UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

DIPARTIMENTO DI FARMACIA

Dottorato di ricerca in Scienza del Farmaco

XXXI ciclo

Deorphanization of Adhesion GPCRs overexpressed in Glioblastoma

Dott.ssa Sara Bottone   Tutors:

                                                 Prof. Mariano Stornaiuolo
                                                 Prof. Ettore Novellino

Coordinatore: Prof.ssa Maria Valeria D'Auria

ANNO ACCADEMICO 2015-2018

Contents

ABSTRACT

G-protein-coupled receptors (GPCRs) represent, nowadays, one of the most productive source of drug targets. Some GPCRs members are, however, poorly studied and hence cannot be targeted for therapies yet. Among these, there are more than 90 GPCRs still orphan, i.e. with no described endogenous ligand and no clearly defined function. The orphan receptors CELSR2 and Frizzled4 (Fz4), have recently drawn the attention of the scientific community in virtue of their involvement in cancer progression. CELSR2 is an adhesion GPCR (class B GPCRs) overexpressed in glioblastoma. Until now, CELSR2 endogenous ligands have not yet been identified. Fz4 belongs to class F GPCRs and it is activated by the lipoproteins WNTs and Norrin. Misregulation of Fz4 activity is involved in tumor proliferation and cancer stem cell genesis in many types of malignancies, such as glioblastoma, colorectal and breast cancer. To date, the existence of low-molecular-weight organic molecules binding to and modulating Fz4 has not been reported. Thus, the identification of ligands/modulators of CELSR2 and Fz4 activity could pave the way for new therapeutic strategies to treat cancer. In this thesis, I attempted to identify CELSR2 and Fz4 ligands by using the "Pharmacological chaperone readout", an innovative screening platform (PC-platform) that identifies ligands in virtue of their potency in affecting the tridimensional structure and the intracellular localization of a protein target.

The first chapter of the thesis describes the results obtained in the attempt of identifying new ligands of Fz4 receptor. To achieve this aim, a cell line expressing a Fz4 mutant, Fz4-L501fsX533, was generated. This frameshift mutation is responsible in vivo for the occurrence of the Familial Exudative Vitreoretinopathy (FEVR), a pathology of the retina. The resulting mutant, here referred to as Fz4-FEVR, aggregates intracellularly in the Endoplasmic Reticulum (ER) without reaching the Plasma Membrane (PM) of the cell where, in contrast, the wt receptor localizes at steady state. To identify Fz4 wt modulators, a library of organic molecules has been screened for pharmacological chaperones of Fz4-FEVR, i.e. for molecules able to rescue the folding and correct localization of Fz4-FEVR at PM. Using such read-out, the organic molecule FzM1 has been thus identified as Fz4-FEVR pharmacological chaperone. The pharmacological chaperone FzM1 acts as allosteric inhibitor of the Fz4-wt receptor, binding directly to the wt receptor and inhibiting the signalling pathway Fz4 is involved in. I also performed a structure-activity relationship (SAR) analysis using FzM1 as lead to identify the first allosteric agonist of Fz4, FzM1.8.

As discussed in the second chapter of this thesis, a specific PC-platform was developed for the identification of ligands of CELSR2. To achieve this aim, a misfolded version of CELSR2 was generated and then a library of metabolites was screened looking for compounds binding to the receptor and correcting its folding. Zebrafish eggs were used as biomass to obtain a library of metabolites which were screened for pharmacological chaperone efficiency. The outcome of the screening identified three compounds acting as modulators of CELSR2: cholesterol, PGE2 and β-carotene. They do not target the orthosteric binding site of CELSR2 but they are located in an hydrophobic region of the receptor at the interface between the TM-bundle and the lipid bilayer. The potential ability of these molecules to allosterically modulate CELSR2 function could have important implications both in physiological and pathological cell conditions.

Chapter 1

INTRODUCTION

Folding is a physical process by which proteins reach their functional three dimensional structure

Proteins are pivotal units for cell function and vitality. To be properly active, a protein must "properly fold", i.e. attain its correct three-dimensional structure. Final and correct "native-conformations" are achieved when unfolded nascent chains, emerging from the ribosomes, undergo a series of conformational changes and structural transitions that characterize their specific folding pathways[1].

The free energy endowed by a nascent chain is very high[2]. At the beginning of the folding process, entropy forces the unfolded chain to assume and switch among all the allowed structural conformations[2]. Later, during the folding process, the enthalpy generated by the interactions established between different amino acids and domains of the protein counterbalances the entropic disorder and stabilize protein native states, reducing total Free Energy to an absolute minimum[3].

The native and the unfolded states have only a minimal difference of free energy (only 5-10 kcal/mole)[4]. However, only in the native conformation a protein can exert the specific function to which it has been destined[2]. This is because only in this state the amino acids contributing to the proteins' active sites are correctly located in the space and relatively to each others.

During the folding process, proteins may assume the so called intermediate conformations[5]. Compared with the native state, these conformations are less stable, have already acquired secondary structures but miss the correct final tertiary structure. One of the partially folded, intermediate states is named the "molten globule"[4].

In vivo, most of the proteins achieve the native state in a time range from milliseconds to seconds [6]. Longer time is required, in contrast, for those proteins characterized by a complex structure (transmembrane and multi-span proteins)[7].

Proteins fold moving in an energy landscape

The energy landscape perspective is the most recent theory explaining the folding process of a protein[8,9]. It describes the free energy endowed by the polypeptide chain as a funnel-like landscape with many local minima corresponding to small energy traps (metastable states) and a single global minimum corresponding to the native state[10]. In the energy funnel, the formation of conformations with lower free energy occurs more likely than those with higher free energy and thus, the unfolded protein irreversibly moves towards the minimal free energy of the native conformation[11].

In the simplest case, the folding process involves only two states: the native state (N) and the unfolded state (U).

$$U \leftrightarrow N$$

In this case the folding process is represented by a smooth energy landscape and the folding and unfolding processes have the same running time (**Fig. 1**).
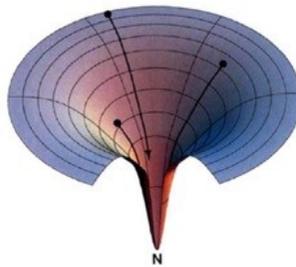
Figure 1. Energy landscape in which each point represents a protein conformation. The vertical axis represents the free energy (N stands for native state). (Adapted from Dill and Chan, 1997).

In most of cases, proteins take more time to roll down a rough funnel characterized by many small hills and valleys[12] (**Fig. 2**).
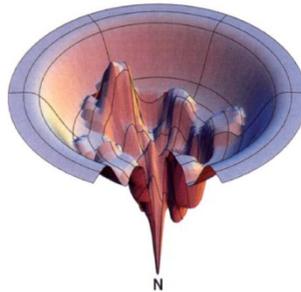
Figure 2. Energy landscape showing energetic hills and traps which hamper the protein folding progress (adapted from Dill and Chan, 1997).

In these cases, the folding process alternates slow and fast phases. The slow phase depends on the presence of local minima in which partially folded intermediates can be trapped. The energy barrier that an intermediate has to overcome to achieve the native state is very high. For this reason, the intermediate state (I) is considered a competitor of the native state[12].

$$U \leftrightarrow I \leftrightarrow N$$
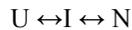
Protein Folding is a multistep process

How a protein reaches its three-dimensional conformation constitutes a central problem in biology. Anfinsen's pioneristic experiments on the *in vitro* refolding of the protein Ribonuclease I[13], suggested that all the information required for a protein to fold are contained in its primary sequence. Although these refolding experiments could be repeated with other small, single domain proteins, large multi-domain proteins normally tend to aggregate when they are refolded in vitro[14]. This reflects what happens in a cell, where most of the proteins requires cellular assistance to fold properly and efficiently[15].

Protein folding is a multistep process:

• The first step includes the formation of secondary structures (α-helix and β-sheets) which fold rapidly because they are stabilized by intramolecular hydrogen bonds[16,17].

• In a second phase of the folding process hydrophobic interactions, protein-solvent interactions and eventually disulfide bridges are established so that the hydrophilic and hydrophobic side chains can face to aqueous solvent and hydrophobic core of the protein, respectively[18].

• Finally, coarse movements of the different protein sub-domains lead to the acquisition of the final native conformation.

Despite enthalpy would push a folded protein toward an unfolded and chaotic organization of the amino-acidic chain, the native conformation is an irreversible state for a protein (at least for those kept at physiological condition of temperature, pH and ionic strength)[19]. Intra-chain interactions, hydrogen bonds, hydrophobic interactions, and disulfide bridges as well as interactions established between side chains and the solvent all help stabilizing the native conformation of the protein and contribute to protein stability[20].

Chaperones assist protein folding

Although all information to reach the native, biologically active state of a protein is present in the amino acid sequence[21], cells invest in a complex network of molecular chaperones which will assist the nascent chain and make the folding process more efficient[15]. Chaperones are nanoscale molecular machines that recognize incompletely or incorrectly folded proteins, arrest or unfold them and then either release them for spontaneous refolding or target them for degradation[22].

In the cytosol as well as in the Endoplasmic Reticulum (the two intracellular locations where protein folding occurs) two classes of chaperones can be found: general chaperones, which assist unfolded proteins by recognizing exposed hydrophobic patches, mobile loops, lack of compactness, or incomplete gycosylation, and private chaperones, which recognize a specific amino acid sequence or structure and bind devotedly to a specific protein (family), like for the chaperone RAP for the LDL-receptor family[23].

Besides chaperones, the ER possesses various enzymes that assist proteins during folding: oxidoreductases which are involved in the formation and rearrangements of disulfide bonds[24]; Peptidyl-Prolyl-Isomerases (PPIases) which catalyze cis/trans isomerization of proline residues[25]. PPIases are extremely important in protein folding and their activity increase enormously the folding rate[26]. This is because the ribosome attaches amino acids in the *trans* conformation, which requires an isomerization step for all *cis*-prolines before the final folded structure can be achieved[26].

In the ER a sensitive surveillance mechanism ensures the degradation of misfolded proteins

The Endoplasmic Reticulum hosts the folding of important proteins such as antibodies, receptors and ion channels[27]. Incorrect folding of these proteins would be extremely harmful not only for the cell, but for the entire organism. To avoid such events of misfolding, a stringent quality control system operates in the ER and discriminates between correctly folded proteins and misfolded proteins [either (partially) unfolded or incompletely assembled protein complexes][28].

