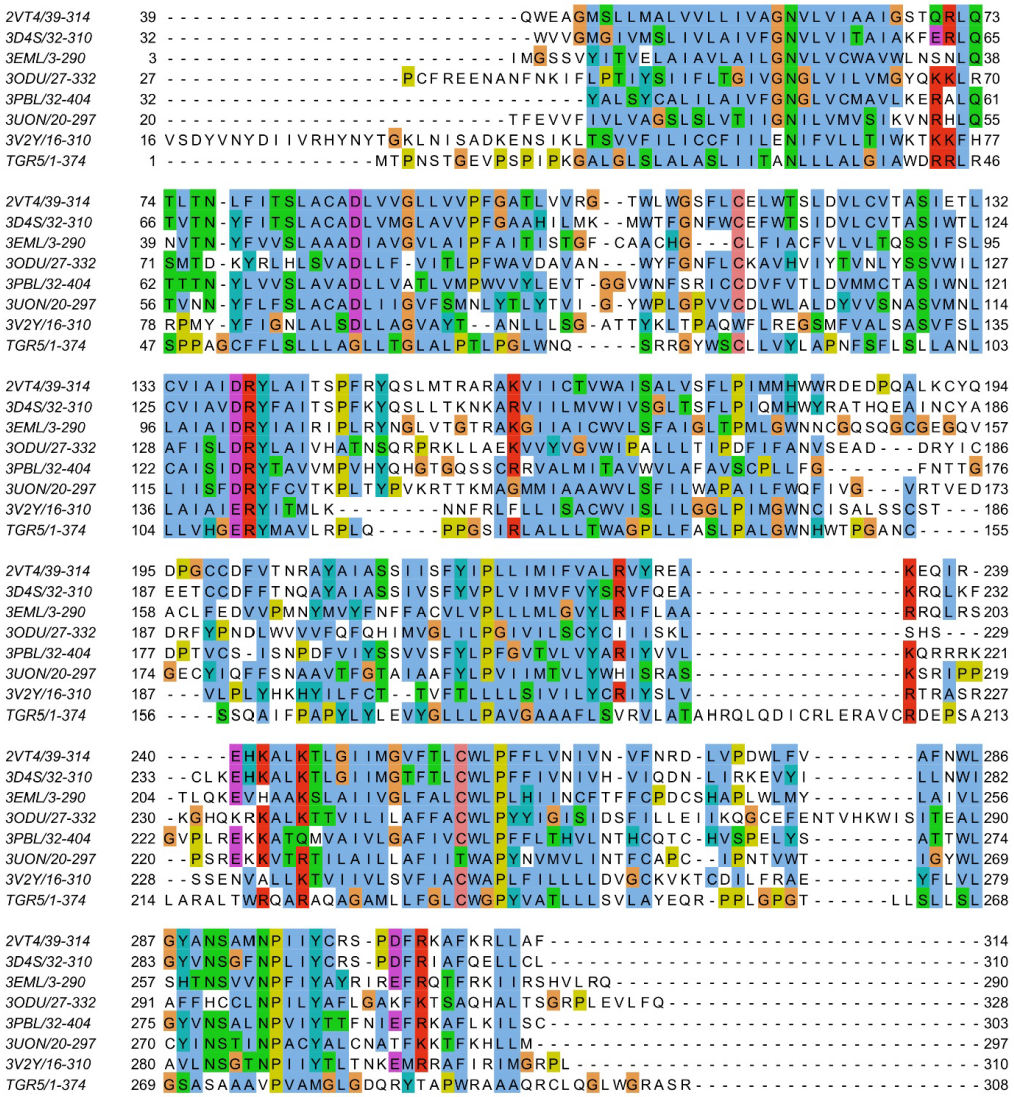


Figure S2. Sequence alignment for the homology modelling for the second alignment alternative (conserved proline-centered).

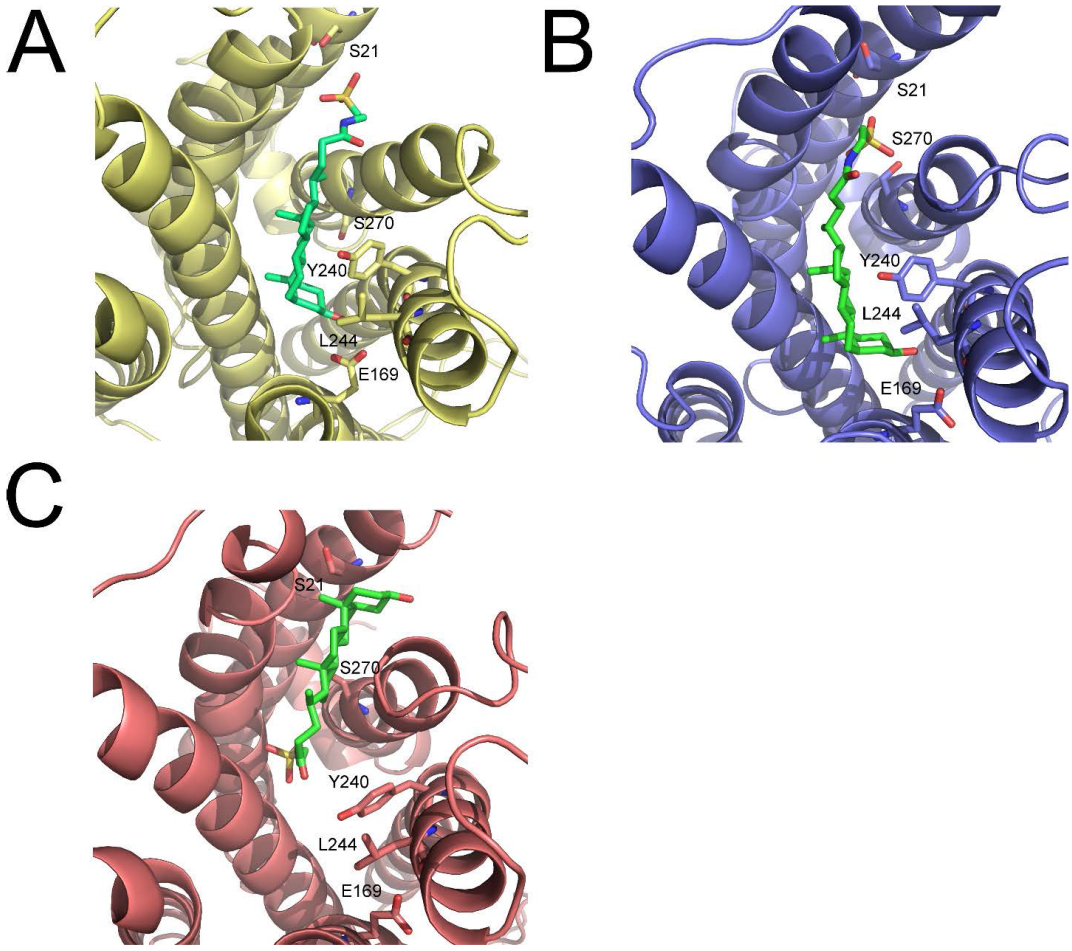


Figure S3. Alignment templates of AFMoC models resulting in $q^2 < 0.03$ (see Table S4), using the docking pose of TLC in **A:** model 42 of the (D/E)X(K/R)-centered alignment, **B:** model 26 of the conserved proline-centered alignment, and **C:** model 34 of the conserved proline-centered alignment.

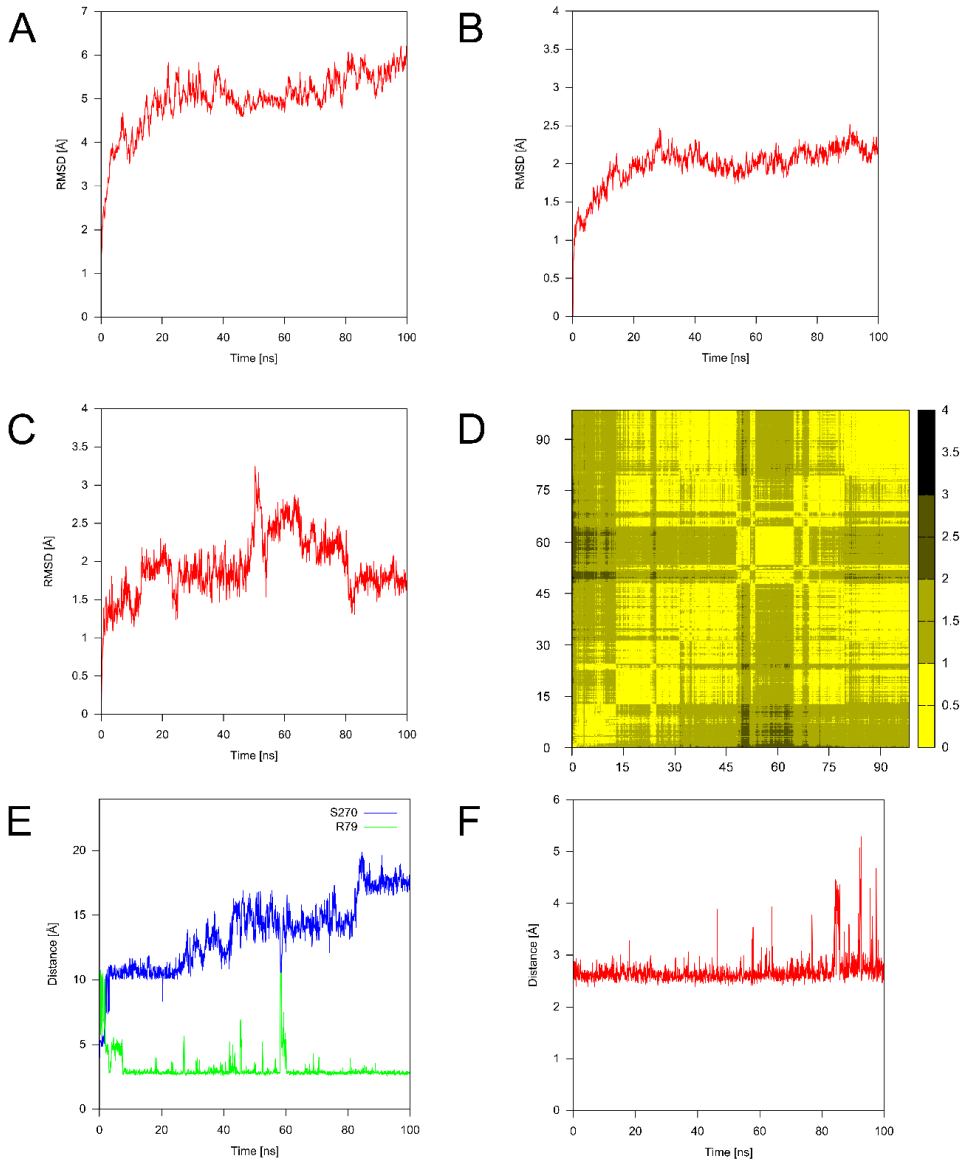


Figure S4. Analysis of the MD simulation of TLC in model 19 of the proline-centered alignment. **A:** All-atom root mean square deviation (RMSD) of TGR5. **B:** RMSD of all C_α-atoms of the transmembrane regions. **C:** All-atom symmetry corrected RMSD of TLC after superimpositioning all protein atoms of TGR5. **D:** 2D-RMSD of TLC after superimpositioning all protein atoms of TGR5. **E:** Shortest distance between any of the oxygen atoms of the sulfonic acid group of TLC and the hydroxyl oxygen of $S270^{7.43}$ (blue) or any of the guanidine nitrogen atoms of $R79^{EL1}$ (green). **F:** Shortest distance between the 3-hydroxyl oxygen of TLC and any of the carboxyl oxygen atoms of $E169^{5.44}$. ss

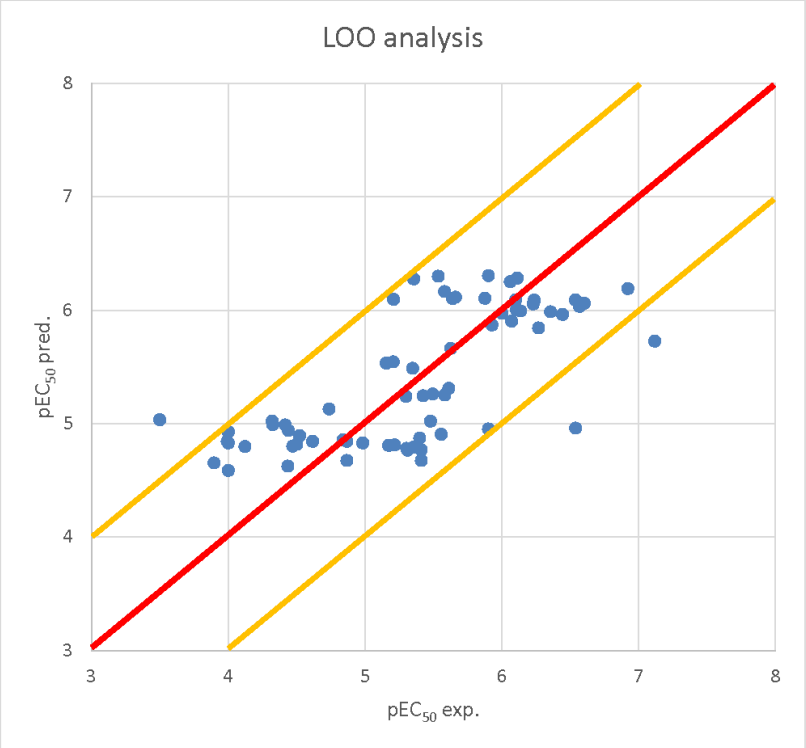


Figure S5. Scatterplot of the predicted versus experimentally determined pEC_{50} values taken from the leave-one-out analysis of the final AFMoC model. The red line indicates a perfect correlation, while the orange lines indicate a deviation by one log-unit.