Only folded proteins are exported toward their final destinations. Unfolded proteins, on the contrary, are retained in the ER by chaperones, that recognizes improperly misfolded proteins[29]. For example, protein carrying incorrect disulfide bonds are recognized as misfolded and retained for further folding, aided by molecular chaperones and folding catalysts such as protein disulfide isomerase [30]. Prolonged retention of misfolded proteins in the ER leads to ubiquitin-mediated degradation by the ER-associated degradation pathway (ERAD)[31]. In the ERAD pathway, unfolded proteins are recognized as aberrant molecules by quality control receptors. These misfolded proteins are then sorted to an ER-membrane-associated dislocation/ubiquitination complex containing adaptor proteins that help retro-translocation of the proteins into the cytosol, where they are polyubiquitinated and then degraded by the 26S proteasome[31] (**Fig. 3**).
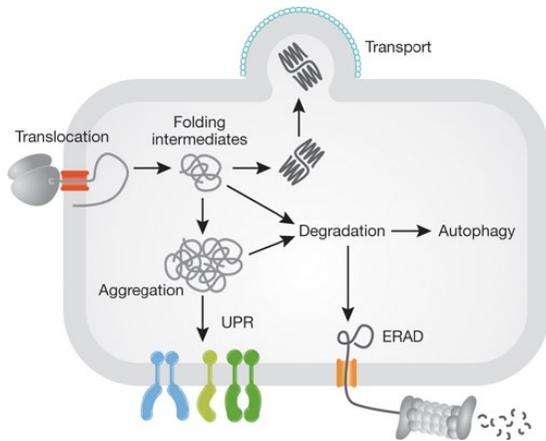
Figure 3. The general ERAD pathway for degradation of misfolded ER proteins.
ERAD is initiated on the recognition of non-native proteins as aberrant molecules by
the unfolded protein response (UPR) receptors. After being extracted from its folding
path, the substrate is retro-translocated in the cytosol where proteins are poly-
ubiquitinated and then degraded from the proteasome (adapted from Douglas M. Cyr
et al, 2009).

Accumulation of misfolded proteins can cause diseases

Despite the presence of a highly active chaperone machinery, in some cases, misfolded proteins cannot be rescued. The presence of misfolded proteins (and especially their accumulation inside or outside the cells) can cause severe diseases[32]. These pathologies are referred to as "Conformational Diseases"[32]. Most of the time, misfolding is due to mutated amino acid residues that contributes negatively to the stability of a protein[32]. More rarely, mutations have been shown to stabilize folding intermediates which in turn slow down the folding process rate[32].

The deleterious effects of misfolded proteins may depend on protein loss of function[33] as observed in Cystic Fibrosis (CF) and Congenital Nephrogenic Diabetes Insipidus (cNDI). Cystic fibrosis is caused by mutations in both copies of the gene coding for the cystic fibrosis transmembrane conductance regulator (CFTR) protein[34]. Although several mutations in the CFTR sequence have been identified, the deletion of three nucleotides coding for the phenylalanine residue at position 508 (ΔF508 CFTR) accounts for 70% of patients diagnosed with CF[35]. The ΔF508 allele of CFTR has been confirmed as a trafficking mutation that blocks maturation of the protein in the ER and targets it for premature proteolysis[36]. Similarly, congenital nephrogenic diabetes insipidus is caused by mutations of the arginine vasopressin receptor (V2R) that cause the trapping of the receptor in the ER[37]. Misfolded V2R is not expressed on the cell surface and this results in a receptor loss of function[37].

In other cases, deleterious effects of misfolded proteins may due to injurious 'gain of function' as seen in many neuro-degenerative diseases, in which protein misfolding results in the formation of harmful amyloid fibrils[33]. Among these pathologies there is the bovine spongiform encephalopathy (BSE), an infectious pathology whose infectious agent is a protein named prion scrapie or Prp^{sc}[38]. Compared to Prp^c, the wt uninfectious version of the prion protein, Prp^{sc} misfolds and forms fibrils that are toxic to the cells[39]. Moreover, Prp^{sc} is able to drive the conformational change of Prp^c proteins that are in its close proximity[40]. This conformational change can be also the result of mutations (single amino acid substitutions, deletions or insertions) of Prp^c primary sequence[41]. The beta-sheet content in Prp^{sc} structure is higher than that of Prp^c that, on the other hand, is mainly enriched in alpha-helical conformation[42].

Beta-sheet conformation seems to be a common feature of misfolded aggregated proteins and it seems the driving force for the formation of the toxic Prp^{sc} fibrils, as

well as of many other toxic fibrils[43]. Amyloid β peptide, α-synuclein and Hungtingtin form amyloid fibrils in Alzheimer's disease (AD), Parkinson's disease (PD) and Huntington's disease (HD), respectively[43]. Beta-sheet conformation also exists in many functional native proteins such as immunoglobulin[44], but the transition from alpha-helix to beta sheet is characteristic of amyloids deposits[45]. The abnormal conformational transition from alpha-helix to beta-sheet exposes hydrophobic amino acid residues and promotes protein aggregation[45] (**Fig. 4**).

There are different pharmacological approaches for conformational diseases. In some cases it is possible to protect the protein native state from misfolding and aggregation events[46]. For BSE, a therapeutic approach is being consisting in the development of antibodies which recognize and protect the native Prp^c against the negative interactions with Prp^{s46}.

In other cases it is possible to simplify the achievement of the correctly folded state. Recent studies have paved the way for the utilization of small molecules named Pharmacological Chaperones (PC)[47,48]. PC are able to enter the cell and interact with a protein to stabilize its native state and prevent its aggregation[48].
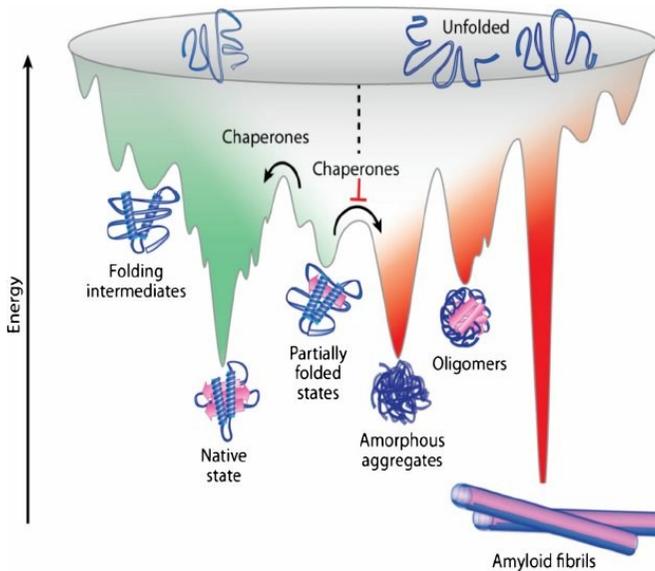
Figure 4. Energy landscape of protein folding and misfolding. Protein folding and protein aggregation are competing reactions. During the folding process, the formation of energetically favorable intramolecular interactions (green) shifts the thermodynamic equilibrium towards the native conformations. However, proteins may adopt energetically favorable but non-native conformations that can slow down the folding process rate. These partially folded or misfolded states are prone to establish intermolecular interactions (red) which result in protein aggregation (amorphous aggregates, β-sheet-rich oligomers, and amyloid fibrils). Chaperones assist in overcoming the free energy barriers. Indeed, they inhibit intermolecular interactions and therefore promote folding to the native state conformation (adapted from Muntau et al. 2014).

Ligands act as pharmacological chaperones

Binding of orthostatic or allosteric drugs to their target proteins may promote the folding of the latter by shifting the thermodynamic equilibrium towards the native conformation[49]. Native states together with folding intermediates, that are endowed with properly organized pockets, are the only conformations able to form ligand binding interactions. Ligand binding lowers the free energy of the proteins, increasing their stability[49]. Moreover, the low energy state of the ligand-protein complex minimizes off-pathway interactions preventing misfolding and aggregation[49].

Ligands have been shown to be successful in rescuing folding of membrane proteins and more specifically of GPCRs. Indeed, GPCRs explore a wide range of possible native conformations, all endowed with very similar energetic minima[50]. Such structure flexibility allows the binding of different class of ligands, all more or less able to stabilize their cognate receptor. For this reason, a large pool of candidates with a potential activity as PC have been identified. For example, β2AR-agonist promotes the folding of its cognate receptor by shifting the thermodynamic equilibrium towards the native state[51] (**Fig. 5**).
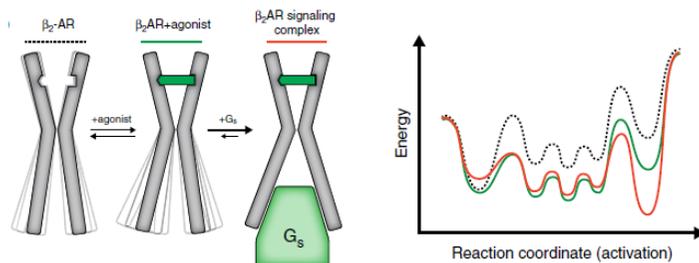
Figure 5. Agonist binding increases β2AR dynamics by decreasing the energy of intermediate and active states. G proteins further stabilize the active conformation and the formation of a signaling complex is required for the complete receptor transition to the active state (adapted from Nygaard, R et al 2012).

In vitro, there are several demonstrations of ligands acting as PC with potential for therapeutic applications.

Misfolded versions of the V2-vasopressin receptor causes the congenital nephrogenic diabetes insipidus (cNDI)[37]. Among these, the mutant R137H is retained in the ER and not expressed on the cell surface[52]. Treatment of cells transiently expressing V2 mutants with the V2R antagonist SR49059, rescued the membrane localization and signalling efficiency of the mutant receptor[53]. Moreover, the SR49059 chaperone activity was tested in cNDI patients carrying R137H mutation[54]. Two-days after administration, SR49059 had beneficial effects on urine volume and osmolality, that are clinical parameters associated with a functional V2R. However, further clinical trials of SR49059 were abandoned because of its potential interference with the cytochrome P450 metabolic pathway[54].

Another study showed that the antagonist Ipsen 17 acts as PC of its cognate receptor melanocortin 4 receptor (MCR4)[55]. Mutations in MCR4 that cause intracellular retention of the receptor are the most common monogenic cause for obesity[56]. Ipsen 17 rescued several intracellularly-retained MCR4 mutants, resulting in increased cAMP production upon agonist stimulation[55]. Thus, the ability of Ipsen 17 to rescue a wide variety of MCR4 mutants makes it particularly promising for therapeutic applications. However it has not yet been tested in animal models.

Janovick et al. showed that a non-peptide antagonist of the gonadotropin releasing hormone receptor (GnRHR) could rescue folding and activity of five intracellularly retained mutants (T32I, E90K, C200Y, C279Y, and L266R)[57]. GnRHR mutations can affect the neural regulation of the reproductive system[58]. In a knock-in mouse model of hypogonadism, treatment with GnRHR agonists restore reproductive parameters, including spermatogenesis and androgen levels[59].

Other examples of G-protein-coupled receptors (GPCRs) whose ligands were shown to rescue the function of mutated receptors are: rhodopsin[60], α1/β1/β2 adrenergic receptors[61], luteinizing hormone receptor (LH) and glucagon receptor[62].

To date, known ligands of a protein were assayed for their pharmacological chaperone potential. Recently, we showed that the potential to rescue misfolded targets can be used as a readout in screening campaigns toward the identification of new ligands[63].

RESULTS

"Pharmacological Chaperone Readout" screening platform

Recently we have set up a new screening platform named "Pharmacological Chaperone Readout" (PC-Platform)[63]. This platform allows to identify ligands in virtue of their ability to interact directly with the target and to affect its tridimensional structure (**Fig. 6**).

By using PC-platform we have identified the first allosteric modulator of the Frizzled 4 (Fz4) receptor, a member of the GPCR class F family.

To identify ligands of Fz4 we have generated a cell line expressing a Fz4 mutant, Fz4-L501fsX533, responsible *in vivo* for the occurrence of the Familial Exudative Vitreoretinopathy (FEVR), a pathology of the retina[64]. The frame-shift mutation L501fs533 dictates a change in the amino acidic sequence of the C-terminal tail of the receptor inhibiting its function. Moreover, the resulting mutant, here referred to as Fz4-FEVR, aggregates intracellularly in the ER without reaching the PM of the cell where, in contrast, the wt receptor localizes at steady state[65].

To identify Fz4 wt modulators, a library of organic molecules has been screened for pharmacological chaperones of Fz4-FEVR, i.e. for molecules able to rescue the folding and correct localization of Fz4-FEVR at PM. Using such read-out the organic molecule FzM1 has been thus identified as Fz4-FEVR pharmacological chaperone[63].

The pharmacological chaperone FzM1 acts as allosteric inhibitor of the Fz4-wt receptor, binding directly to the wt receptor and inhibiting the signalling pathway Fz4 is involved in[63]. Structure-Activity Relation analysis performed using FzM1 as lead allowed us to identify the first allosteric agonist of Fz4, FzM1.8[66].
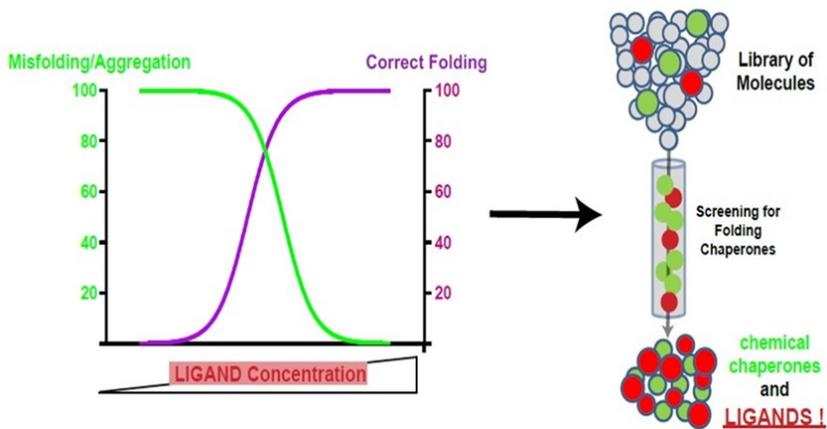
Figure 6. Schematic representation of "Pharmacological Chaperone Readout" platform.

FzM1 acts as a pharmacological chaperone of the mutant Fz4 receptor

The frameshift mutation that characterizes Fz4-FEVR C-terminal tail, affects receptor stability and induces its aggregation and retention in the ER[65]. Notwithstanding, Fz4-FEVR localization at the PM can be rescued by overexpressing the protein chaperone α-B crystalline[67]. Fz4-FEVR can be thus included among the proteins that are responsible for conformational diseases whose phenotype can be rescued by strategies aiming to improve folding. The marked difference in localization between the wt and mutant receptor makes Fz4-FEVR an ideal platform to screen for folding chaperones by offering an unambiguous readout.

In order to identify pharmacological chaperones of Fz4-FEVR, we used as biological platform Hek293 cells stably expressing (HA)-tagged wt Fz4 (HA-Fz4-WT) or Fz4-FEVR (HA-Fz4-FEVR). As expected, immunofluorescence data showed that HA-Fz4-WT was localized in the Golgi complex and at the PM, whereas HA-Fz4-FEVR appeared to be trapped intracellularly, mainly in the ER (**Fig. 7 A**). HA-Fz4-FEVR Hek293 expressing cells were treated with a library of molecules for 48h (**Supplementary Table S1**). Chaperone activity was then assayed by measuring the recovery of HA-Fz4-FEVR localization at the PM.

We showed that FzM1 rescued HA-Fz4-FEVR PM localization with an efficiency of 15% and with a half-maximum effective concentration (EC50) in the micromolar range (**Fig. 7 A, B, C**). FzM1 is an ureido derivative harbouring a phenol, a thiophene and a naphthyl functional groups (**Fig. 7 B**)

21
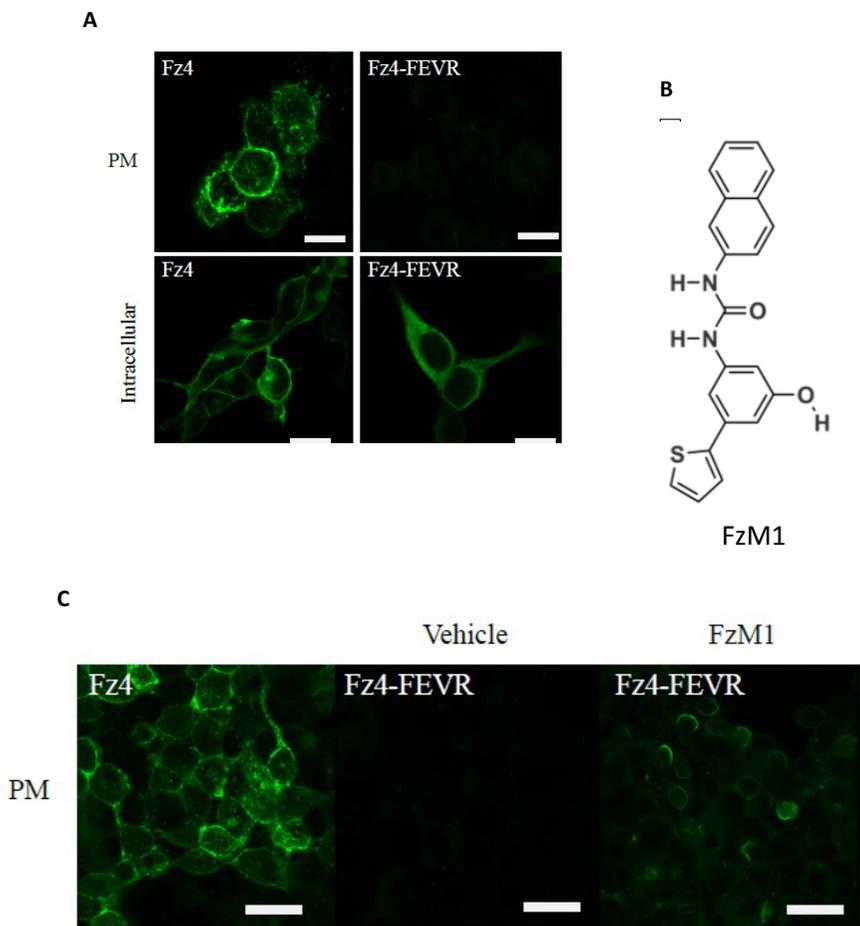
Figure 7. **FzM1 rescues Fz4-FEVR PM localization in Hek293 cells**. **(A)** Cellular localization of Fz4 and Fz4-FEVR expressed in Hek293 cells. Confocal immunofluorescence of non permeabilized (PM) and permeabilized (intracellular) cells showing PM and ER localization of Fz4 and Fz4-FEVR (green), respectively. **(B)** Chemical structure of FzM1. **(C)** Rescue of Fz4-FEVR localization at the PM upon FzM1 treatment. Confocal immunofluorescence of non permeabilized cells showing PM localization of Fz4-FEVR upon treatment with FzM1. Fz4-WT expressing cells are shown for comparison.

FzM1 binding site on Fz4-wt receptor

We identified FzM1, a molecule acting as pharmacological chaperone of the mutant receptor Fz4-FEVR. We then confirmed that the molecule selected by the screening platform indeed addresses our target. In order to do this we identified the FzM1 binding site on Fz4-wt.

To achieve our aim, we modified the original FzM1 molecule by adding a reactive alkyne moiety. This new compound was called FzM1alk (**Fig. 8**). Alkyne groups may receive nucleophilic attacks from sulfhydrylic or hydroxylic groups of amino acids and generate covalent adducts. Thus FzM1alk would bind to serines, threonines and cysteines located at its binding sites.

HA-Fz4-WT Hek293 cells were cultured in the presence of FzM1alk. Fz4 was immunopurified, digested with trypsin and analyzed by LC/MS.

HPLC profiles of samples from FzM1alk-treated cells and untreated cells were compared. These almost totally overlapped, with the exception of a fraction eluting with a retention time of 17.2 min and present only in samples obtained from FzM1alk-treated cells (**Fig. 9**). Among the peptides eluting in this fraction, the one with an *m/z* of 1,386.9 Da corresponded, with a deviation (Δmass) of 1.6 Da from the theoretical mass, to amino acids 418–426 of Fz4 ICL3 (theoretical mass of 990 Da) carrying FzM1alk (molecular weight 398.5 Da) covalently linked either to S418 or to T425 (**Fig. 10 A** theoretical mass [M-H-FzM1alk]+ of 1,388.5 Da). A peptide with this *m/z* was present exclusively in the FzM1alk-treated sample. In contrast, the peptides that co-eluted with it (**Fig. 10 B**) were not unique as they were also present in the untreated sample in similar abundances, and thus they were not further considered.