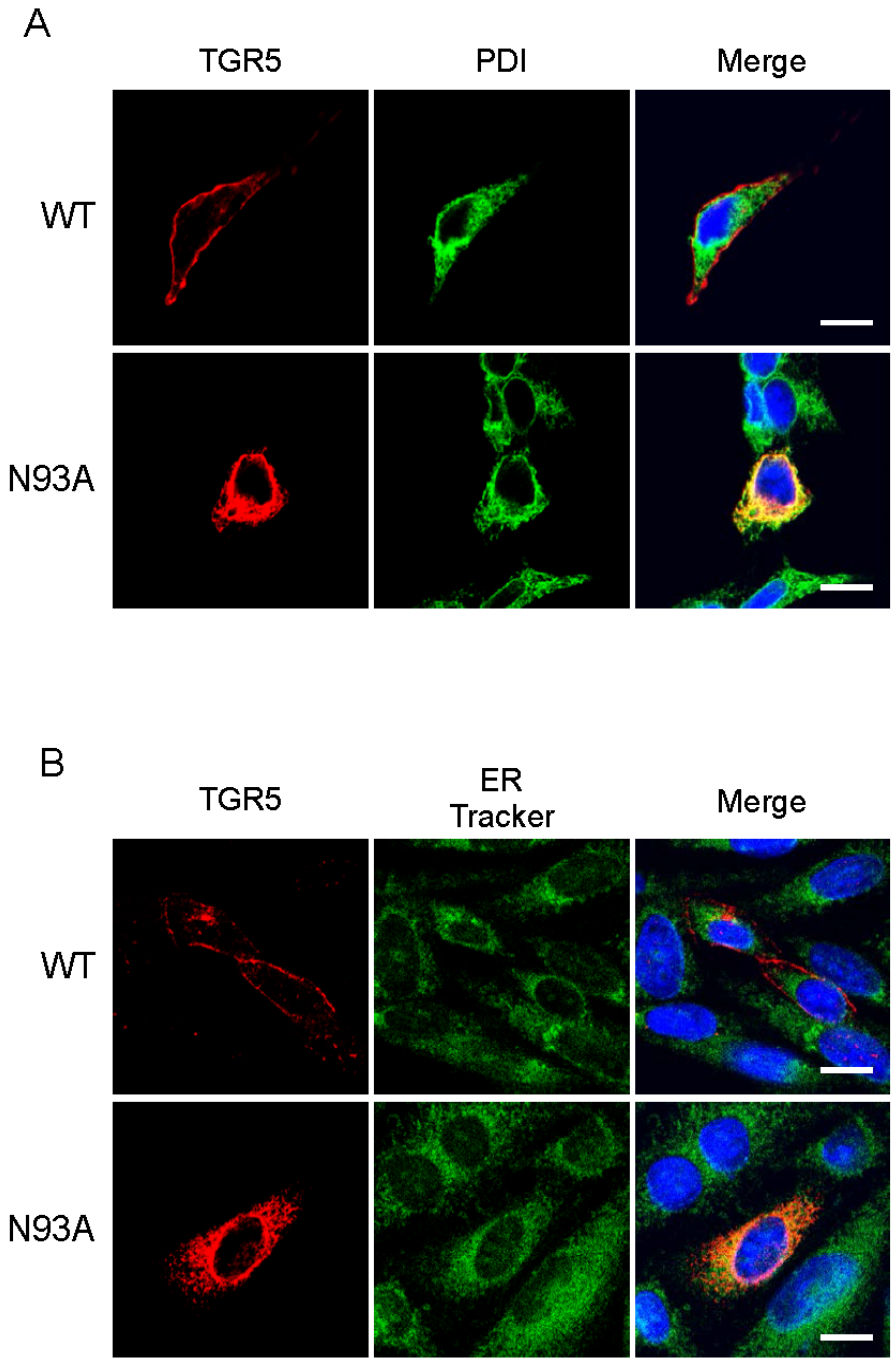


Figure S6. Localization of TGR5 by confocal laser scanning microscopy. HEK293 cells (**A**) and CHO cells (**B**) were transiently transfected with TGR5 wildtype (WT) and the N93A variant in

the pcDNA3.1+ vector. The TGR5 protein was made visible using an anti-TGR5 antibody (RVLR2, red). Compared to the WT protein, which was mainly located in the plasma membrane, the N93A variant was predominately retained intracellularly. Using an antibody against protein disulfide isomerase (PDI, in A) or the ER-Tracker (Invitrogen, in B) N93A fluorescence was colocalized with these markers, indicating that the mutant receptor is retained in the endoplasmic reticulum of both transfected cell lines. Nuclei were stained with Hoechst (in blue). Bars = 10 µm.

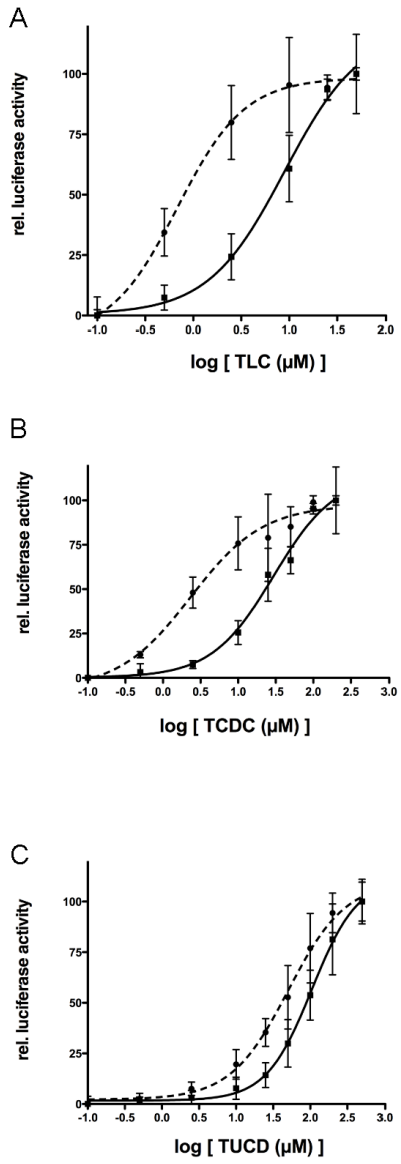


Figure S7. The receptor responsiveness of TGR5 WT (dotted line) and the Y89A$^{3.29}$ variant (straight line) towards TLC (**A**), TCDC (**B**), and TUDC (**C**) was measured using a cAMP responsive luciferase assay. The normalized signal was plotted against the log of the compound concentration. The corresponding EC_{50} values were calculated using the Prism software. Each data point was averaged over 3 to 10 independent measurements.

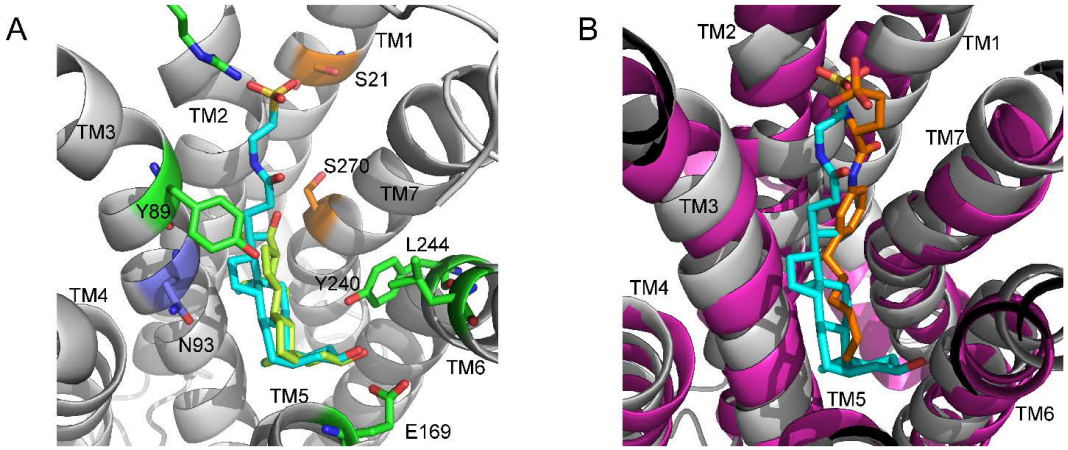


Figure S8. A: Comparison of the binding mode of TLC (cyan) and pregnandiol (limegreen) in TGR5 (grey). **B:** Comparison of the binding mode of TLC (cyan) in TGR5 (grey) and the sphingolipid mimic (orange) in human sphingosine 1-phosphate receptor 1 (S1P1) (magenta) (PDB code: 3V2Y). The sulfonic acid and phosphonate moieties of TLC and the sphingolipid mimic, respectively, occupy the same region, and the hydrophobic scaffolds also exhibit a high positional overlap.

Supplemental References

[1] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck - a Program to Check the Stereochemical Quality of Protein Structures, J. Appl. Crystallogr., 26 (1993) 283-291.

[2] H. Sato, A. Macchiarulo, C. Thomas, A. Gioiello, M. Une, A.F. Hofmann, R. Saladin, K. Schoonjans, R. Pellicciari, J. Auwerx, Novel potent and selective bile acid derivatives as TGR5 agonists: biological screening, structure− activity relationships, and molecular modeling studies, J. Med. Chem., 51 (2008) 1831-1841.

CURRICULUM VITAE

Personal information:

Name: Christoph Gerhard Wilhelm Gertzen

Date of birth: 24.09.1986 in Oberhausen (Germany)

Education:

1997 – 2006 *High school degree*
Freiherr-vom-Stein-Gymnasium, Oberhausen

2006 – 2010 *University degree (Staatsexamen)*
Heinrich-Heine-Universität Düsseldorf, Düsseldorf

2011 – 2012 *University degree (Dipl. Pharm.)*
Christian-Albrechts-Universität zu Kiel, Kiel

2012 *Approbation (Apotheker)*
Düsseldorf

2012 – now *PhD student*
Heinrich-Heine-Universität Düsseldorf, Düsseldorf

Work experience:

2010 – 2011 *Practical year*
ABC-Apotheke, Oberhausen

Peer-reviewed publications:

Gertzen, C.G.W., Gohlke H.; *From Hansch-Fujita analysis to AFMoC: A road to structure-based QSAR*. **Mol Info.** (**2012**), 31, 698–704.
Impact factor reported in 2010: 2.338

[§]Spomer, L., [§]**Gertzen, C.G.W.**, Schmitz, B., Häussinger, D., Gohlke, H., Keitel, V; *A membrane-proximal, C-terminal alpha-helix is required for TGR5 plasma membrane localization and function*. **J Biol Chem** (**2014**), 289, 3689-3702.
Impact factor reported in 2014: 4.573
[§] Both authors contributed equally to this work.

[§]**Gertzen, C.G.W.**, [§]Spomer, L., Smits, S.H.J., Häussinger, D., Keitel, V., Gohlke, H.; *Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists*. **Eur J Med Chem** (**2015**), 104, 57-72.
Impact factor reported in 2014: 3.447

Alchab, F., Ettouati, L., Bouaziz, Z., Bollacke, A., Delcros, J.-G., **Gertzen, C.G.W.**, Gohlke, H., Pinaud, N., Marchivie, M., Guillon, J., Fenet, B., Jose, J., Le Borgne, M.; *Synthesis, biological evaluation, and molecular modeling of substituted indeno[1,2-b]indoles as inhibitors of human protein kinase CK2.* **Pharmaceuticals** (**2015**), 8, 279-302.

Greife A., Felekyan S., Ma Q., **Gertzen C.G.W.**, Spomer L., Dimura M., Peulen T.O., Wöhler C., Häussinger D., Gohlke H., Keitel V., Seidel C.A.M.; *Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modeling study TGR5 in live cells*. **Sci Rep (2016)**, 6, 36792.
Impact factor reported in 2015: 5.228

Diedrich D., Hamacher A., **Gertzen C.G.W.**, Reiss G. J., Alves Avelar L. A., Kurz T., Gohlke H., Kassack M. U., Hansen F. K.; *Rational design and diversity-oriented synthesis of peptoid-based selective HDAC6 inhibitors*. **Chem Commun** (**2016**), 52, 3219 – 3222.
Impact factor reported in 2014: 6.834

Stenzel K., Hamacher A., Hansen F.K., Marek L., **Gertzen C.G.W.**, Senger J., Jung M., Gohlke H., Kassack M. U., Kurz T.; *Alkoxyurea-based Histone Deacetylase (HDAC) Inhibitors with Quinoline Derivates as Cap Group*. **Submitted**.

Syntschewsk V., Hamacher A., **Gertzen C.G.W.**, Senger J., Marek M., Romier C., Kurz T., Jung M., Gohlke H., Kassack M. U., Hansen F. K.; *Design, multicomponent synthesis and anticancer activity of a focused library of HDAC inhibitors with peptoid-based cap groups*. **In preparation.**

[§] Both authors contributed equally to this work.

Other publications:

Koch, A., Ma, Q., Frohnapfel, M., Spomer, L., Keitel, V., **Gertzen, C.G.W.**, Gohlke, H., Seidel, C.A.M.; *High-precision FRET analysis of the G-protein coupled receptor TGR5 in live cells*. **Eur J Med Res** (**2014**), 19(Suppl 1), S12.
Impact factor reported in 2014: 1.495

Gertzen, C.G.W., Spomer, L., Schmitz, B., Häussinger, D., Keitel, V, Gohlke, H.; *The secondary structure of the TGR5 membrane-proximal C-terminus determines plasma membrane localization and responsiveness towards extracellular ligands*. **Eur J Med Res** (**2014**), 19(Suppl 1), S14.
Impact factor reported in 2014: 1.495

Posters:

Gertzen, C.G.W., Spomer, L., Smits, S.H.J., Häussinger, D., Keitel, V., Gohlke, H.; *Binding mode of agonists of the GPCR TGR5*. **GASL Annual Meeting 2016**, Düsseldorf, Germany, January 21 - January 23, 2016.

Gertzen, C.G.W., Spomer, L., Smits, S.H.J., Häussinger, D., Keitel, V., Gohlke, H.; *Binding mode of agonists of the GPCR TGR5*. **DPhG Annual Meeting 2015**, Düsseldorf, Germany, September 23 - September 25, 2015.

Gertzen, C.G.W., Spomer, L., Häussinger, D., Keitel, V., Gohlke, H.; *Binding mode of TGR5 agonists*. **VII FALK GASTRO-CONFERENCE: XXIII International Bile Acid Meeting**, Freiburg i. Br., Germany, October 8. - October 9, 2014.