Thus, the spectra indicated that FzM1alk interacted with Fz4 ICL3 by contacting at least S418 or T425.
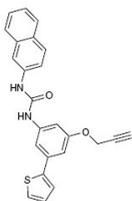
Figure 8. FzM1alkchemical structure.
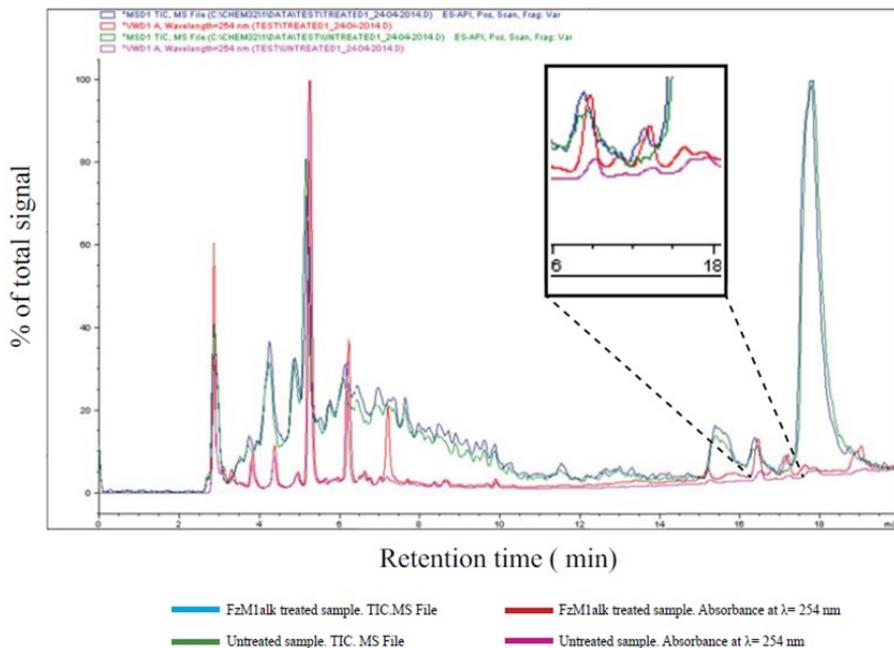
FzM1alk

Figure 9. LC/MS analysis of FzM1alk treated Fz4. HA-Fz4-wt Hek293 cells were treated with FzM1alk (10 μM) for 24 hours to be then lysed. After immunoisolation, Fz4 was digested with Trypsin and analyzed by LC/MS. Comparison between FzM1alk treated vs untreated Fz4 MS spectra enables to identify peptides presenting FzM1alk covalently attached (ΔM= + 398.5 Da). HPLC profiles of FzM1alk treated (blue curve) and untreated samples (green curve). Absorbance at λ=254 nm FzM1alk (red curve) and untreated (magenta curve) samples are reported. The boxes shows an enlargement of the region (retention time from 16.0 to 18.0 min) where a difference in the two HPLC profiles could be identified**;**

24
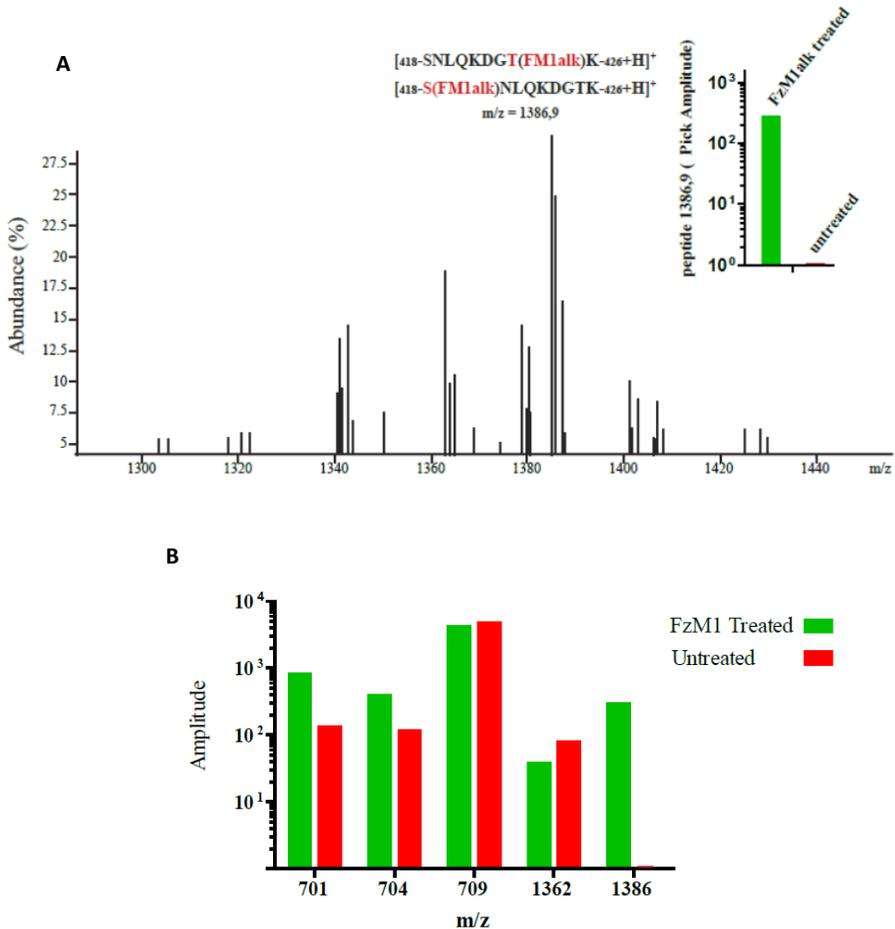
Figure 10. LC/MS analysis of FzM1alk treated Fz4. (A) MS spectra of the FzM1alk treated sample fraction eluting at retention time of 17.2 minutes. **(B)** Amplitudes of peptides with the indicated m/z in FzM1alk treated (green bars) and untreated (red bars) samples. Among the peptides eluting between 16.0 and 18.0 min of retention time, the pick corresponding to m/z of 1386-1387 was the only one to be exclusively present in FzM1alk treated sample.

25

Binding of FzM1alk or FzM1 to Fz4 ICL3 should cause a change in solvent accessibility in this region of the receptor. These changes can be highlighted by measuring the rate of hydrogen-to-deuterium exchange (HDX) at this site.

In the absence of FzM1, almost the entire receptor exchanged its hydrogen atoms with deuterium (with the exception of a helix of the ECD domain, amino acids 154–164; **Fig. 11**). As expected, the C-terminal tail and the ECD of the receptor showed the higher rate of exchange compared to the TMD (**Fig. 12 A**). Upon FzM1 binding, solvent accessibility of ICL3, TM5–7 and the C-terminal tail of the receptor was reduced, whereas the rest of the protein became more solvent accessible (**Fig. 12 B**). The burial of the ICL3 from solvent upon ligand binding appears compatible with ICL3 being contacted by FzM1 and with a change in the conformation of this loop. A burial of the C-terminal tail can be also hypothesized from the HDX data. The latter could be due to a tighter interaction of the tail with ICL3 or, more likely, as seen for many other GPCRs, with the tail interacting with the detergent micelles.

We envisaged that the diminished solvent accessibility of ICL3 upon FzM1 binding could have been the mechanism behind the folding chaperone effect of the molecule on Fz4-FEVR. To verify this, we replaced ICL3 residues with alanines to affect the fold of this ICL and its interactions with the solvent. We mutated some amino acids in the FZ4-FEVR ICL3 to then test the effect of such mutations on the PM localization of the receptor. Notably, the resulting mutant receptor regained PM localization only when T425 and K426 of FZ4-FEVR were mutated (**Fig. 13).** Thus, by interfering with the interactions established between the solvent and the amino acid involved in FzM1 and FzM1alk binding, we obtained the rescue of Fz4-FEVR folding and localization.

By identifying the FzM1alk binding site at the ICL3 region, we proved that this pharmacological chaperone is indeed a ligand of Fz4-wt. FzM1 or FzM1alk binding induces a conformational change in Fz4-wt that reduces the solvent accessibility at the ICL3. Moreover, they act as pharmacological chaperones by stabilizing the Fz4-FEVR receptor and hampering its aggregation. Though the FEVR C-terminal tail induces Fz4-FEVR aggregation, ICL3 can counterbalance the effect of the tail, influencing the susceptibility of the full receptor to the aggregation. When FzM1 binds to ICL3 or T425A-K426A are mutated, the negative effects of the C-terminal tail on the entire receptor are compromised and the aggregation process is abolished.

.
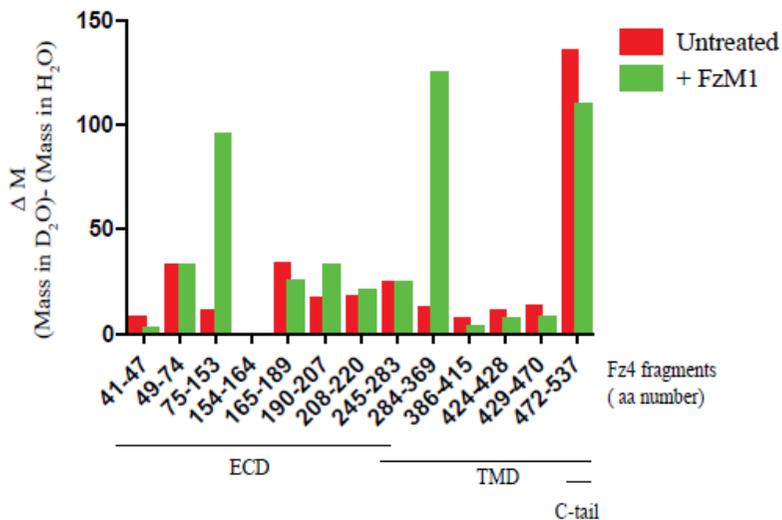
Figure 11. Analysis of FzM1 binding by HD exchange. Immuno-purified Fz4-wt, treated or not with FzM1, was incubated on ice with D_2O to allow H to D exchange. The protein was then digested at low pH with Formic Acid and run on LC/MS (pH of the run less than 2.5 to minimize back exchange). Hydrogen to deuterium exchange rate (difference of peptide masses in D_2O and in H_2O) of Fz4 peptides in the presence (green bars) or in the absence (red bars) of FzM1. Peptides are indicated by their aa numbering and by their position in Fz4 (Extracellular Domain, (ECD)), Transmembrane Domain (TMD) and cytosolic tail (C-tail);
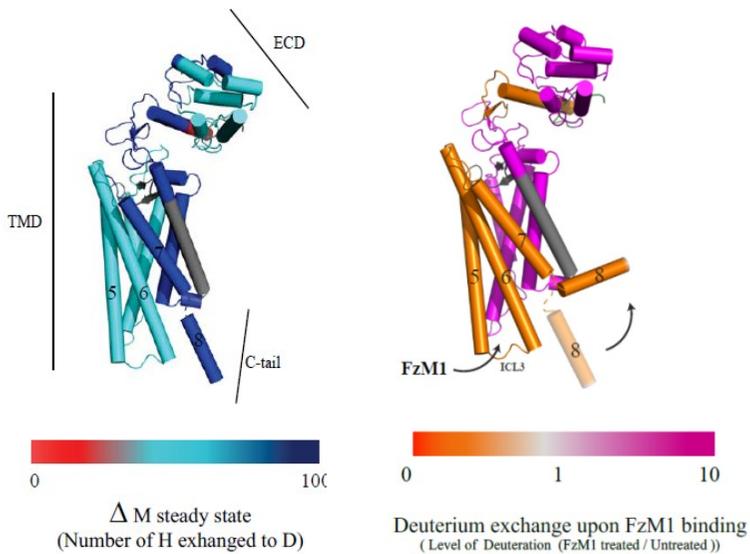
Figure 12. **Analysis of FzM1 binding by HD exchange**. **(A)** Rate of deuterium exchange (change of mass upon incubation with D_2O, Da) in untreated Fz4. **(B)** Difference in rate of deuterium exchange upon FzM1 binding. The arrow indicates a possible movement of the C-tail toward the lipid bilayer. Segments in gray represent not detected peptides.
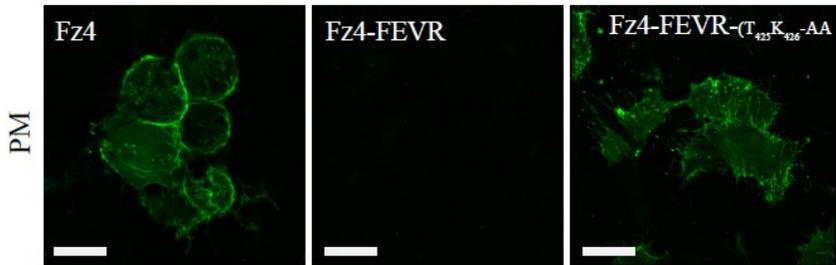
Figure 13. Cellular localization of HA-Fz4, HA-Fz4-FEVR and HA-Fz4^{T425AK426A}-FEVR transfected in Hek293 cells. Immunofluorescence of non permeabilized cells showing the PM localization of HA-Fz4-WT and HA-Fz4^{T425AK426A}-FEVR but not of HA-Fz4-FEVR. Scale bars, 10 μm.

FzM1 is an allosteric inhibitor of Fz4

By identifying the FzM1 binding site, we proved that this pharmacological chaperone is a ligand of Fz4-wt. Next, we confirmed that FzM1 modulates Fz4-wt activity.

Fz4 signaling pathway is activated by Wnt ligands. Wnt binding to Fz4 recruits the scaffold protein Dishevelled (Dvl) at the PM[68]. The complex Wnt-Fz4-Dvl inactivates the Adenomatous Polyposis Coli (APC) destruction complex [formed by Glycogen synthase kinase 3 (GSK3β), Axin, and Adenomatous Polyposis Coli] and leads to the intracellular accumulation and nuclear translocation of β-catenin. In the nucleus, β-catenin regulates T-cell factor /lymphoid enhancer factor (TCF/LEF) dependent transcription of Wnt/ β-catenin target genes like *cyclin D1*, *c-myc,* and *lgr5*[68].

We tested the ability of FzM1 to modulate the Wnt/β-catenin pathway with a luciferase assay that measures the TCF/LEF dependent transcription activation. We performed the assay on Hek293 cells transfected with HA-Fz4-wt. These cells do not express Fz4 endogenous ligands and, as expected, did not have a TCF/LEF basal activity. However TCF/LEF activity can be induced upon treatment with Norrin, an high affinity ligand of Fz4[69].

In absence of the Fz4 agonist Norrin, FM1 did not induce the TCF/LEF dependent transcription (**Fig. 14 B**). Thus, FzM1 did not act as a Wnt/ β-catenin pathway activator.

On the contrary, after FzM1 treatment, Norrin-dependent TCF/LEF activation was abolished, suggesting that this molecule could be an inhibitor of the Wnt–β-catenin pathway (**Fig. 14 A**).

We further confirmed the FzM1 activity as Wnt pathway inhibitor by using LiCl, a GSK-3β inhibitor, that causes the β-catenin nuclear translocation (**Fig. 14 B**).

We also hypothesized the molecular mechanism behind Wnt pathway inhibition by FzM1. ICL3 is the region involved in the Dsh recruitment by Fz4. Contacting ICL3, FzM1 induced a conformational change of Fz4 which no longer was able to recruit Dvl, ultimately blocking β-catenin C-terminal phosphorylation played by CK1 and the consequent β-catenin nuclear translocation.

A

B
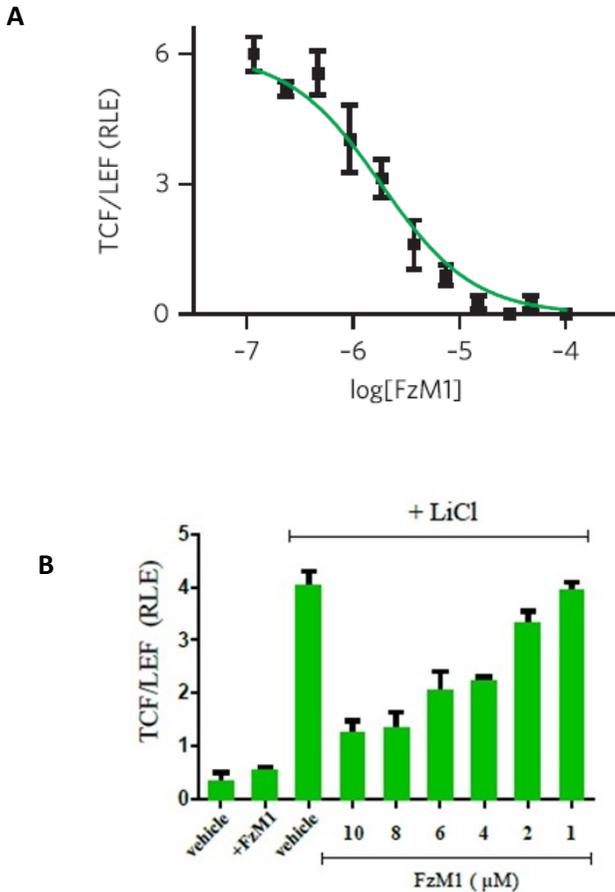
Figure 14. (**A**) Dose-response curve for FzM1 inhibition of Wnt signaling. Hek293 cells were transiently transfected with HA-Fz4-WT and treated with Norrin (40 ng ml^{-1}) in the presence of the indicated concentrations of FzM1. Measurements were done in triplicate and are shown as mean ± s.d. RLE, relative luciferase expression. (**B**) Hek293 cells were transiently transfected with HA-Fz4-WT treated with LiCl (30 mM) and or with the indicated concentrations of FzM1.

FzM1.8, a FzM1 derivate compound, acts as an allosteric agonist

We showed that FzM1 is an allosteric inhibitor of Fz4. With the aim to generate new Wnt/β-catenin pathway modulators we used FzM1 as lead compound to obtain a small library of FzM1 analogs.

To evaluate the activity of the new compounds we used Hek293 cells transiently expressing both Fz4 and a Wnt reporter construct, the latter presenting the coding sequence of GFP under the control of an optimized Wnt Responsive Element (WRE-wt). Compared to mock transfected cells, transient expression of FZD4 did not induce activation of the Wnt reporter construct indicating a low level of basal activity of FZD4 in the absence of ligands. In contrast, activation of Wnt/β-catenin canonical pathway by the GSK3-β inhibitor LiCl (**Fig. 15 A-C**) induced GFP expression and increases the fluorescence of the cells.

We showed that the replacement of FzM1 thiophene with a carboxylic group transformed FzM1 in to a new allosteric agonist of Fz4. In the absence of any orthosteric ligand, the new compound 3-Hydroxy-5-(3-(naphthalen-2-yl)ureido) benzoic acid (FzM1.8, **Fig. 16 and Fig. 17 A, B**) was able to increase WRE activity in Fz4 expressing Hek293 cells.

Similarly to FzM1, FzM1.8 did not increase WRE activity in cells not expressing Fz4 (**Fig. 17 C-D**) indicating that the effect of the compound is Fz4 dependent. Moreover, FzM1.8 did not induce GFP expression in cells transiently expressing both Fz4 and either a mutated nonfunctional WRE reporter (WRE-mut, **Fig. 17 E-F**). These results confirm that FzM1.8 does not act as general transcriptional activator and that, instead, its activity relies on a functional WRE. Interestingly, in the range of concentrations tested, FzM1.8 modulated WRE activity with values fitting hormetic dose-response curves (described in Equation1). We measured both a stimulatory (EC50 of activation ($\log EC50act$) ± s.e.m.=-6.4 ± 0.2, n=16) and an inhibitory activity (EC50 of inhibition ($\log EC50inh$) ± s.e.m. = -5.5± 0.1, n=16) (**Fig. 17 B**).