Gertzen, C.G.W., Spomer, L., Schmitz B., Häussinger, D., Keitel, V., Gohlke, H.; *Why does TGR5 go to the membrane?* **1. International Conference of Collaborative Research Center SFB 974 "Liver Damage and Regeneration"**, Düsseldorf, Germany, November 15 – November 16, 2013.

Gertzen, C.G.W., Spomer, L., Schmitz B., Häussinger, D., Keitel, V., Gohlke, H.; *Why does TGR5 go to the membrane?* **1. International Conference of Collaborative Research Center SFB 974 "Liver Damage and Regeneration"**, Düsseldorf, Germany, November 15 – November 16, 2013.

Gertzen, C.G.W., Spomer, L., Schmitz B., Häussinger, D., Keitel, V., Gohlke, H.; *Why does TGR5 go to the membrane?* **3rd Indo-German conference on Modelling Chemical and Biological (Re) activity (MCBR3)**, Chandigarh, India, February 26 - March 1, 2013.

Gertzen, C.G.W., Spomer, L., Schmitz B., Nisius B., Häussinger, D., Keitel, V., Gohlke, H.;
A helical structure of the N-terminal part of the TGR5 C-terminus is required for receptor membrane localization. **22nd International Bile Acid Meeting (Hepatic and Extrahepatic Targets of Bile Acid Signaling)**, Vienna, Austria, September 14 - September 15, 2012.

Presentations:

An experimentally validated binding mode model of TGR5 agonists. **29th Molecular Modeling Workshop 2015**, Erlangen, Germany, March 9 - March 11, 2015.

Why Does the GPCR TGR5 Go to the Plasma Membrane? **COMPUTER SIMULATION AND THEORY OF MACROMOLECULES 2014**, Hünfeld, Germany, April 11 - April 12, 2014.

Awards:

2nd Lecture Award

An experimentally validated binding mode model of TGR5 agonists. **29th Molecular Modeling Workshop 2015**, Erlangen, Germany, March 9 - March 11, 2015.

REFERENCES

1. Bjarnadóttir, T. K.; Gloriam, D. E.; Hellstrand, S. H.; Kristiansson, H.; Fredriksson, R.; Schiöth, H. B. Comprehensive Repertoire and Phylogenetic Analysis of the G Protein-Coupled Receptors in Human and Mouse. *Genomics* **2006,** 88, 263-273.

2. Vassilatis, D. K.; Hohmann, J. G.; Zeng, H.; Li, F.; Ranchalis, J. E.; Mortrud, M. T.; Brown, A.; Rodriguez, S. S.; Weller, J. R.; Wright, A. C.; Bergmann, J. E.; Gaitanaris, G. A. The G Protein-Coupled Receptor Repertoires of Human and Mouse. *P Natl Acad Sci* **2003,** 100, 4903-4908.

3. Richelson, E.; Nelson, A. Antagonism by Antidepressants of Neurotransmitter Receptors of Normal Human Brain in Vitro. *J Pharmacol Exp Ther* **1984,** 230, 94-102.

4. Kanba, S.; Suzuki, E.; Nomura, S.; Nakaki, T.; Yagi, G.; Asai, M.; Richelson, E. Affinity of Neuroleptics for D1 Receptor of Human Brain Striatum. *J Neuropsychiatry Clin Neurosci* **1994,** 19, 265-269.

5. Fukui, H.; Fujimoto, K.; Mizuguchi, H.; Sakamoto, K.; Horio, Y.; Takai, S.; Yamada, K.; Ito, S. Molecular Cloning of the Human Histamine H1 Receptor Gene. *Biochem Bioph Res Co* **1994,** 201, 894-901.

6. Ruat, M.; Traiffort, E.; Bouthenet, M. L.; Schwartz, J. C.; Hirschfeld, J.; Buschauer, A.; Schunack, W. Reversible and Irreversible Labeling and Autoradiographic Localization of the Cerebral Histamine H2 Receptor Using [125I]Iodinated Probes. *P Natl Acad Sci* **1990,** 87, 1658-1662.

7. Osterberg, G. Topography of the Layer of Rod and Cones in the Human Retina. *Acta Ophthalmol, Kbh, Suppl* **1935,** 6, 103.

8. Basinger, S.; Bok, D.; Hall, M. Rhodopsin in the Rod Outer Segment Plasma Membrane. *J Cell Biol* **1976,** 69, 29-42.

9. Barnes, P. J. Distribution of Receptor Targets in the Lung. *Pro Am Thorac Soc* **2004,** 1, 345-351.

10. Gugger, M.; White, R.; Song, S.; Waser, B.; Cescato, R.; Rivi're, P.; Reubi, J. C. GPR87 Is an Overexpressed G-Protein Coupled Receptor in Squamous Cell Carcinoma of the Lung. *Disease Markers* **2008,** 24.

11. Foster, S. R.; Porrello, E. R.; Purdue, B.; Chan, H.-W.; Voigt, A.; Frenzel, S.; Hannan, R. D.; Moritz, K. M.; Simmons, D. G.; Molenaar, P.; Roura, E.; Boehm, U.; Meyerhof, W.; Thomas, W. G. Expression, Regulation and Putative Nutrient-Sensing Function of Taste GPCRs in the Heart. *PLoS ONE* **2013,** 8, e64579.

12. Tang, C.-M.; Insel, P. A. GPCR Expression in the Heart: "New" Receptors in Myocytes and Fibroblasts. *Trends Cardiovasc Med* **2004,** 14, 94-99.

13. Lautrette, A.; Li, S.; Alili, R.; Sunnarborg, S. W.; Burtin, M.; Lee, D. C.; Friedlander, G.; Terzi, F. Angiotensin II and EGF Receptor Cross-Talk in Chronic Kidney Diseases: A New Therapeutic Approach. *Nat Med* **2005,** 11, 867-874.

14. Brown, D.; Breton, S.; Ausiello, D. A.; Marshansky, V. Sensing, Signaling and Sorting Events in Kidney Epithelial Cell Physiology. *Traffic* **2009,** 10, 275-284.

15. Dyer, J.; Daly, K.; Salmon, K.; Arora, D.; Kokrashvili, Z.; Margolskee, R.; Shirazi-Beechey, S. Intestinal Glucose Sensing and Regulation of Intestinal Glucose Absorption. *Biochem Soc T* **2007,** 35, 1191-1194.

16. Bohn, L. M.; Raehal, K. M. Opioid Receptor Signaling: Relevance for Gastrointestinal Therapy. *Curr Opin Phrarmacol* **2006,** 6, 559-563.

17. Shahbazian, A.; Heinemann, A.; Schmidhammer, H.; Beubler, E.; Holzer-Petsche, U.; Holzer, P. Involvement of M- and K-, but Not Δ-, Opioid Receptors in the Peristaltic Motor

Depression Caused by Endogenous and Exogenous Opioids in the Guinea-Pig Intestine. *Brit J Pharmacol* **2002,** 135, 741-750.

18. Katritch, V.; Cherezov, V.; Stevens, R. C. Structure-Function of the G-Protein-Coupled Receptor Superfamily. *Annu Rev Pharmacol Toxicol* **2013,** 53, 531.

19. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. Opinion - How Many Drug Targets Are There? *Nat Rev Drug Discov* **2006,** 5, 993-996.

20. Tyndall, J. D.; Sandilya, R. GPCR Agonists and Antagonists in the Clinic. *Med Chem* **2005,** 1, 405-421.

21. Rask-Andersen, M.; Almén, M. S.; Schiöth, H. B. Trends in the Exploitation of Novel Drug Targets. *Nat Rev Drug Discov* **2011,** 10, 579-590.

22. Goodman, M. Market Watch: Sales of Biologics to Show Robust Growth through to 2013. *Nat Rev Drug Discov* **2009,** 8, 837-837.

23. Lawrence, S. Biotech Drug Market Steadily Expands. *Nat Biotech* **2005,** 23, 1466-1466.

24. Weber, F.; Sedelmeier, G. Top 200 Pharmawirkstoffe. *Angew Chem* **2014,** 62, 997-997.

25. Raiteri, M. Functional Pharmacology in Human Brain. *Pharmacol Rev* **2006,** 58, 162-193.

26. EllisClare. The State of GPCR Research in 2004. *Nat Rev Drug Discov* **2004,** 3, 577-626.

27. Lappano, R.; Maggiolini, M. G Protein-Coupled Receptors: Novel Targets for Drug Discovery in Cancer. *Nat Rev Drug Discov* **2011,** 10, 47-60.

28. Pierce, K. L.; Premont, R. T.; Lefkowitz, R. J. Seven-Transmembrane Receptors. *Nat Rev Mol Cell Biol* **2002,** 3, 639-650.

29. Pin, J.-P.; Galvez, T.; Prézeau, L. Evolution, Structure, and Activation Mechanism of Family 3/C G-Protein-Coupled Receptors. *Pharmacol Ther* **2003,** 98, 325-354.

30. Pal, K.; Melcher, K.; Xu, H. E. Structure and Mechanism for Recognition of Peptide Hormones by Class B G-Protein-Coupled Receptors. *Acta Pharmacol Sin* **2012,** 33, 300-311.

31. Siu, F. Y.; He, M.; de Graaf, C.; Han, G. W.; Yang, D.; Zhang, Z.; Zhou, C.; Xu, Q.; Wacker, D.; Joseph, J. S.; Liu, W.; Lau, J.; Cherezov, V.; Katritch, V.; Wang, M.-W.; Stevens, R. C. Structure of the Human Glucagon Class B G-Protein-Coupled Receptor. *Nature* **2013,** 499, 444-449.

32. Hollenstein, K.; Kean, J.; Bortolato, A.; Cheng, R. K. Y.; Dore, A. S.; Jazayeri, A.; Cooke, R. M.; Weir, M.; Marshall, F. H. Structure of Class B GPCR Corticotropin-Releasing Factor Receptor 1. *Nature* **2013,** 499, 438-443.

33. Chun, L.; Zhang, W.-h.; Liu, J.-f. Structure and Ligand Recognition of Class C GPCRs. *Acta Pharmacol Sin* **2012,** 33, 312-323.

34. Spehr, M.; Gisselmann, G.; Poplawski, A.; Riffell, J. A.; Wetzel, C. H.; Zimmer, R. K.; Hatt, H. Identification of a Testicular Odorant Receptor Mediating Human Sperm Chemotaxis. *Science* **2003,** 299, 2054-2058.

35. Neuhaus, E. M.; Zhang, W.; Gelis, L.; Deng, Y.; Noldus, J.; Hatt, H. Activation of an Olfactory Receptor Inhibits Proliferation of Prostate Cancer Cells. *J Biol Chem* **2009,** 284, 16218-16225.

36. Thomas, C.; Gioiello, A.; Noriega, L.; Strehle, A.; Oury, J.; Rizzo, G.; Macchiarulo, A.; Yamamoto, H.; Mataki, C.; Pruzanski, M.; Pellicciari, R.; Auwerx, J.; Schoonjans, K. TGR5-Mediated Bile Acid Sensing Controls Glucose Homeostasis. *Cell Metab* **2009,** 10, 167-177.

37. Alemi, F.; Poole, D. P.; Chiu, J.; Schoonjans, K.; Cattaruzza, F.; Grider, J. R.; Bunnett, N. W.; Corvera, C. U. The Receptor TGR5 Mediates the Prokinetic Actions of Intestinal Bile Acids and Is Required for Normal Defecation in Mice. *Gastroenterology* **2013,** 144, 145-154.

38. Kawamata, Y.; Fujii, R.; Hosoya, M.; Harada, M.; Yoshida, H.; Miwa, M.; Fukusumi, S.; Habata, Y.; Itoh, T.; Shintani, Y. A G Protein-Coupled Receptor Responsive to Bile Acids. *J Biol Chem* **2003,** 278, 9435-9440.

39. Maruyama, T.; Miyamoto, Y.; Nakamura, T.; Tamai, Y.; Okada, H.; Sugiyama, E.; Nakamura, T.; Itadani, H.; Tanaka, K. Identification of Membrane-Type Receptor for Bile Acids (M-BAR). *Biochem Bioph Res Co* **2002,** 298, 714-719.

40. Maruyama, T.; Tanaka, K.; Suzuki, J.; Miyoshi, H.; Harada, N.; Nakamura, T.; Miyamoto, Y.; Kanatani, A.; Tamai, Y. Targeted Disruption of G Protein-Coupled Bile Acid Receptor 1 (GPBAR1/M-BAR) in Mice. *J Endocrinol* **2006,** 191, 197-205.

41. Vassileva, G.; Golovko, A.; Markowitz, L.; Abbondanzo, S.; Zeng, M.; Yang, S.; Hoos, L.; Tetzloff, G.; Levitan, D.; Murgolo, N. Targeted Deletion of GPBAR1 Protects Mice from Cholesterol Gallstone Formation. *Biochem J* **2006,** 398, 423-430.

42. Spomer, L.; Gertzen, C. G. W.; Schmitz, B.; Häussinger, D.; Gohlke, H.; Keitel, V. A Membrane-Proximal, C-Terminal A-Helix Is Required for Plasma Membrane Localization and Function of the G Protein-Coupled Receptor (GPCR) TGR5. *J Biol Chem* **2014,** 289, 3689-3702.

43. Keitel, V.; Donner, M.; Winandy, S.; Kubitz, R.; Häussinger, D. Expression and Function of the Bile Acid Receptor TGR5 in Kupffer Cells. *Biochem Bioph Res Co* **2008,** 372, 78-84.

44. Keitel, V.; Reinehr, R.; Gatsios, P.; Rupprecht, C.; Görg, B.; Selbach, O.; Häussinger, D.; Kubitz, R. The G-Protein Coupled Bile Salt Receptor TGR5 Is Expressed in Liver Sinusoidal Endothelial Cells. *Hepatology* **2007,** 45, 695-704.

45. Keitel, V.; Cupisti, K.; Ullmer, C.; Knoefel, W. T.; Kubitz, R.; Häussinger, D. The Membrane-Bound Bile Acid Receptor TGR5 Is Localized in the Epithelium of Human Gallbladders. *Hepatology* **2009,** 50, 861-870.