Despite presenting a switch in its activity, FzM1.8 contacts the same binding sites of the original lead. As already shown for FzM1, Hek293 cells transfected with the Fz4 mutant FZD4T425A only partially responded to FzM1.8 treatment (**Fig. 18**). Moreover, HDX analysis reveals that FzM1.8 binding reduced solvent accessibility of ICL3. Interestingly, both the ligands affected the conformation of TM6, with FzM1.8

increasing the solvent accessibility of this section of the receptor more than FzM1.8 (**Fig. 19**).
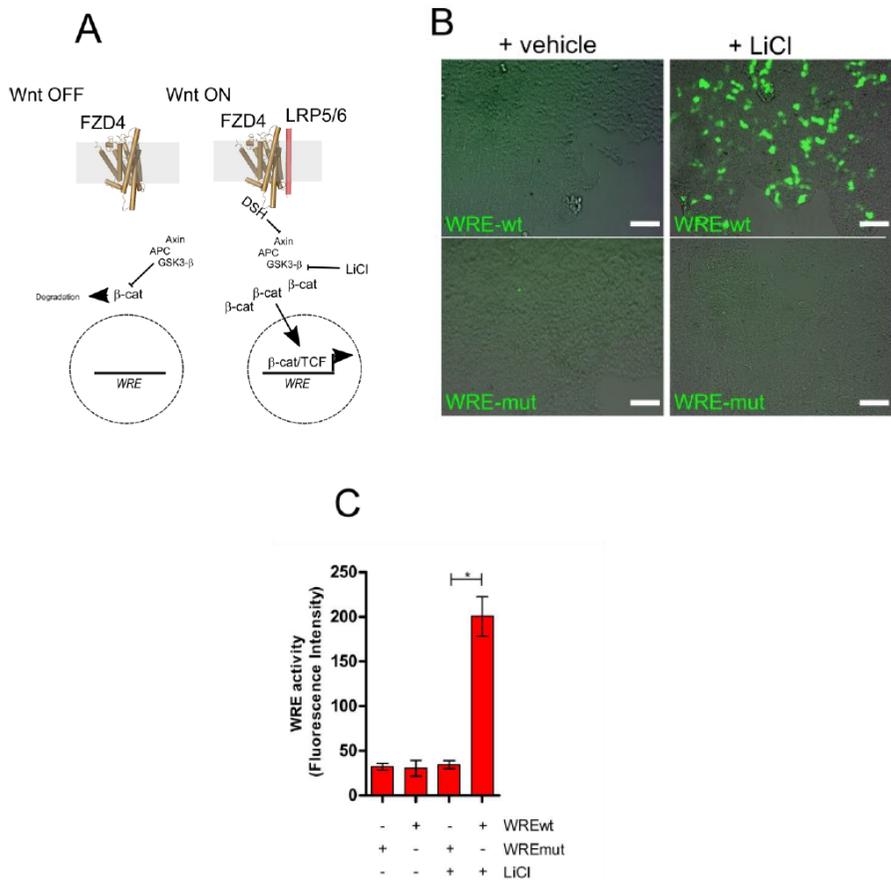
Figure 15. Experimental platform used to measure Wnt/β-catenin pathway activity.
(**A**) Schematic cartoon depicting Fz4 involvement in Wnt/β-catenin pathway; (**B**) Hek293
cells transfected with the constructs WRE-wt and WRE-mut (Merge of Bright Field and
GFP channel, Bar = 50μm). Activation of the pathway by LiCl induces GFP expression in
cells transfected with WRE-wt but not in those transfected with mutated nonfunctional
WRE (WRE-mut); (**C**) The histogram shows the increase in WRE activity after treatment
with LiCl. [Values are reported as mean +/- s.e.m. (n=5) *P<0.05].
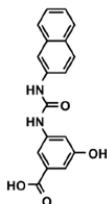
Figure 16. Chemical structure of FzM1.8
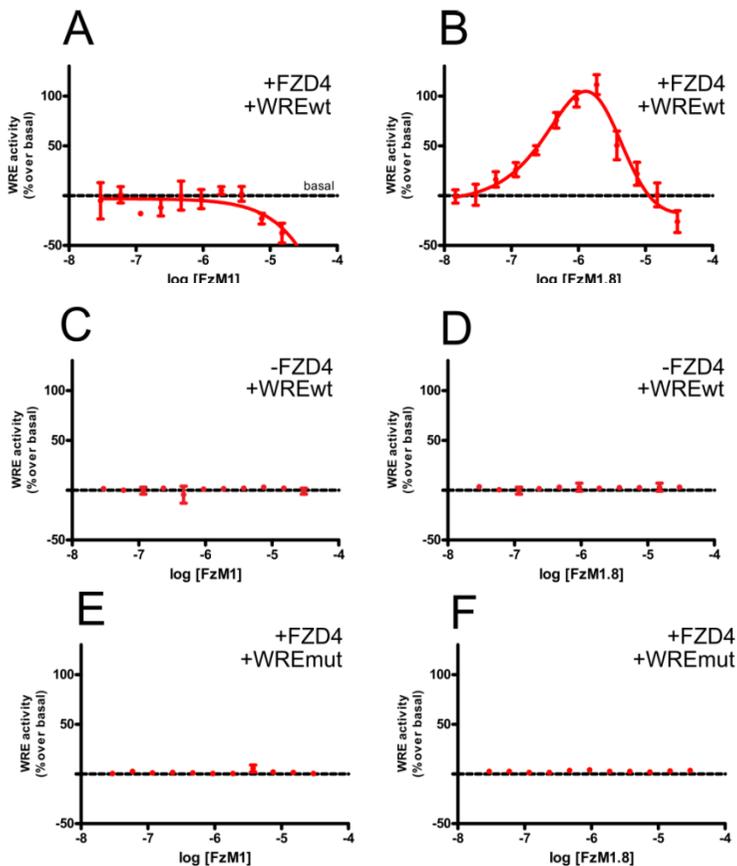
Figure 17. FzM1.8 is an allosteric agonist of Fz4 and activates the Wnt/β-catenin pathway. (**A-B**) Dose-response curves for FzM1 and FzM1.8 modulation of Wnt/β-catenin pathway in Hek293 cells co-transfected with both Fz4-wt cDNA and the WRE-wt reporter construct. Values indicate changes in the WRE activity (expressed as the percentage of change over basal activity). (**C-D**) Dose-response curves for FzM1 and FzM1.8 modulation of WRE activity in Hek293 in the absence of Fz4; (**E-F**) Dose-response curves for FzM1 and FzM1.8 modulation of WRE activity in Hek293 cells co-transfected with Fz4-wt and WRE-mut.

35
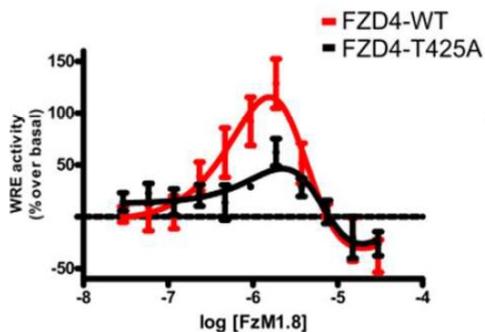
Figure 18. Dose-response curves for FzM1.8 modulation of WRE activity in Hek293 cells expressing Fz4-wt (red curve) or FZD4-T425A (black curve);
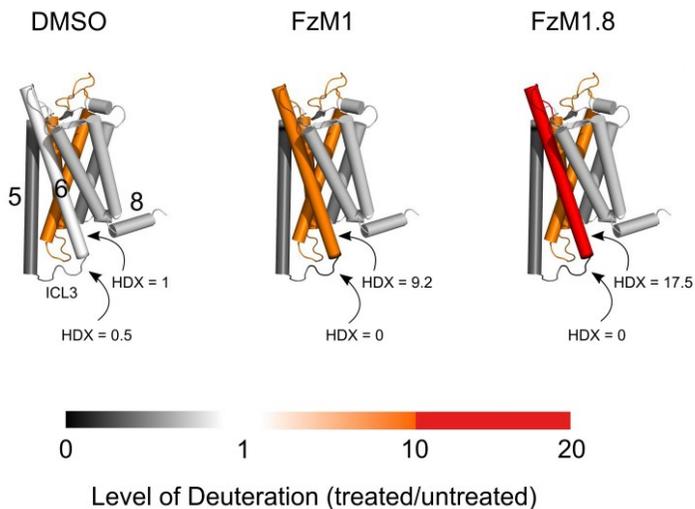
Figure 19. FzM1.8 addresses Fz4 ICL3. Solvent accessibility of Fz4 upon binding to DMSO, FzM1 or FzM1.8 measured by HDX analysis. The color scale reflects the level of deuterium exchange measured as difference in peptide mass in the presence or in the absence of vehicle or of ligands. The arrows indicate the ICL3 and the TM6 of Fz4 and the level of deuteration upon vehicle, FzM1 or FzM1.8 binding. Numbers indicate TMD helices (5–6) and the C-terminal tail helix (8).

FzM1.8 mechanism of action

We showed that FzM1.8, upon binding to Fz4, initiates a signaling route resulting in an increased WRE activity. Thus, we started testing if FzM1.8, by either affecting Fz4/Dvl interaction or by inhibiting β-catenin proteasomal degradation, was acting via Wnt canonical pathway (**Fig. 20 A**).

As shown in Figure **20 B**, in the presence of FzM1.8, Fz4 is still able of recruiting Dvl. Furthermore, differently from Wnt ligands or LiCl, FzM1.8 did not increase intracellular levels of β-catenin (**Fig. 20 C**). Finally, in the presence of CIK 7, a pan-inhibitor of Casein Kinase I[70], a protein involved in activation of the Fz4/Dvl complex[71], FzM1.8 stimulatory activity was mostly preserved (**Fig. 20 D**). Thus FzM1.8/Fz4 complex stimulates WRE activity but does not affect proteins taking part in the Fz4/Dvl/APC axis.

To explain FzM1.8 activity, we thus hypothesized the involvement of a non-canonical pathway. The only pathway that has been so far described as positive modulator of the Wnt canonical pathway is the one involving Phosphatidylinositol-4,5-bisphosphate 3-kinases (PI3Ks)[72]. PI3K has been already shown to affect Wnt canonical pathway on multiple levels. In differentiated myofibers, binding of the orthosteric ligand Wnt7a to Fz7 directly activates PI3K pathway and induces myofiber hypertrophy[73]. Moreover, via AKT, PI3K inhibits GSK3-β and leads to β-catenin stabilization[74]. Alternatively, AKT has been shown to promote the activity of the histone acetyltransferases CBP and P300, both trans-activators of the β–catenin/TCF transcription complex[75]. These remodel the chromatin-bound to Wnt responsive genes, ultimately boosting the activity of the β–catenin/TCF transcription complex. Indeed, in the presence of the PI3K inhibitor LY294002[76], FzM1.8 lost its ability to induce WRE activity (**Fig 21 A**). WB analysis of cell treated with FzM1.8 revealed an increase in AKT phosphorylation at Thr308 upon 30 min and 2 hour of incubation with the ligand, confirming the involvement of PI3K in the pathway activated by FzM1.8. Differently, we did not measure an increase in AKT phosphorylation at Ser473 (**Fig. 22 A, B**). Finally, Hek293 cells expressing both Fz4 and the AKT dominant negative mutant AKT K179M were insensitive to FzM1.8, further proving the involvement of the protein kinase in the WRE activation elicited by the ligand (**Fig. 22 C**).