46. Keitel, V.; Görg, B.; Bidmon, H. J.; Zemtsova, I.; Spomer, L.; Zilles, K.; Häussinger, D. The Bile Acid Receptor TGR5 (GPBAR-1) Acts as a Neurosteroid Receptor in Brain. *Glia* **2010,** 58, 1794-1805.

47. Perino, A.; Pols, T. W. H.; Nomura, M.; Stein, S.; Pellicciari, R.; Schoonjans, K. TGR5 Reduces Macrophage Migration through Mtor-Induced C/EBPβ Differential Translation. *J Clin Invest* **2014,** 124, 5424-5436.

48. Keitel, V.; Cupisti, K.; Ullmer, C.; Knoefel, W. T.; Kubitz, R.; Haeussinger, D. The Membrane-Bound Bile Acid Receptor TGR5 Is Localized in the Epithelium of Human Gallbladders. *Hepatology* **2009,** 50, 861-870.

49. Keitel, V.; Häussinger, D. Perspective: TGR5 (GPBAR-1) in Liver Physiology and Disease. *Clin Res Hepatol Gastroenterol* **2012,** 36, 412-419.

50. Kusuhara, H.; Sugiyama, Y. Efflux Transport Systems for Drugs at the Blood–Brain Barrier and Blood–Cerebrospinal Fluid Barrier (Part 1). *Drug Discov Today* **2001,** 6, 150-156.

51. Kitazawa, T.; Terasaki, T.; Suzuki, H.; Kakee, A.; Sugiyama, Y. Efflux of Taurocholic Acid across the Blood-Brain Barrier: Interaction with Cyclic Peptides. *J Pharmacol Exp Ther* **1998,** 286, 890-895.

52. Pols, T. W. H.; Noriega, L. G.; Nomura, M.; Auwerx, J.; Schoonjans, K. The Bile Acid Membrane Receptor TGR5 as an Emerging Target in Metabolism and Inflammation. *J Hepatol* **2011,** 54, 1263-1272.

53. Pols, T. W. H.; Nomura, M.; Harach, T.; Lo Sasso, G.; Oosterveer, Maaike H.; Thomas, C.; Rizzo, G.; Gioiello, A.; Adorini, L.; Pellicciari, R.; Auwerx, J.; Schoonjans, K. TGR5 Activation Inhibits Atherosclerosis by Reducing Macrophage Inflammation and Lipid Loading. *Cell Metab* **2011,** 14, 747-757.

54. Calmus, Y.; Poupon, R. Shaping Macrophages Function and Innate Immunity by Bile Acids: Mechanisms and Implication in Cholestatic Liver Diseases. *Clin Res Hepatol Gastroenterol* **2014,** 38, 550-556.

55. Li, T.; Chiang, J. Y. L. Bile Acid Signaling in Metabolic Disease and Drug Therapy. *Pharmacol Rev* **2014,** 66, 948-983.

56. Duboc, H.; Tache, Y.; Hofmann, A. F. The Bile Acid TGR5 Membrane Receptor: From Basic Research to Clinical Application. *Dig Liver Dis* **2014,** 46, 302-312.

57. Li, T.; Holmstrom, S. R.; Kir, S.; Umetani, M.; Schmidt, D. R.; Kliewer, S. A.; Mangelsdorf, D. J. The G Protein-Coupled Bile Acid Receptor, TGR5, Stimulates Gallbladder Filling. *Mol Endocrinol* **2011,** 25, 1066-1071.

58. Keitel, V.; Reinehr, R.; Reich, M.; Sommerfeld, A.; Cupisti, K.; Knoefel, W. T.; Häussinger, D. TGR5 (GPBAR-1) Is Expressed in Cholangiocarcinomas and Confers Apopotosis Resistance in Isolated Cholangiocytes. *Z Gastroenterol* **2012,** 50, P5_24.

59. Hong, J.; Behar, J.; Wands, J.; Resnick, M.; Wang, L. J.; DeLellis, R. A.; Lambeth, D.; Souza, R. F.; Spechler, S. J.; Cao, W. Role of a Novel Bile Acid Receptor TGR5 in the Development of Oesophageal Adenocarcinoma. *Gut* **2010,** 59, 170-180.

60. Cao, W.; Tian, W.; Hong, J.; Li, D.; Tavares, R.; Noble, L.; Moss, S. F.; Resnick, M. B. Expression of Bile Acid Receptor TGR5 in Gastric Adenocarcinoma. *Am J Physiol Gastrointest Liver Physiol* **2013,** 304, G322-327.

61. Anwer, M. S. Intracellular Signaling by Bile Acids. *J Biosci (Rajshari)* **2012,** 20, 1-23.

62. Lieu, T.; Jayaweera, G.; Zhao, P.; Poole, D. P.; Jensen, D.; Grace, M.; McIntyre, P.; Bron, R.; Wilson, Y. M.; Krappitz, M.; Haerteis, S.; Korbmacher, C.; Steinhoff, M. S.; Nassini, R.; Materazzi, S.; Geppetti, P.; Corvera, C. U.; Bunnett, N. W. The Bile Acid Receptor TGR5 Activates the TRPA1 Channel to Induce Itch in Mice. *Gastroenterology* **2014,** 147, 1417-1428.

63. Alemi, F.; Kwon, E.; Poole, D. P.; Lieu, T.; Lyo, V.; Cattaruzza, F.; Cevikbas, F.; Steinhoff, M.; Nassini, R.; Materazzi, S.; Guerrero-Alba, R.; Valdez-Morales, E.; Cottrell, G. S.; Schoonjans, K.; Geppetti, P.; Vanner, S. J.; Bunnett, N. W.; Corvera, C. U. The TGR5 Receptor Mediates Bile Acid-Induced Itch and Analgesia. *J Clin Invest* **2013,** 123, 1513-1530.

64. Zou, Q.; Duan, H.; Ning, M.; Liu, J.; Feng, Y.; Zhang, L.; Zhu, J.; Leng, Y.; Shen, J. 4-Benzofuranyloxynicotinamide Derivatives Are Novel Potent and Orally Available TGR5 Agonists. *Eur J Med Chem* **2014,** 82, 1-15.

65. Sato, H.; Macchiarulo, A.; Thomas, C.; Gioiello, A.; Une, M.; Hofmann, A. F.; Saladin, R.; Schoonjans, K.; Pellicciari, R.; Auwerx, J. Novel Potent and Selective Bile Acid Derivatives as TGR5 Agonists: Biological Screening, Structure− Activity Relationships, and Molecular Modeling Studies. *J Med Chem* **2008,** 51, 1831-1841.

66. Macchiarulo, A.; Gioiello, A.; Thomas, C.; Pols, T. W.; Nuti, R.; Ferrari, C.; Giacchè, N.; De Franco, F.; Pruzanski, M.; Auwerx, J. Probing the Binding Site of Bile Acids in TGR5. *ACS Med Chem Lett* **2013,** 4, 1158-1162.

67. Gioiello, A.; Rosatelli, E.; Nuti, R.; Macchiarulo, A.; Pellicciari, R. Patented TGR5 Modulators: A Review (2006 – Present). *Expert Opin Ther Pat* **2012,** 22, 1399-1414.

68. Pellicciari, R.; Gioiello, A.; Macchiarulo, A.; Thomas, C.; Rosatelli, E.; Natalini, B.; Sardella, R.; Pruzanski, M.; Roda, A.; Pastorini, E.; Schoonjans, K.; Auwerx, J. Discovery of 6α-Ethyl-23(S)-Methylcholic Acid (S-EMCA, Int-777) as a Potent and Selective Agonist for the TGR5 Receptor, a Novel Target for Diabesity. *J Med Chem* **2009,** 52, 7958-7961.

69. Herbert, M. R.; Siegel, D. L.; Staszewski, L.; Cayanan, C.; Banerjee, U.; Dhamija, S.; Anderson, J.; Fan, A.; Wang, L.; Rix, P.; Shiau, A. K.; Rao, T. S.; Noble, S. A.; Heyman, R. A.; Bischoff, E.; Guha, M.; Kabakibi, A.; Pinkerton, A. B. Synthesis and SAR of 2-Aryl-3-Aminomethylquinolines as Agonists of the Bile Acid Receptor TGR5. *Bioorg Med Chem Lett* **2010,** 20, 5718-5721.

70. Budzik, B. W.; Evans, K. A.; Wisnoski, D. D.; Jin, J.; Rivero, R. A.; Szewczyk, G. R.; Jayawickreme, C.; Moncol, D. L.; Yu, H. Synthesis and Structure–Activity Relationships of a Series of 3-Aryl-4-Isoxazolecarboxamides as a New Class of TGR5 Agonists. *Bioorg Med Chem Lett* **2010,** 20, 1363-1367.

71. D'Amore, C.; Di Leva, F. S.; Sepe, V.; Renga, B.; Del Gaudio, C.; D'Auria, M. V.; Zampella, A.; Fiorucci, S.; Limongelli, V. Design, Synthesis, and Biological Evaluation of Potent Dual Agonists of Nuclear and Membrane Bile Acid Receptors. *J Med Chem* **2014,** 57, 937-954.

72. Genet, C.; Strehle, A.; Schmidt, C.; Boudjelal, G.; Lobstein, A.; Schoonjans, K.; Souchet, M.; Auwerx, J.; Saladin, R.; Wagner, A. Structure− Activity Relationship Study of Betulinic Acid, a Novel and Selective TGR5 Agonist, and Its Synthetic Derivatives: Potential Impact in Diabetes. *J Med Chem* **2009,** 53, 178-190.

73. Pellicciari, R.; Gioiello, A.; Sabbatini, P.; Venturoni, F.; Nuti, R.; Colliva, C.; Rizzo, G.; Adorini, L.; Pruzanski, M.; Roda, A. Avicholic Acid: A Lead Compound from Birds on the Route to Potent TGR5 Modulators. *ACS Med Chem Lett* **2012,** 3, 273-277.

74. Donna, D. Y.; Sousa, K. M.; Mattern, D. L.; Wagner, J.; Fu, X.; Vaidehi, N.; Forman, B. M.; Huang, W. Stereoselective Synthesis, Biological Evaluation, and Modeling of Novel Bile Acid-Derived G-Protein Coupled Bile Acid Receptor 1 (GPBAR1, TGR5) Agonists. *Bioorg Med Chem* **2015,** 23, 1613-1628.

75. Pellicciari, R.; Sato, H.; Gioiello, A.; Costantino, G.; Macchiarulo, A.; Sadeghpour, B. M.; Giorgi, G.; Schoonjans, K.; Auwerx, J. Nongenomic Actions of Bile Acids. Synthesis and Preliminary Characterization of 23- and 6,23-Alkyl-Substituted Bile Acid Derivatives as Selective Modulators for the G-Protein Coupled Receptor TGR5. *J Med Chem* **2007,** 50, 4265-4268.

76. Gertzen, C. G. W.; Spomer, L.; Smits, S. H.; Haussinger, D.; Keitel, V.; Gohlke, H. Mutational Mapping of the Transmembrane Binding Site of the G-Protein Coupled Receptor TGR5 and Binding Mode Prediction of TGR5 Agonists. *Eur J Med Chem* **2015,** 104, 57-72.

77. Illergård, K.; Ardell, D. H.; Elofsson, A. Structure Is Three to Ten Times More Conserved Than Sequence—a Study of Structural Response in Protein Cores. *Proteins: Struct, Funct, Bioinf* **2009,** 77, 499-508.

78. Pierce, K. L.; Premont, R. T.; Lefkowitz, R. J. Seven-Transmembrane Receptors. *Nat Rev Mol Cell Biol* **2002,** 3, 639-650.

79. Michino, M.; Abola, E.; Brooks, C. L.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-Wide Assessment of GPCR Structure Modelling and Ligand Docking: GPCR Dock 2008. *Nat Rev Drug Discov* **2009,** 8, 455-463.

80. Kufareva, I.; Katritch, V.; Stevens, Raymond C.; Abagyan, R. Advances in GPCR Modeling Evaluated by the GPCR Dock 2013 Assessment: Meeting New Challenges. *Structure* **2014,** 22, 1120-1139.

81. Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S. A Critical Assessment of Docking Programs and Scoring Functions. *J Med Chem* **2006,** 49, 5912-5931.

82. Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions. *Proteins: Struct, Funct, Bioinf* **2002,** 47, 409-443.

83. Korb, O.; Stutzle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein− Ligand Docking with Plants. *J Chem Inf Model* **2009,** 49, 84-96.

84. Huang, S.-Y.; Grinter, S. Z.; Zou, X. Scoring Functions and Their Evaluation Methods for Protein–Ligand Docking: Recent Advances and Future Directions. *Phys Chem Chem Phys* **2010,** 12, 12899-12908.

85. Sotriffer, C. A.; Gohlke, H.; Klebe, G. Docking into Knowledge-Based Potential Fields: A Comparative Evaluation of DrugScore. *J Med Chem* **2002,** 45, 1967-1970.

86. Gohlke, H.; Klebe, G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew Chem, Int Ed* **2002,** 41, 2644-2676.

87. Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R.; Participants, G. D. Status of GPCR Modeling and Docking as Reflected by Community-Wide GPCR Dock 2010 Assessment. *Structure* **2011,** 19, 1108-1126.

88. Kufareva, I.; Katritch, V.; Stevens, Raymond C.; Abagyan, R. Advances in GPCR Modeling Evaluated by the GPCR Dock 2013 Assessment: Meeting New Challenges. *Structure* 22, 1120-1139.

89. Cavasotto, C. N.; Orry, A. J. W.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; O'Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J. Discovery of Novel Chemotypes to a G-Protein-Coupled Receptor through Ligand-Steered Homology Modeling and Structure-Based Virtual Screening. *J Med Chem* **2008,** 51, 581-588.

90. Evers, A.; Klebe, G. Successful Virtual Screening for a Submicromolar Antagonist of the Neurokinin-1 Receptor Based on a Ligand-Supported Homology Model. *J Med Chem* **2004,** 47, 5381-5392.

91. Evers, A.; Klabunde, T. Structure-Based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the α1a Adrenergic Receptor. *J Med Chem* **2005,** 48, 1088-1097.