Since FzM1.8 does not induce β-catenin stabilization (**Fig. 20 C**) we tested whether WRE activation by FzM1.8 was dependent on CBP/p300. In the presence of the

CBP/p300 specific inhibitor I-CBP 112[77], FzM1.8 activity was drastically reduced (**Fig. 21 B**) confirming the involvement of PI3K and CBP/p300 downstream the pathway activated by FzM1.8. The involvement of CBP/p300 indicates also that the Fz4/PI3K axis elicited by FzM1.8 is "ancillary" to the canonical Wnt pathway. Even if it bypasses the Wnt/Fz4/APC canonical route, the Fz4/PI3K axis acts ultimately on the nuclear pool of β-catenin (**Fig. 21 C**).
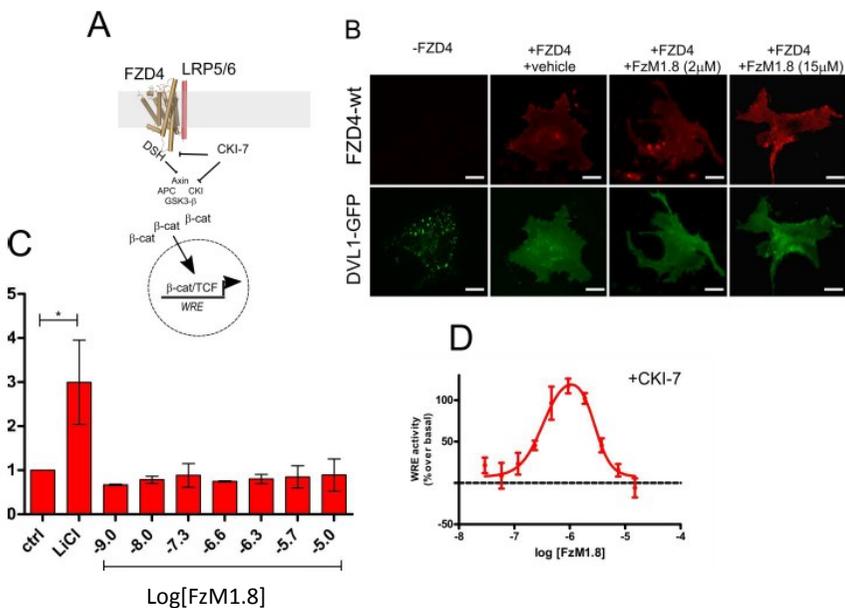
Figure 20. FzM1.8 does not affect canonical Wnt/β-catenin pathway. (**A**) Schematic cartoon depicting Fz4 involvement in Wnt/β-catenin pathway and the molecular target of the Casein Kinase I inhibitor CKI-7; (**B**) Non-confocal immunofluorescence showing recruitment of Dvl-GFP at the PM of HuH7.5 cells, in the presence or not of Fz4-wt and of the indicated amount of FzM1.8; (**C**) The histogram shows the intracellular levels of β-catenin measured in Fz4 expressing Hek293 cells upon treatment with LiCl (30 mM) or the indicated amount of FzM1.8; (**D**) Dose-response curves for FzM1.8 modulation of WRE activity in cells expressing Fz4-wt in the presence of CKI-7; [Values are reported as mean +/- s.e.m. (n=5) *P<0.05].

39
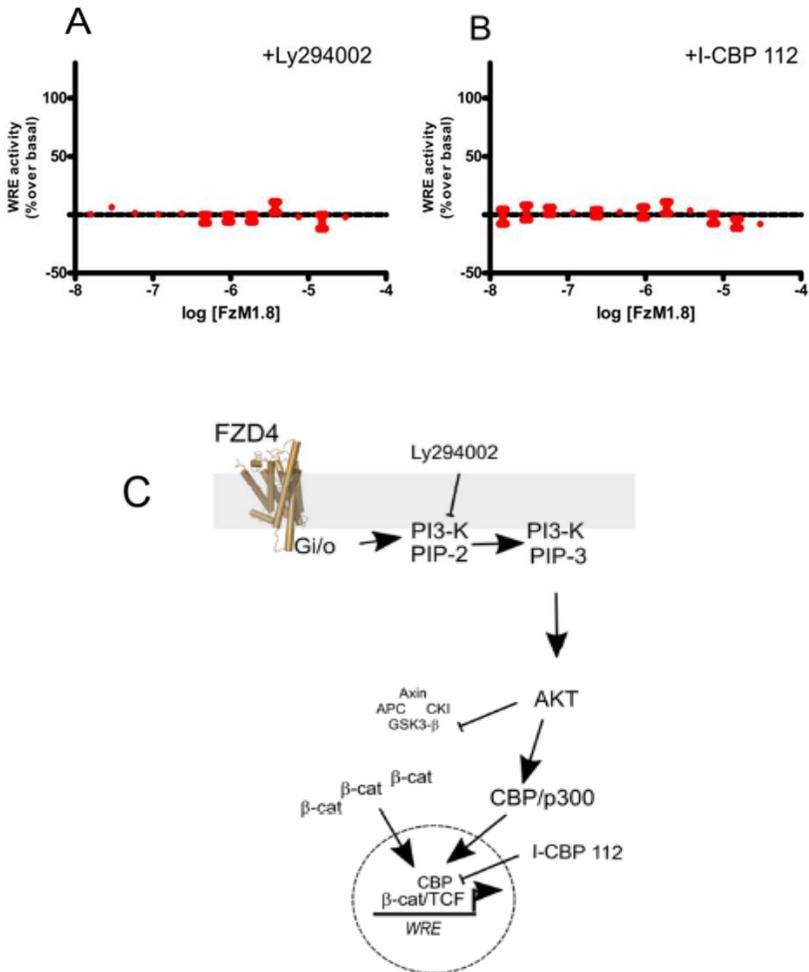
Figure 21. Fz4/FzM1.8 complex activates PI3K. (**A-B**) Dose-response curves for FzM1.8 modulation of WRE activity in cells expressing Fz4-wt in the presence of either PI3K (**A**) or CBP/p300 (**B**) inhibitors; (**C**) Schematic cartoon depicting Fz4 involvement in Wnt/PI3K pathway. [Values are reported as mean +/- s.e.m. (n=5).*P<0.05].
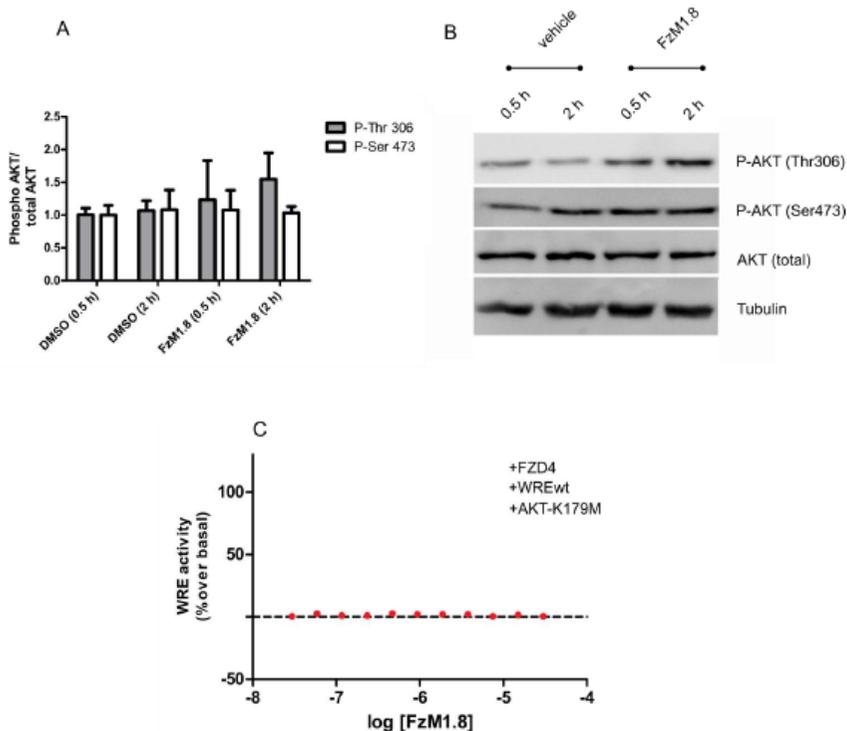
Figure 22. FzM1.8 signaling involves AKT. (A) Intracellular levels of AKT phosphorylated at Thr306 (P-Thr306), or at Ser473 (P-Ser473) in Hek293 cells transfected with Fz4 and treated for the indicated period of times with FzM1.8 (2 μM) or with the corresponding amount of vehicle (DMSO). In **A** the ratio between Phosphorylated and total AKT are indicated for each condition. [Values are reported as mean +/- s.e.m. (n=3)]. Panel **B** shows sections of gels representative of the replicates used for the quantification reported in **A**. The intensity of each band was normalized by decorating each filters with an anti α-tubulin antibody. Only the relevant portion of the gels are shown. (**C**) Dose-response curves for FzM1.8 modulation of WRE activity in Hek293 cells co-transfected with Fz4-wt and with the dominant negative AKT mutant AKT-K176M. [Representative of three independent experiments].

41

Effects of FzM1 and FzM1.8 on tumor cells

We also investigated if FzM1 and FzM1.8 by modulating the Fz4 receptor activity could affect cancer cells growth. Indeed, the intracellular signaling pathway activated by Fz4 culminates in the activation of genes that have been related to tumor cell survival, differentiation and invasiveness[78].

As first we looked at the effects of FzM1 treatment on growth, differentiation and migration of U87MG glioblastoma cells. Fz4 expression in these cells has already been shown to relate to invasiveness and the differentiation state of the cells[79]. This cell line, when cultured in dish, arranges to form neurospheres that are spherical cell clusters containing cells positive for Nestin, a neural stem cell marker (**Fig. 23 A, B**). Treatment with FzM1 hampers neurosphere formation and induces cells to acquire a more elongated or neuronal appearance (**Fig. 23 and Fig. 24)**. Moreover, FzM1 reduces the number of cells positive for Nestin, suggesting that, upon FzM1 treatment, U87MG cells transitioned toward a more differentiated phenotype (**Fig. 25 A, B**).

We confirmed the ability of FzM1 to act as inhibitor of tumor cell growth in intestinal Caco2 cell line. Similarly to U87MG cells, Caco2 cells are known to rely on an active Fz4-dependent pathway to maintain their undifferentiated proliferative state. FzM1 treatment promotes differentiation of Caco2 cells as shown by the marked change in their morphology (increase in cytoplasm volume) and by the expression of the differentiation marker E-cadherin (**Fig. 26 A, B**).

We also evaluated the effects of FzM1.8 treatment on Caco2 cell line. Differentially from FzM1, FzM1.8 does not change the differentiation state of these cells. However, the treatment causes an increase of the cells duplication rate (**Fig. 27 B**). Thus, we investigated if FzM1.8 influenced the Fz4-depedent gene regulation. We showed that FzM1.8 causes a Fz4-dependent upregulation of genes stimulating cell proliferation such as *c-myc* and *cyclin-D* (**Fig. 27 A, B**). On the contrary, FzM1.8 treatment does not induce changes in the colon stem cell marker *lgr5*, indicating the FzM1.8 actives Fz4 pathways that stimulates proliferation preserving stemness of the cells (**Fig. 27 A, B**). It has been further showed that the ability of FzM1.8 to activate Fz4 plays an important role in tissue repair lung cells[66,80]. Indeed, Fz4 is a positive modulator of cell proliferation and wound healing, crucial processes involved in tissue repair.