92. Ananthan, S.; Zhang, W.; Hobrath, J. V. Recent Advances in Structure-Based Virtual Screening of G-Protein Coupled Receptors. *The AAPS Journal* **2009,** 11, 178-185.

93. Cozzini, P.; Kellogg, G. E.; Spyrakis, F.; Abraham, D. J.; Costantino, G.; Emerson, A.; Fanelli, F.; Gohlke, H.; Kuhn, L. A.; Morris, G. M.; Orozco, M.; Pertinhez, T. A.; Rizzi, M.; Sotriffer, C. A. Target Flexibility: An Emerging Consideration in Drug Discovery and Design†. *J Med Chem* **2008,** 51, 6237-6255.

94. Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of Mm-Pbsa in Reproducing the Binding Free Energies to Hiv-1 Rt of Tibo Derivatives and Predicting the Binding Mode to Hiv-1 Rt of Efavirenz by Docking and Mm-Pbsa. *J Am Chem Soc* **2001,** 123, 5221-5230.

95. Koch, A.; Ma, Q.; Frohnapfel, M.; Spomer, L.; Keitel-Anselmino, V.; Gertzen, C.; Gohlke, H.; Seidel, C. A. High-Precision FRET Analysis of the G-Protein Coupled Receptor TGR5 in Live Cells. *Eur J Med Res* **2014,** 19, 1-2.

96. Hannah, M. TGR5 Ligands as Potential Therapeutics in Inflammatory Diseases. *Int J Interferon Cytokine Mediat Res* **2014,** 6, 27-38.

97. Perino, A.; Schoonjans, K. TGR5 and Immunometabolism: Insights from Physiology and Pharmacology. *Trends Pharmacol Sci*.

98. Stepanov, V.; Stankov, K.; Mikov, M. The Bile Acid Membrane Receptor TGR5: A Novel Pharmacological Target in Metabolic, Inflammatory and Neoplastic Disorders. *J Recept Signal Transduct* **2013,** 33, 213-223.

99. Kuipers, F.; Bloks, V. W.; Groen, A. K. Beyond Intestinal Soap - Bile Acids in Metabolic Control. *Nat Rev Endocrinol* **2014,** 10, 488–498.

100. Kawamata, Y.; Fujii, R.; Hosoya, M.; Harada, M.; Yoshida, H.; Miwa, M.; Fukusumi, S.; Habata, Y.; Itoh, T.; Shintani, Y.; Hinuma, S.; Fujisawa, Y.; Fujino, M. A G Protein-Coupled Receptor Responsive to Bile Acids. *J Biol Chem* **2003,** 278, 9435-9440.

101. Consortium, U. The Universal Protein Resource (Uniprot). *Nucleic Acids Res* **2008,** 36, D190-D195.

102. Consortium, U. Reorganizing the Protein Space at the Universal Protein Resource (Uniprot). *Nucleic Acids Res* **2011**, gkr981.

103. G-Protein-Coupled Receptors Database (GPCRDB). www.gpcr.org/7tm/

104. Horn, F.; Weare, J.; Beukers, M. W.; Horsch, S.; Bairoch, A.; Chen, W.; Edvardsen, O.; Campagne, F.; Vriend, G. GPCRDB: An Information System for G Protein-Coupled Receptors. *Nucleic Acids Res* **1998,** 26, 275-279.

105. Isberg, V.; Vroling, B.; van der Kant, R.; Li, K.; Vriend, G.; Gloriam, D. GPCRDB: An Information System for G Protein-Coupled Receptors. *Nucleic Acids Res* **2014,** 42, D422-D425.

106. Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the Cxcr4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **2010,** 330, 1066-1071.

107. Duvernay, M. T.; Zhou, F.; Wu, G. A Conserved Motif for the Transport of G Protein-Coupled Receptors from the Endoplasmic Reticulum to the Cell Surface. *J Biol Chem* **2004,** 279, 30741-30750.

108. Seachrist, J. L.; Ferguson, S. S. Regulation of G Protein-Coupled Receptor Endocytosis and Trafficking by Rab GTPases. *Life sciences* **2003,** 74, 225-235.

109. Di Leva, F. S.; Festa, C.; Renga, B.; Sepe, V.; Novellino, E.; Fiorucci, S.; Zampella, A.; Limongelli, V. Structure-Based Drug Design Targeting the Cell Membrane Receptor GPBAR1: Exploiting the Bile Acid Scaffold Towards Selective Agonism. *Sci Rep* **2015,** 5, 16605.

110. Zhou, X.; Li, D.; Resnick, M. B.; Wands, J.; Cao, W. NADPH Oxidase NOX5-S and Nuclear Factor Kb1 Mediate Acid-Induced Microsomal Prostaglandin E Synthase-1 Expression in Barrett's Esophageal Adenocarcinoma Cells. *Mol Pharmacol* **2013,** 83, 978-990.

111. Yasuda, H.; Hirata, S.; Inoue, K.; Mashima, H.; Ohnishi, H.; Yoshiba, M. Involvement of Membrane-Type Bile Acid Receptor M-Bar/TGR5 in Bile Acid-Induced Activation of Epidermal Growth Factor Receptor and Mitogen-Activated Protein Kinases in Gastric Carcinoma Cells. *Biochem Bioph Res Co* **2007,** 354, 154-159.

112. Watanabe, M.; Houten, S. M.; Mataki, C.; Christoffolete, M. A.; Kim, B. W.; Sato, H.; Messaddeq, N.; Harney, J. W.; Ezaki, O.; Kodama, T.; Schoonjans, K.; Bianco, A. C.; Auwerx, J. Bile Acids Induce Energy Expenditure by Promoting Intracellular Thyroid Hormone Activation. *Nature* **2006,** 439, 484-489.

113. Lavoie, B.; Balemba, O. B.; Godfrey, C.; Watson, C. A.; Vassileva, G.; Corvera, C. U.; Nelson, M. T.; Mawe, G. M. Hydrophobic Bile Salts Inhibit Gallbladder Smooth Muscle Function Via Stimulation of GPBAR1 Receptors and Activation of Katp Channels. *J Physiol* **2010,** 588, 3295-3305.

114. Reich, M.; Deutschmann, K.; Sommerfeld, A.; Klindt, C.; Kluge, S.; Kubitz, R.; Ullmer, C.; Knoefel, W. T.; Herebian, D.; Mayatepek, E.; Häussinger, D.; Keitel, V. TGR5 Is Essential for Bile Acid-Dependent Cholangiocyte Proliferation in Vivo and in Vitro. *Gut* **2015**.

115. Kida, T.; Tsubosaka, Y.; Hori, M.; Ozaki, H.; Murata, T. Bile Acid Receptor TGR5 Agonism Induces NO Production and Reduces Monocyte Adhesion in Vascular Endothelial Cells. *Arteriosclerosis, thrombosis, and vascular biology* **2013,** 33, 1663-1669.

116. Svensson, P.-A.; Olsson, M.; Andersson-Assarsson, J. C.; Taube, M.; Pereira, M. J.; Froguel, P.; Jacobson, P. The TGR5 Gene Is Expressed in Human Subcutaneous Adipose Tissue and Is Associated with Obesity, Weight Loss and Resting Metabolic Rate. *Biochem Bioph Res Co* **2013,** 433, 563-566.

117. Suzuki, T.; Aoyama, J.; Hashimoto, M.; Ohara, M.; Futami-Suda, S.; Suzuki, K.; Ouchi, M.; Igari, Y.; Watanabe, K.; Nakano, H. Correlation between Postprandial Bile Acids and Body Fat Mass in Healthy Normal-Weight Subjects. *Clin Biochem* **2014,** 47, 1128-1131.

118. Sakanaka, T.; Inoue, T.; Yorifuji, N.; Iguchi, M.; Fujiwara, K.; Narabayashi, K.; Kakimoto, K.; Nouda, S.; Okada, T.; Kuramoto, T.; Ishida, K.; Abe, Y.; Takeuchi, T.; Umegaki, E.; Akiba, Y.; Kaunitz, J. D.; Higuchi, K. The Effects of a TGR5 Agonist and a Dipeptidyl Peptidase IV Inhibitor on Dextran Sulfate Sodium-Induced Colitis in Mice. *J Gastroenterol Hepatol* **2015,** 30, 60-65.

119. Skibicka, K. P. The Central Glp-1: Implications for Food and Drug Reward. *Front Neurosci* **2013,** 7.

120. Willms, B.; Werner, J.; Holst, J. J.; Orskov, C.; Creutzfeldt, W.; Nauck, M. A. Gastric Emptying, Glucose Responses, and Insulin Secretion after a Liquid Test Meal: Effects of Exogenous Glucagon-Like Peptide-1 (GLP-1)-(7-36) Amide in Type 2 (Noninsulin-Dependent) Diabetic Patients. *J Clin Endocrinol Metabol* **1996,** 81, 327-332.

121. Han, V. K. M.; Hynes, M. A.; Jin, C.; Towle, A. C.; Lauder, J. M.; Lund, P. K. Cellular Localization of Proglucagon/Glucagon-Like Peptide I Messenger Rnas in Rat Brain. *J Neurosci Res* **1986,** 16, 97-107.

122. Kreymann, B.; Ghatei, M. A.; Williams, G.; Bloom, S. R. Glucagon-Like Peptide-1 7-36: A Physiological Incretin in Man. *The Lancet* **1987,** 330, 1300-1304.

123. Ichikawa, R.; Takayama, T.; Yoneno, K.; Kamada, N.; Kitazume, M. T.; Higuchi, H.; Matsuoka, K.; Watanabe, M.; Itoh, H.; Kanai, T. Bile Acids Induce Monocyte Differentiation toward Interleukin-12 Hypo-Producing Dendritic Cells Via a TGR5-Dependent Pathway. *Immunology* **2012,** 136, 153-162.

124. Sun, L.; Ye, R. D. Role of G Protein-Coupled Receptors in Inflammation. *Acta Pharmacol Sin* **2012,** 33, 342-350.

125. Tang, X.-l.; Wang, Y.; Li, D.-l.; Luo, J.; Liu, M.-y. Orphan G Protein-Coupled Receptors (GPCRs): Biological Functions and Potential Drug Targets. *Acta Pharmacol Sin* **2012,** 33, 363-371.

126. McMillin, M.; Frampton, G.; Tobin, R.; Dusio, G.; Smith, J.; Shin, H.; Newell-Rogers, K.; Grant, S.; DeMorrow, S. TGR5 Signaling Reduces Neuroinflammation During Hepatic Encephalopathy. *J Neurochem* **2015,** 135, 565-576.

127. Cipriani, S.; Mencarelli, A.; Chini, M. G.; Distrutti, E.; Renga, B.; Bifulco, G.; Baldelli, F.; Donini, A.; Fiorucci, S. The Bile Acid Receptor GPBAR-1 (TGR5) Modulates Integrity of Intestinal Barrier and Immune Response to Experimental Colitis. *PLoS One* **2011,** 6, e25637.

128. Yoneno, K.; Hisamatsu, T.; Shimamura, K.; Kamada, N.; Ichikawa, R.; Kitazume, M. T.; Mori, M.; Uo, M.; Namikawa, Y.; Matsuoka, K. TGR5 Signalling Inhibits the Production of Pro-Inflammatory Cytokines by in Vitro Differentiated Inflammatory and Intestinal Macrophages in Crohn's Disease. *Immunology* **2013,** 139, 19-29.

129. Beuers, U.; Hohenester, S.; de Buy Wenniger, L. J. M.; Kremer, A. E.; Jansen, P. L.; Elferink, R. P. The Biliary Hco3− Umbrella: A Unifying Hypothesis on Pathogenetic and Therapeutic Aspects of Fibrosing Cholangiopathies. *Hepatology* **2010,** 52, 1489-1496.

130. Sodeman, T.; Bronk, S. F.; Roberts, P. J.; Miyoshi, H.; Gores, G. J. Bile Salts Mediate Hepatocyte Apoptosis by Increasing Cell Surface Trafficking of Fas. *Am J Physiol Gastrointest Liver Physiol* **2000,** 278, G992-G999.

131. Reinehr, R.; Graf, D.; Fischer, R.; Schliess, F.; Häussinger, D. Hyperosmolarity Triggers CD95 Membrane Trafficking and Sensitizes Rat Hepatocytes toward Cd95l-Induced Apoptosis. *Hepatology* **2002,** 36, 602-614.

132. Jensen, D. D.; Godfrey, C. B.; Niklas, C.; Canals, M.; Kocan, M.; Poole, D. P.; Murphy, J. E.; Alemi, F.; Cottrell, G. S.; Korbmacher, C.; Lambert, N. A.; Bunnett, N. W.; Corvera, C. U. The Bile Acid Receptor TGR5 Does Not Interact with Beta-Arrestins or Traffic to Endosomes but Transmits Sustained Signals from Plasma Membrane Rafts. *J Biol Chem* **2013,** 288, 22942-22960.

133. Que, F. G.; Phan, V. A.; Phan, V. H.; Celli, A.; Batts, K.; LaRusso, N. F.; Gores, G. J. Cholangiocarcinomas Express Fas Ligand and Disable the Fas Receptor. *Hepatology* **1999,** 30, 1398-1404.

134. Powell, J.; McConkey, C. C. The Rising Trend in Oesophageal Adenocarcinoma and Gastric Cardia. *Eur J Cancer Prev* **1992,** 1, 265-270.

135. Lindblad, M.; Rodríguez, L. G.; Lagergren, J. Body Mass, Tobacco and Alcohol and Risk of Esophageal, Gastric Cardia, and Gastric Non-Cardia Adenocarcinoma among Men and Women in a Nested Case-Control Study. *Cancer Causes Control* **2005,** 16, 285-294.

136. Hölscher, A. H.; Bollschweiler, E.; Schneider, P. M.; Siewert, J. R. Early Adenocarcinoma in Barrett's Oesophagus. *British Journal of Surgery* **1997,** 84, 1470-1473.

137. Shaib, Y.; El-Serag, H. B. In *The Epidemiology of Cholangiocarcinoma*, Seminars in liver disease, 2004; 2004; pp 115-125.

138. Patel, T. Worldwide Trends in Mortality from Biliary Tract Malignancies. *BMC cancer* **2002,** 2, 10.

139. Oldham, W. M.; Hamm, H. E. Heterotrimeric G Protein Activation by G-Protein-Coupled Receptors. *Nat Rev Mol Cell Biol* **2008,** 9, 60-71.

140. Katritch, V.; Cherezov, V.; Stevens, R. C. Diversity and Modularity of G Protein-Coupled Receptor Structures. *Trends Pharmacol Sci* **2012,** 33, 17-27.

141. Venkatakrishnan, A. J.; Deupi, X.; Lebon, G.; Tate, C. G.; Schertler, G. F.; Babu, M. M. Molecular Signatures of G-Protein-Coupled Receptors. *Nature* **2013,** 494, 185-194.

142. Marchese, A.; George, S. R.; Kolakowski Jr, L. F.; Lynch, K. R.; O'Dowd, B. F. Novel GPCRs and Their Endogenous Ligands: Expanding the Boundaries of Physiology and Pharmacology. *Trends Pharmacol Sci* **1999,** 20, 370-375.

143. Wu, H.; Wang, C.; Gregory, K. J.; Han, G. W.; Cho, H. P.; Xia, Y.; Niswender, C. M.; Katritch, V.; Meiler, J.; Cherezov, V. Structure of a Class C GPCR Metabotropic Glutamate Receptor 1 Bound to an Allosteric Modulator. *Science* **2014,** 344, 58-64.

144. Katritch, V.; Abagyan, R. GPCR Agonist Binding Revealed by Modeling and Crystallography. *Trends Pharmacol Sci* **2011,** 32, 637-643.

145. Rosenbaum, D. M.; Rasmussen, S. G. F.; Kobilka, B. K. The Structure and Function of G-Protein-Coupled Receptors. *Nature* **2009,** 459, 356-363.

146. Wheatley, M.; Wootten, D.; Conner, M. T.; Simms, J.; Kendrick, R.; Logan, R. T.; Poyner, D. R.; Barwell, J. Lifting the Lid on GPCRs: The Role of Extracellular Loops. *Br J Pharmacol* **2012,** 165, 1688-1703.

147. Hanson, M. A.; Roth, C. B.; Jo, E.; Griffith, M. T.; Scott, F. L.; Reinhart, G.; Desale, H.; Clemons, B.; Cahalan, S. M.; Schuerer, S. C.; Sanna, M. G.; Han, G. W.; Kuhn, P.; Rosen, H.; Stevens, R. C. Crystal Structure of a Lipid G Protein–Coupled Receptor. *Science* **2012,** 335, 851-855.

148. Gether, U.; Asmar, F.; Meinild, A. K.; Rasmussen, S. G. F. Structural Basis for Activation of G-Protein-Coupled Receptors. *Pharmacol Toxicol* **2002,** 91, 304-312.

149. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science (New York, NY)* **2000,** 289, 739-745.

150. Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a B1 Adrenergic G-Protein-Coupled Receptor. *Nature* **2008,** 454, 486-U482.

151. Dror, R. O.; Arlow, D. H.; Maragakis, P.; Mildorf, T. J.; Pan, A. C.; Xu, H. F.; Borhani, D. W.; Shaw, D. E. Activation Mechanism of the Beta(2)-Adrenergic Receptor. *P Natl Acad Sci USA* **2011,** 108, 18684-18689.

152. Schneider, M.; Wolf, S.; Schlitter, J.; Gerwert, K. The Structure of Active Opsin as a Basis for Identification of GPCR Agonists by Dynamic Homology Modelling and Virtual Screening Assays. *Febs Lett* **2011,** 585, 3587-3592.

153. Swaminath, G.; Deupi, X.; Lee, T. W.; Zhu, W.; Thian, F. S.; Kobilka, T. S.; Kobilka, B. Probing the B2 Adrenoceptor Binding Site with Catechol Reveals Differences in Binding and Activation by Agonists and Partial Agonists. *J Biol Chem* **2005,** 280, 22165-22171.

154. Ballesteros, J. A.; Weinstein, H. [19] Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. In *Methods Neurosci*, Stuart, C. S., Ed. Academic Press: 1995; Vol. Volume 25, pp 366-428.

155. Pollock, N. J.; Manelli, A. M.; Hutchins, C. W.; Steffey, M. E.; MacKenzie, R. G.; Frail, D. E. Serine Mutations in Transmembrane V of the Dopamine D1 Receptor Affect Ligand Interactions and Receptor Activation. *J Biol Chem* **1992,** 267, 17780-17786.

156. Gao, Z. G.; Chen, A.; Barak, D.; Kim, S. K.; Muller, C. E.; Jacobson, K. A. Identification by Site-Directed Mutagenesis of Residues Involved in Ligand Recognition and Activation of the Human a(3) Adenosine Receptor. *J Biol Chem* **2002,** 277, 19056-19063.

157. Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the Beta2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2011,** 477, 549-555.

158. Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Jacobson, K. A.; Gao, Z.-G.; Cherezov, V.; Stevens, R. C. Structure of an Agonist-Bound Human A2a Adenosine Receptor. *Science* **2011,** 332, 322-327.

159. Wang, C.; Jiang, Y.; Ma, J.; Wu, H.; Wacker, D.; Katritch, V.; Han, G. W.; Liu, W.; Huang, X.-P.; Vardy, E.; McCorvy, J. D.; Gao, X.; Zhou, X. E.; Melcher, K.; Zhang, C.; Bai, F.; Yang, H.; Yang, L.; Jiang, H.; Roth, B. L.; Cherezov, V.; Stevens, R. C.; Xu, H. E. Structural Basis for Molecular Recognition at Serotonin Receptors. *Science* **2013,** 340, 610-614.

160. Hasbi, A.; Perreault, M. L.; Shen, M. Y. F.; Zhang, L.; To, R.; Fan, T.; Nguyen, T.; Ji, X.; O'Dowd, B. F.; George, S. R. A Peptide Targeting an Interaction Interface Disrupts the Dopamine D1-D2 Receptor Heteromer to Block Signaling and Function in Vitro and in Vivo: Effective Selective Antagonism. *The FASEB Journal* **2014,** 28, 4806-4820.

161. Kim, J. H.; Wess, J.; Vanrhee, A. M.; Schoneberg, T.; Jacobson, K. A. Site-Directed Mutagenesis Identifies Residues Involved in Ligand Recognition in the Human A(2a) Adenosine Receptor. *J Biol Chem* **1995,** 270, 13987-13997.

162. Zylberg, J.; Ecke, D.; Fischer, B.; Reiser, G. Structure and Ligand-Binding Site Characteristics of the Human P2Y11 Nucleotide Receptor Deduced from Computational Modelling and Mutational Analysis. *Biochem J* **2007,** 405, 277-286.

163. Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; IJzerman, A. P.; Stevens, R. C. The 2.6 Angstrom Crystal Structure of a Human A2a Adenosine Receptor Bound to an Antagonist. *Science* **2008,** 322, 1211-1217.

164. Studer, E.; Zhou, X. Q.; Zhao, R. P.; Wang, Y.; Takabe, K.; Nagahashi, M.; Pandak, W. M.; Dent, P.; Spiegel, S.; Shi, R. H.; Xu, W. R.; Liu, X. Y.; Bohdan, P.; Zhang, L. Y.; Zhou, H. P.; Hylemon, P. B. Conjugated Bile Acids Activate the Sphingosine-1-Phosphate Receptor 2 in Primary Rodent Hepatocytes. *Hepatology* **2012,** 55, 267-276.

165. Deupi, X.; Edwards, P.; Singhal, A.; Nickle, B.; Oprian, D.; Schertler, G.; Standfuss, J. Stabilized G Protein Binding Site in the Structure of Constitutively Active Metarhodopsin-II. *P Natl Acad Sci USA* **2012,** 109, 119-124.

166. Ghosh, E.; Kumari, P.; Jaiman, D.; Shukla, A. K. Methodological Advances: The Unsung Heroes of the GPCR Structural Revolution. *Nat Rev Mol Cell Biol* **2015,** 16, 69-81.

167. Provasi, D.; Artacho, M. C.; Negri, A.; Mobarec, J. C.; Filizola, M. Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques. *PLoS Comput Biol* **2011,** 7.

168. Tautermann, C. S.; Pautsch, A. The Implication of the First Agonist Bound Activated GPCR X-Ray Structure on GPCR in Silico Modeling. *ACS Med Chem Lett* **2011,** 2, 414-418.

169. Leioatts, N.; Suresh, P.; Romo, T. D.; Grossfield, A. Structure-Based Simulations Reveal Concerted Dynamics of GPCR Activation. *Proteins: Struct, Funct, Bioinf* **2014**, n/a-n/a.

170. Ghanouni, P.; Gryczynski, Z.; Steenhuis, J. J.; Lee, T. W.; Farrens, D. L.; Lakowicz, J. R.; Kobilka, B. K. Functionally Different Agonists Induce Distinct Conformations in the G Protein Coupling Domain of the β2 Adrenergic Receptor. *J Biol Chem* **2001,** 276, 24433-24436.

171. Chien, E. Y. T.; Liu, W.; Zhao, Q.; Katritch, V.; Won Han, G.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science* **2010,** 330, 1091-1095.

172. Bhattacharya, S.; Hall, S. E.; Li, H.; Vaidehi, N. Ligand-Stabilized Conformational States of Human β2 Adrenergic Receptor: Insight into G-Protein-Coupled Receptor Activation. *Biophys J* **2008,** 94, 2027-2042.

173. Moreira, I. S. Structural Features of the G-Protein/GPCR Interactions. *Biochim Biophys Acta* **2014,** 1840, 16-33.

174. Flock, T.; Ravarani, C. N. J.; Sun, D.; Venkatakrishnan, A. J.; Kayikci, M.; Tate, C. G.; Veprintsev, D. B.; Babu, M. M. Universal Allosteric Mechanism for Gα Activation by GPCRs. *Nature* **2015,** 524, 173-179.

175. Mnpotra, J. S.; Qiao, Z.; Cai, J.; Lynch, D. L.; Grossfield, A.; Leioatts, N.; Hurst, D. P.; Pitman, M. C.; Song, Z.-H.; Reggio, P. H. Structural Basis of G Protein-Coupled Receptor- Gi Protein Interaction: Formation of the Cannabinoid CB2 Receptor / Gi Protein Complex. *J Biol Chem* **2014**.

176. Chung, K. Y.; Rasmussen, S. G. F.; Liu, T.; Li, S.; DeVree, B. T.; Chae, P. S.; Calinski, D.; Kobilka, B. K.; Woods, V. L.; Sunahara, R. K. Conformational Changes in the G Protein Gs Induced by the β2 Adrenergic Receptor. *Nature* **2011,** 477, 611-615.

177. Möller, S.; Vilo, J.; Croning, M. D. R. Prediction of the Coupling Specificity of G Protein Coupled Receptors to Their G Proteins. *Bioinformatics* **2001,** 17, S174-S181.

178. Hara, M. R.; Kovacs, J. J.; Whalen, E. J.; Rajagopal, S.; Strachan, R. T.; Grant, W.; Towers, A. J.; Williams, B.; Lam, C. M.; Xiao, K.; Shenoy, S. K.; Gregory, S. G.; Ahn, S.; Duckett, D. R.; Lefkowitz, R. J. A Stress Response Pathway Regulates DNA Damage through β2-Adrenoreceptors and β-Arrestin-1. *Nature* **2011**.

179. Kang, Y.; Zhou, X. E.; Gao, X.; He, Y.; Liu, W.; Ishchenko, A.; Barty, A.; White, T. A.; Yefanov, O.; Han, G. W.; Xu, Q.; de Waal, P. W.; Ke, J.; Tan, M. H. E.; Zhang, C.; Moeller, A.; West, G. M.; Pascal, B. D.; Van Eps, N.; Caro, L. N.; Vishnivetskiy, S. A.; Lee, R. J.; Suino-Powell, K. M.; Gu, X.; Pal, K.; Ma, J.; Zhi, X.; Boutet, S.; Williams, G. J.; Messerschmidt, M.; Gati, C.; Zatsepin, N. A.; Wang, D.; James, D.; Basu, S.; Roy-Chowdhury, S.; Conrad, C. E.; Coe, J.; Liu, H.; Lisova, S.; Kupitz, C.; Grotjohann, I.; Fromme, R.; Jiang, Y.; Tan, M.; Yang, H.; Li, J.; Wang, M.; Zheng, Z.; Li, D.; Howe, N.; Zhao, Y.; Standfuss, J.; Diederichs, K.; Dong, Y.; Potter, C. S.; Carragher, B.; Caffrey, M.; Jiang, H.; Chapman, H. N.; Spence, J. C. H.; Fromme, P.; Weierstall, U.; Ernst, O. P.; Katritch, V.; Gurevich, V. V.; Griffin, P. R.; Hubbell, W. L.; Stevens, R. C.; Cherezov, V.; Melcher, K.; Xu, H. E. Crystal Structure of Rhodopsin Bound to Arrestin by Femtosecond X-Ray Laser. *Nature* **2015,** 523, 561-567.

180. Black, J. B.; Premont, R. T.; Daaka, Y. In *Feedback Regulation of G Protein-Coupled Receptor Signaling by Grks and Arrestins*, Seminars in cell & developmental biology, 2016; Elsevier: 2016.

181. Nobles, K. N.; Xiao, K.; Ahn, S.; Shukla, A. K.; Lam, C. M.; Rajagopal, S.; Strachan, R. T.; Huang, T.-Y.; Bressler, E. A.; Hara, M. R.; Shenoy, S. K.; Gygi, S. P.; Lefkowitz, R. J. Distinct Phosphorylation Sites on the β2-Adrenergic Receptor Establish a Barcode That Encodes Differential Functions of B-Arrestin. *Sci Signal* **2011,** 4, ra51.

182. Qian, J.; Wu, C.; Chen, X.; Li, X.; Ying, G.; Jin, L.; Ma, Q.; Li, G.; Shi, Y.; Zhang, G. Differential Requirements of Arrestin-3 and Clathrin for Ligand-Dependent and-Independent Internalization of Human G Protein-Coupled Receptor 40. *Cell Signalling* **2014,** 26, 2412-2423.

183. Kang, D. S.; Tian, X.; Benovic, J. L. Role of β-Arrestins and Arrestin Domain-Containing Proteins in G Protein-Coupled Receptor Trafficking. *Curr Opin cell Biol* **2014,** 27, 63-71.

184. Irannejad, R.; Tomshine, J. C.; Tomshine, J. R.; Chevalier, M.; Mahoney, J. P.; Steyaert, J.; Rasmussen, S. G. F.; Sunahara, R. K.; El-Samad, H.; Huang, B.; von Zastrow, M. Conformational Biosensors Reveal GPCR Signalling from Endosomes. *Nature* **2013,** 495, 534-538.

185. Evron, T.; Peterson, S. M.; Urs, N. M.; Bai, Y.; Rochelle, L. K.; Caron, M. G.; Barak, L. S. G Protein and β-Arrestin Signaling Bias at the Ghrelin Receptor. *J Biol Chem* **2014,** 289, 33442-33455.

186. Esseltine, J. L.; Ferguson, S. S. Regulation of G Protein-Coupled Receptor Trafficking and Signaling by Rab GTPases. *Small GTPases* **2013,** 4, 132-135.

187. Wang, G.; Wu, G. Small GTPase Regulation of GPCR Anterograde Trafficking. *Trends Pharmacol Sci* **2012,** 33, 28-34.

188. Lachance, V.; Degrandmaison, J.; Marois, S.; Robitaille, M.; Génier, S.; Nadeau, S.; Angers, S.; Parent, J.-L. Ubiquitylation and Activation of a Rab GTPase Is Promoted by a β2AR–HACE1 Complex. *J Cell Sci* **2014,** 127, 111-123.

189. Magalhaes, A. C.; Dunn, H.; Ferguson, S. S. Regulation of GPCR Activity, Trafficking and Localization by GPCR-Interacting Proteins. *Brit J Pharmacol* **2012,** 165, 1717-1736.

190. Duvernay, M.; Dong, C.; Zhang, X.; Zhou, F.; Nichols, C.; Wu, G. Anterograde Trafficking of G Protein-Coupled Receptors: Function of the C-Terminal F(X)6LL Motif in Export from the Endoplasmic Reticulum. *Mol Pharmacol* **2009,** 75, 751 - 761.

191. Bornert, O.; Møller, T. C.; Boeuf, J.; Candusso, M.-P.; Wagner, R.; Martinez, K. L.; Simonin, F. Identification of a Novel Protein-Protein Interaction Motif Mediating Interaction of GPCR-Associated Sorting Proteins with G Protein-Coupled Receptors. *PLoS ONE* **2013,** 8, e56336.

192. Abascal-Palacios, G.; Schindler, C.; Rojas, A. L.; Bonifacino, J. S.; Hierro, A. Structural Basis for the Interaction of the Golgi-Associated Retrograde Protein Complex with the T-SNARE Syntaxin 6. *Structure* **2013,** 21, 1698-1706.

193. Bonifacino, J. S.; Glick, B. S. The Mechanisms of Vesicle Budding and Fusion. *Cell* **2004,** 116, 153-166.

194. Eggink, H.; Soeters, M.; Pols, T. TGR5 Ligands as Potential Therapeutics in Inflammatory Diseases. *Intl J Interferon Cytokine Mediat Res* **2014,** 6, 27-38.

195. Pauling, L.; Corey, R. B.; Branson, H. R. The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *P Natl Acad Sci* **1951,** 37, 205-211.

196. Zheng, H.; Pearsall, E. A.; Hurst, D. P.; Zhang, Y. H.; Chu, J.; Zhou, Y. L.; Reggio, P. H.; Loh, H. H.; Law, P. Y. Palmitoylation and Membrane Cholesterol Stabilize Mu-Opioid Receptor Homodimerization and G Protein Coupling. *BMC Cell Biol* **2012,** 13.

197. Uddin, M. S.; Kim, H.; Deyo, A.; Naider, F.; Becker, J. M. Identification of Residues Involved in Homodimer Formation Located within a Beta-Strand Region of the N-Terminus of a Yeast G Protein-Coupled Receptor. *J Recept Signal Transduct* **2012,** 32, 65-75.

198. Gurevich, V. V.; Gurevich, E. V. How and Why Do GPCRs Dimerize? *Trends Pharmacol Sci* **2008,** 29, 234-240.

199. Knepp, A. M.; Periole, X.; Marrink, S.-J.; Sakmar, T. P.; Huber, T. Rhodopsin Forms a Dimer with Cytoplasmic Helix 8 Contacts in Native Membranes. *Biochemistry* **2012,** 51, 1819-1821.

200. Milligan, G. G Protein-Coupled Receptor Dimerization: Function and Ligand Pharmacology. *Mol Pharmacol* **2004,** 66, 1 - 7.

201. Selent, J.; Kaczor, A. A. Oligomerization of G Protein-Coupled Receptors: Computational Methods. *Curr Med Chem* **2011,** 18, 4588-4605.

202. Kaczor, A. A.; Selent, J. Oligomerization of G Protein-Coupled Receptors: Biochemical and Biophysical Methods. *Curr Med Chem* **2011,** 18, 4606-4634.

203. Lee, S.; O'Dowd, B.; Ng, G.; Varghese, G.; Akil, H.; Mansour, A.; Nguyen, T.; George, S. Inhibition of Cell Surface Expression by Mutant Receptors Demonstrates That D2 Dopamine Receptors Exist as Oligomers in the Cell. *Mol Pharmacol* **2000,** 58, 120 - 128.

204. Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. Organization of the G Protein-Coupled Receptors Rhodopsin and Opsin in Native Membranes. *J Biol Chem* **2003,** 278, 21655-21662.

205. Terrillon, S.; Bouvier, M. Roles of G-Protein-Coupled Receptor Dimerization. *EMBO Reports* **2004,** 5, 30-34.

206. White, J. H.; Wise, A.; Main, M. J.; Green, A.; Fraser, N. J.; Disney, G. H.; Barnes, A. A.; Emson, P.; Foord, S. M.; Marshall, F. H. Heterodimerization Is Required for the Formation of a Functional GABAb Receptor. *Nature* **1998,** 396, 679-682.

207. Milligan, G. G Protein-Coupled Receptor Dimerization: Function and Ligand Pharmacology. *Mol Pharmacol* **2004,** 66, 1-7.

208. Teichmann, A.; Gibert, A.; Lampe, A.; Grzesik, P.; Rutz, C.; Furkert, J.; Schmoranzer, J.; Krause, G.; Wiesner, B.; Schülein, R. The Specific Monomer/Dimer Equilibrium of the Corticotropin-Releasing Factor Receptor Type 1 Is Established in the Endoplasmic Reticulum. *J Biol Chem* **2014**.

209. Ward, R. J.; Pediani, J. D.; Godin, A. G.; Milligan, G. Regulation of Oligomeric Organization of the Serotonin 5-HT2c Receptor Observed by Spatial Intensity Distribution Analysis. *J Biol Chem* **2015,** 290, 12844-12857.

210. Rozenfeld, R.; Bushlin, I.; Gomes, I.; Tzavaras, N.; Gupta, A.; Neves, S.; Battini, L.; Gusella, G. L.; Lachmann, A.; Ma'ayan, A.; Blitzer, R. D.; Devi, L. A. Receptor Heteromerization Expands the Repertoire of Cannabinoid Signaling in Rodent Neurons. *PLoS ONE* **2012,** 7, e29239.

211. AbdAlla, S.; Lother, H.; Quitterer, U. At1-Receptor Heterodimers Show Enhanced G-Protein Activation and Altered Receptor Sequestration. *Nature* **2000,** 407, 94 - 98.

212. Waldhoer, M.; Fong, J.; Jones, R.; Lunzer, M.; Sharma, S.; Kostenis, E.; Portoghese, P.; Whistler, J. A Heterodimer-Selective Agonist Shows in Vivo Relevance of G Protein-Coupled Receptor Dimers. *P Natl Acad Sci USA* **2005,** 102, 9050 - 9055.

213. Armando, S.; Quoyer, J.; Lukashova, V.; Maiga, A.; Percherancier, Y.; Heveker, N.; Pin, J.-P.; Prézeau, L.; Bouvier, M. The Chemokine CXC4 and CC2 Receptors Form Homo- and Heterooligomers That Can Engage Their Signaling G-Protein Effectors and Barrestin. *The FASEB Journal* **2014,** 28, 4509-4523.

214. Hiller, C.; Kühhorn, J.; Gmeiner, P. Class a G-Protein-Coupled Receptor (GPCR) Dimers and Bivalent Ligands. *J Med Chem* **2013,** 56, 6542-6559.

215. Taddese, B.; Simpson, L. M.; Wall, I. D.; Blaney, F. E.; Kidley, N. J.; Clark, H. S. X.; Smith, R. E.; Upton, G. J. G.; Gouldson, P. R.; Psaroudakis, G.; Bywater, R. P.; Reynolds, C. A. G-Protein-Coupled Receptor Dynamics: Dimerization and Activation Models Compared with Experiment. *Biochem Soc T* **2012,** 40, 394-399.

216. Christopoulos, A. Advances in GPCR Allostery: From Function to Structure. *Mol Pharmacol* **2014**.

217. Lavoie, C.; Mercier, J.; Salahpour, A.; Umapathy, D.; Breit, A.; Villeneuve, L.; Zhu, W.; Xiao, R.; Lakatta, E.; Bouvier, M. β1/ β2-Adrenergic Receptor Heterodimerization Regulates β2-Adrenergic Receptor Internalization and Erk Signaling Efficacy. *J Biol Chem* **2002,** 277, 35402 - 35410.

218. Pfeiffer, M.; Koch, T.; Schröder, H.; Laugsch, M.; Höllt, V.; Schulz, S. Heterodimerization of Somatostatin and Opioid Receptors Cross-Modulates Phosphorylation, Internalization, and Desensitization. *J Biol Chem* **2002,** 277, 19762-19772.

219. Jordan, B.; Trapaidze, N.; Gomes, I.; Nivarthi, R.; Devi, L. Oligomerization of Opioid Receptors with β2-Adrenergic Receptors: A Role in Trafficking and Mitogen-Activated Protein Kinase Activation. *P Natl Acad Sci* **2001,** 98, 343-348.

220. Terrillon, S.; Barberis, C.; Bouvier, M. Heterodimerization of V1a and V2 Vasopressin Receptors Determines the Interaction with β-Arrestin and Their Trafficking Patterns. *P Natl Acad Sci USA* **2004,** 101, 1548-1553.

221. Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X.-P.; Carroll, F. I. Structure of the Human κ-Opioid Receptor in Complex with JDTIC. *Nature* **2012,** 485, 327-332.

222. Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S. Crystal Structure of the Mu-Opioid Receptor Bound to a Morphinan Antagonist. *Nature* **2012,** 485, 321-U170.

223. Johnston, J. M.; Wang, H.; Provasi, D.; Filizola, M. Assessing the Relative Stability of Dimer Interfaces in G Protein-Coupled Receptors. *PLoS Comput Biol* **2012,** 8.

224. Chini, B.; Parenti, M. G-Protein-Coupled Receptors, Cholesterol and Palmitoylation: Facts About Fats. *J Mol Endocrinol* **2009,** 42, 371-379.

225. Goddard, A.; Watts, A. Regulation of G Protein-Coupled Receptors by Palmitoylation and Cholesterol. *BMC Biology* **2012,** 10, 27.

226. Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor. *Structure* **2008,** 16, 897-905.

227. Ecke, D.; Hanck, T.; Tulapurkar, M. E.; Schafer, R.; Kassack, M.; Stricker, R.; Reiser, G. Hetero-Oligomerization of the P2Y11 Receptor with the P2Y1 Receptor Controls the Internalization and Ligand Selectivity of the P2Y11 Receptor. *Biochem J* **2008,** 409, 107-116.

228. George, S.; Fan, T.; Xie, Z.; Tse, R.; Tam, V.; Varghese, G.; O'Dowd, B. Oligomerization of Mu- and Delta-Opioid Receptors. Generation of Novel Functional Properties. *J Biol Chem* **2000,** 275, 26128 - 26135.

229. Pisterzi, L.; Jansma, D.; Georgiou, J.; Woodside, M.; Chou, J.; Angers, S.; Raicu, V.; Wells, J. Oligomeric Size of the M2 Muscarinic Receptor in Live Cells as Determined by Quantitative Fluorescence Resonance Energy Transfer. *J Biol Chem* **2010,** 285, 16723 - 16738.

230. Hammad, M.; Dupre, D. Chaperones Contribute to G Protein Coupled Receptor Oligomerization, but Do Not Participate in Assembly of the G Protein with the Receptor Signaling Complex. *J Mol Signaling* **2010,** 5, 16.

231. Albizu, L.; Cottet, M.; Kralikova, M.; Stoev, S.; Seyer, R.; Brabet, I.; Roux, T.; Bazin, H.; Bourrier, E.; Lamarque, L. Time-Resolved FRET between GPCR Ligands Reveals Oligomers in Native Tissues. *Nat Chem Biol* **2010,** 6, 587-594.

232. Ciruela, F.; Vilardaga, J. P.; Fernandez-Duenas, V. Lighting up Multiprotein Complexes: Lessons from GPCR Oligomerization. *Trends Biotechnol* **2010,** 28, 407-415.

233. Sali, A.; Berman, H. M.; Schwede, T.; Trewhella, J.; Kleywegt, G.; Burley, S. K.; Markley, J.; Nakamura, H.; Adams, P.; Bonvin, A. M. Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **2015,** 23, 1156-1167.

234. Becker, O.; Shacham, S.; Marantz, Y.; Noiman, S. Modeling the 3D Structure of GPCRs: Advances and Application to Drug Discovery. *Curr Opin Drug Discovery Dev* **2003,** 6, 353-361.

235. Cavasotto, C. N.; Phatak, S. S. Homology Modeling in Drug Discovery: Current Trends and Applications. *Drug Discov Today* **2009,** 14, 676-683.

236. Michino, M.; Abola, E.; Dock, G.; Brooks, C. L.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-Wide Assessment of GPCR Structure Modelling and Ligand Docking: GPCR Dock 2008. *Nat Rev Drug Discov* **2009,** 8, 455-463.

237. Bronson, J. E.; Fei, J.; Hofman, J. M.; Gonzalez, R. L.; Wiggins, C. H. Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *Biophys J* **2009,** 97, 3196-3205.

238. Alber, F.; Förster, F.; Korkin, D.; Topf, M.; Sali, A. Integrating Diverse Data for Structure Determination of Macromolecular Assemblies. *Annu Rev Biochem* **2008,** 77, 443-477.

239. Zhang, H.-M.; Kazazic, S. a.; Schaub, T. M.; Tipton, J. D.; Emmett, M. R.; Marshall, A. G. Enhanced Digestion Efficiency, Peptide Ionization Efficiency, and Sequence Resolution for Protein Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem* **2008,** 80, 9034-9041.

240. Kazazic, S.; Zhang, H.-M.; Schaub, T. M.; Emmett, M. R.; Hendrickson, C. L.; Blakney, G. T.; Marshall, A. G. Automated Data Reduction for Hydrogen/Deuterium Exchange Experiments, Enabled by High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J Am Soc Mass Spectrom* **2010,** 21, 550-558.

241. Wang, Y. S.; Liu, D.; Wyss, D. F. Competition Std Nmr for the Detection of High-Affinity Ligands and Nmr-Based Screening. *Magn Reson Chem* **2004,** 42, 485-489.

242. Angulo, J.; Enríquez-Navas, P. M.; Nieto, P. M. Ligand–Receptor Binding Affinities from Saturation Transfer Difference (STD) NMR Spectroscopy: The Binding Isotherm of STD Initial Growth Rates. *Chem - Eur J* **2010,** 16, 7803-7812.

243. Salim, K.; Fenton, T.; Bacha, J.; Urien-Rodriguez, H.; Bonnert, T.; Skynner, H. A.; Watts, E.; Kerby, J.; Heald, A.; Beer, M. Oligomerization of G-Protein-Coupled Receptors Shown by Selective Co-Immunoprecipitation. *J Biol Chem* **2002,** 277, 15482-15485.

244. Baragli, A.; Alturaihi, H.; Watt, H. L.; Abdallah, A.; Kumar, U. Heterooligomerization of Human Dopamine Receptor 2 and Somatostatin Receptor 2: Co-Immunoprecipitation and Fluorescence Resonance Energy Transfer Analysis. *Cell Signalling* **2007,** 19, 2304-2316.

245. Jares-Erijman, E. A.; Jovin, T. M. FRET Imaging. *Nat Biotechnol* **2003,** 21, 1387-1395.

246. Roy, R.; Hohng, S.; Ha, T. A Practical Guide to Single-Molecule FRET. *Nat Methods* **2008,** 5, 507-516.

247. Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. A Toolkit and Benchmark Study for FRET-Restrained High-Precision Structural Modeling. *Nat Methods* **2012,** 9, 1218-1225.

248. Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Müller, S.; Seidel, C. A. Accurate Distance Determination of Nucleic Acids Via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J Am Chem Soc* **2011,** 133, 2463-2480.

249. Khokhlov, A. R. *Statistical Physics of Macromolecules*. Amer Inst of Physics: 1994.

250. Zhuang, X.; Kim, H.; Pereira, M. J.; Babcock, H. P.; Walter, N. G.; Chu, S. Correlating Structural Dynamics and Function in Single Ribozyme Molecules. *Science* **2002,** 296, 1473-1476.

251. Sali, A. W., B. *Modeller*, 9v9; 2011.

252. Eswar, N.; Marti-Renom, M. A.; Webb, B.; Madhusudhan, M. S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A. Comparative Protein Structure Modeling with Modeller. *Curr Protoc Bioinformatics* **2006**, 5.6.1-5.6.30.

253. Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use? *J Med Chem* **2009,** 52, 5207-5216.

254. Moult, J. A Decade of Casp: Progress, Bottlenecks and Prognosis in Protein Structure Prediction. *Curr Opin Struct Biol* **2005,** 15, 285-289.

255. Moult, J.; Fidelis, K.; Kryshtafovych, A.; Schwede, T.; Tramontano, A. Critical Assessment of Methods of Protein Structure Prediction (Casp)—Round X. *Proteins: Struct, Funct, Bioinf* **2014,** 82, 1-6.

256. Bower, M. J.; Cohen, F. E.; Dunbrack, R. L. Prediction of Protein Side-Chain Rotamers from a Backbone-Dependent Rotamer Library: A New Homology Modeling Tool. *J Mol Biol* **1997,** 267, 1268-1282.

257. Larsson, P.; Wallner, B.; Lindahl, E.; Elofsson, A. Using Multiple Templates to Improve Quality of Homology Models in Automated Homology Modeling. *Protein Sci* **2008,** 17, 990-1002.

258. Gertzen, C. G. W. Struktur-Aktivitaetsbeziehungen Von P2Y11-Antagonisten. Diploma, Christian-Albrechts-Universitaet zu Kiel, 2011.

259. Novoa, E. M.; Pouplana, L. R. d.; Barril, X.; Orozco, M. Ensemble Docking from Homology Models. *J Chem Theory Comput* **2010,** 6, 2547-2557.

260. Floudas, C.; Klepeis, J.; Pardalos, P. Global Optimization Approaches in Protein Folding and Peptide Docking. *DIMACS series in discrete mathematics and theoretical computer science* **1999,** 47, 141-171.

261. Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated Docking of Flexible Ligands: Applications of Autodock. *J Mol Recognit* **1996,** 9, 1-5.

262. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J Comput Chem* **1998,** 19, 1639-1662.

263. Velec, H. F. G.; Gohlke, H.; Klebe, G. DrugScore(CSD)-Knowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of near-Native Ligand Poses and Better Affinity Prediction. *J Med Chem* **2005,** 48, 6296-6303.

264. Sotriffer, C. A.; Gohlke, H.; Klebe, G. Docking into Knowledge-Based Potential Fields: A Comparative Evaluation of DrugScore. *J Med Chem* **2002,** 45, 1967-1970.

265. Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions. *J Mol Biol* **2000,** 295, 337-356.

266. Moro, S.; Guo, D.; Camaioni, E.; Boyer, J. L.; Harden, T. K.; Jacobson, K. A. Human P2Y1 Receptor: Molecular Modeling and Site-Directed Mutagenesis as Tools to Identify Agonist and Antagonist Recognition Sites. *J Med Chem* **1998,** 41, 1456-1466.

267. de Graaf, C.; Pospisil, P.; Pos, W.; Folkers, G.; Vermeulen, N. P. E. Binding Mode Prediction of Cytochrome P450 and Thymidine Kinase Protein−Ligand Complexes by Consideration of Water and Rescoring in Automated Docking. *J Med Chem* **2005,** 48, 2308-2318.

268. Ragno, R.; Simeoni, S.; Castellano, S.; Vicidomini, C.; Mai, A.; Caroli, A.; Tramontano, A.; Bonaccini, C.; Trojer, P.; Bauer, I.; Brosch, G.; Sbardella, G. Small Molecule Inhibitors of Histone Arginine Methyltransferases: Homology Modeling, Molecular Docking, Binding Mode Analysis, and Biological Evaluations. *J Med Chem* **2007,** 50, 1241-1253.

269. Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved Binding Mode of Human β2 Adrenergic Receptor Inverse Agonists and Antagonist Revealed by X-Ray Crystallography. *J Am Chem Soc* **2010,** 132, 11443-11445.

270. Kufareva, I.; Abagyan, R. Type-II Kinase Inhibitor Docking, Screening, and Profiling Using Modified Structures of Active Kinase States. *J Med Chem* **2008,** 51, 7921-7932.

271. Thompson, A. J.; Price, K. L.; Reeves, D. C.; Chan, S. L.; Chau, P.-L.; Lummis, S. C. Locating an Antagonist in the 5-HT3 Receptor Binding Site Using Modeling and Radioligand Binding. *J Biol Chem* **2005,** 280, 20476-20482.

272. Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-Based Discovery of β2-Adrenergic Receptor Ligands. *P Natl Acad Sci* **2009,** 106, 6843-6848.

273. Hirst, W. D.; Abrahamsen, B.; Blaney, F. E.; Calver, A. R.; Aloj, L.; Price, G. W.; Medhurst, A. D. Differences in the Central Nervous System Distribution and Pharmacology of the Mouse 5-Hydroxytryptamine-6 Receptor Compared with Rat and Human Receptors Investigated by Radioligand Binding, Site-Directed Mutagenesis, and Molecular Modeling. *Mol Pharmacol* **2003,** 64, 1295-1308.

274. Ulens, C.; Akdemir, A.; Jongejan, A.; van Elk, R.; Bertrand, S.; Perrakis, A.; Leurs, R.; Smit, A. B.; Sixma, T. K.; Bertrand, D. Use of Acetylcholine Binding Protein in the Search for Novel A7 Nicotinic Receptor Ligands. In Silico Docking, Pharmacological Screening, and X-Ray Analysis†. *J Med Chem* **2009,** 52, 2372-2383.

275. Macchiarulo, A.; Gioiello, A.; Thomas, C.; Pols, T. W. H.; Nuti, R.; Ferrari, C.; Giacche, N.; De Franco, F.; Pruzanski, M.; Auwerx, J.; Schoonjans, K.; Pellicciari, R. Probing the Binding Site of Bile Acids in TGR5. *ACS Med Chem Lett* **2013,** 4, 1158-1162.

276. Gertzen, C. G. W.; Gohlke, H. From Hansch-Fujita Analysis to AFMoC: A Road to Structure-Based QSAR. *Mol Inf* **2012,** 31, 698-704.

277. Gohlke, H.; Klebe, G. DrugScore Meets CoMFA: Adaptation of Fields for Molecular Comparison (AFMoC) or How to Tailor Knowledge-Based Pair-Potentials to a Particular Protein. *J Med Chem* **2002,** 45, 4153-4170.

278. Breu, B.; Silber, K.; Gohlke, H. Consensus Adaptation of Fields for Molecular Comparison (AFMoC) Models Incorporate Ligand and Receptor Conformational Variability into Tailor-Made Scoring Functions. *J Chem Inf Model* **2007,** 47, 2383-2400.

279. Hillebrecht, A.; Supuran, C. T.; Klebe, G. Integrated Approach Using Protein and Ligand Information to Analyze Selectivity- and Affinity-Determining Features of Carbonic Anhydrase Isozymes. *ChemMedChem* **2006,** 1, 839-853.

280. Matter, H.; Will, D. W.; Nazare, M.; Schreuder, H.; Laux, V.; Wehner, V. Structural Requirements for Factor Xa Inhibition by 3-Oxybenzamides with Neutral P1 Substituents: Combining X-Ray Crystallography, 3D-QSAR, and Tailored Scoring Functions. *J Med Chem* **2005,** 48, 3290-3312.

281. Silber, K.; Heidler, P.; Kurz, T.; Klebe, G. AFMoC Enhances Predictivity of 3D QSAR: A Case Study with DOXP-Reductoisomerase. *J Med Chem* **2005,** 48, 3547-3563.

282. Cho, S. J.; Tropsha, A. Cross-Validated r2-Guided Region Selection for Comparative Molecular Field Analysis: A Simple Method to Achieve Consistent Results. *J Med Chem* **1995,** 38, 1060-1066.

283. Tropsha, A.; Cho, S. J. Cross-Validated r2 Guided Region Selection for CoMFA Studies. In *3D QSAR in Drug Design*, Springer: 1998; pp 57-69.

284. Rodriguez, M. C.; Xie, Y.-B.; Wang, H.; Collison, K.; Segaloff, D. L. Effects of Truncations of the Cytoplasmic Tail of the Luteinizing Hormone/Chorionic Gonadotropin Receptor on Receptor-Mediated Hormone Internalization. *Mol Endocrinol* **1992,** 6, 327-336.

285. Oksche, A.; Dehe, M.; Schülein, R.; Wiesner, B.; Rosenthal, W. Folding and Cell Surface Expression of the Vasopressin V2 Receptor: Requirement of the Intracellular C-Terminus. *Febs Lett* **1998,** 424, 57-62.

286. Pankevych, H.; Korkhov, V.; Freissmuth, M.; Nanoff, C. Truncation of the A1 Adenosine Receptor Reveals Distinct Roles of the Membrane-Proximal Carboxyl Terminus in Receptor Folding and G Protein Coupling. *J Biol Chem* **2003,** 278, 30283-30293.

287. Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. Atomic-Force Microscopy: Rhodopsin Dimers in Native Disc Membranes. *Nature* **2003,** 421, 127-128.

288. Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Le Trong, I.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. Crystal Structure of a Photoactivated Deprotonated Intermediate of Rhodopsin. *P Natl Acad Sci USA* **2006,** 103, 16123-16128.

289. Homeyer, N.; Gohlke, H. Extension of the Free Energy Workflow Few Towards Implicit Solvent/Implicit Membrane MM–PBSA Calculations. *Biochim Biophys Acta, Gen Subj* **2015,** 1850, 972–982.

290. Metz, A.; Pfleger, C.; Kopitz, H.; Pfeiffer-Marek, S.; Baringhaus, K.-H.; Gohlke, H. Hot Spots and Transient Pockets: Predicting the Determinants of Small-Molecule Binding to a Protein–Protein Interface. *J Chem Inf Model* **2011,** 52, 120-133.

291. Metz, A.; Ciglia, E.; Gohlke, H. Modulating Protein-Protein Interactions: From Structural Determinants of Binding to Druggability Prediction to Application. *Curr Pharm Design* **2012,** 18, 4630-4647.

292. Gohlke, H.; Case, D. A. Converging Free Energy Estimates: MM-PB (GB) SA Studies on the Protein–Protein Complex Ras–Raf. *J Comput Chem* **2004,** 25, 238-250.

293. Miller III, B. R.; McGee Jr, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **2012,** 8, 3314-3321.

294. Craig, I. R.; Pfleger, C.; Gohlke, H.; Essex, J. W.; Spiegel, K. Pocket-Space Maps to Identify Novel Binding-Site Conformations in Proteins. *J Chem Inf Model* **2011,** 51, 2666-2679.

295. Homeyer, N.; Gohlke, H. FEW: A Workflow Tool for Free Energy Calculations of Ligand Binding. *J Comput Chem* **2013,** 34, 965-973.