A
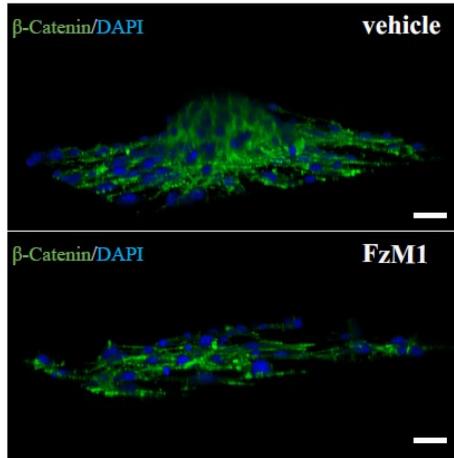
B
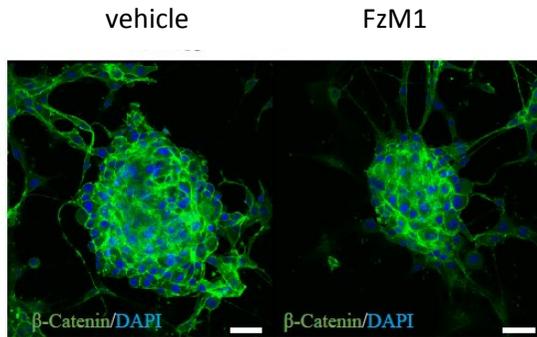
Figure 23. Effects of FzM1 treatment on U87MG glioblastoma cells. (A)
U87MG cells were cultured in presence of vehicle (DMSO 0,1%) or FzM1 (10
µM). β-catenin is in green and nuclei are in blue (DAPI). Immunofluorescence
shows vehicle-cells form peculiar neurospheres; upon compounds treatment they
disassemble. Scale bar: 20µm. **(B)** U87MG cells immunofluorescence, β-catenin
is in green and nuclei are in blue (DAPI). FzM1 treatment reduces dimension of
neurospheres. Scale bar: 20µm.

vehicle   FzM1
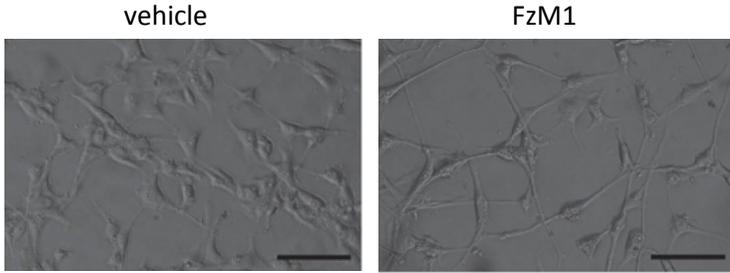
Figure 24. Morphological effect of FzM1 on U87MG cells . FzM1 (10 µM)
affects U87MG cell morphology. Upon treatment, cells acquire a more
differentiated phenotype. Bright field images show nuclei. Scale bar: 100 µm

A

vehicle   FzM1

B
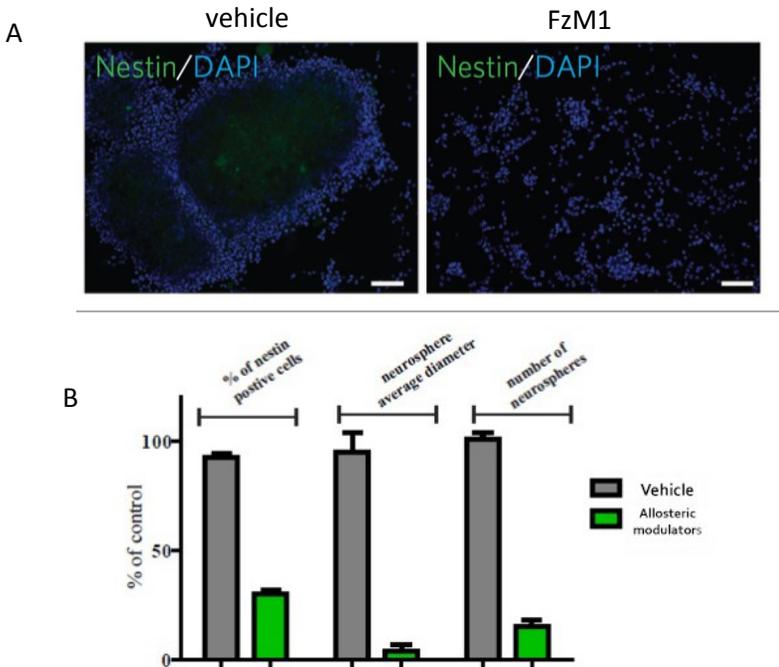
Figure 25. **FzM1 influences on U87MG differentiation. (A)** Immunofluorescence
shows neural stem cell marker nestin in green and nuclei in blue (DAPI). FzM1
treatment reduces the number of nestin-positive U87MG cells. Scale bar: 100 µm (**B**)
Upon FzM1 treatment U87MG cells show radical decrease of: % of Nestin-positive
cells, neurosphere average diameter and number of neurospheres.

vehicle   FzM1

A

B
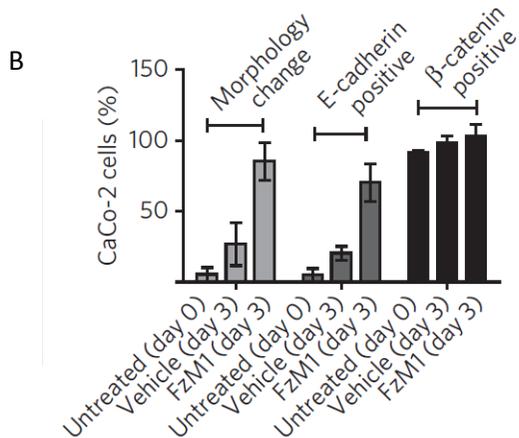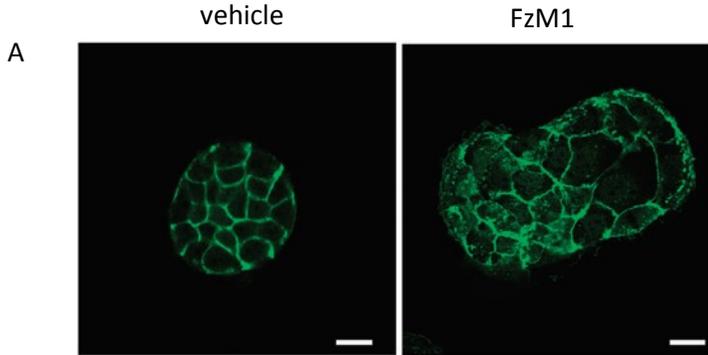
Figure 26. **Effects of FzM1 on Caco2 differentiation.** **(A)** FzM1 accelerates CaCo-2 cell differentiation. CaCo-2 cells were treated with vehicle (DMSO 0.1%, 3 d) or with FzM1 (10 μM, 3 d). Upon treatment, cells acquire a more differentiated phenotype (β-catenin is in green). Scale bars, 20 μm. **(B)** Quantification of differentiation of CaCo-2 (number of cells changed in morphology, E-cadherin–positive and β-catenin–positive cells) upon treatment with FzM1 (10 μM, 3 d); $n = 3$, data represent mean ± s.e.m.
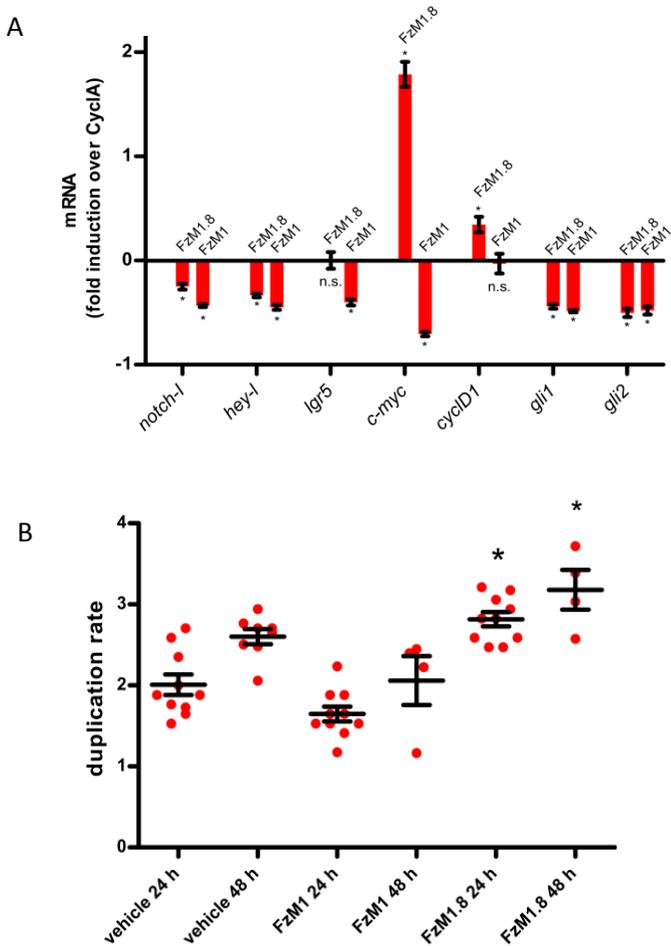
Figure 27. **FzM1.8 increases proliferation of undifferentiated colon cancer cells.**
(A) The effect of FzM1.8 or FzM.1 on the expression levels of genes regulated by
Fz4 signalling pathway. Genes stimulating cell proliferation such as *c-myc* and
cyclin-D are up-regulated. **(B)** Duplication rate of human colonic biopsies in culture
medium supplemented with FzM1, FzM1.8 or with vehicle.

DISCUSSION

There are several evidences of the ability of ligands to act as pharmacological chaperones in rescuing the folding of their target proteins. We proved that, by transitive property, the potential to rescue misfolded targets can be used as a readout in screening campaigns toward the identification of new ligands. We screened a library of small molecules for their ability to rescue the ER-trapped folding defective mutant of Fz4, Fz4-FEVR. In this way we identified FzM1, the first allosteric modulator of Fz4 that physically interacts with the ICL3 of Fz4. In our opinion, the change in the conformation of ICL3 upon FzM1 binding is the mechanism behind the pharmacological chaperone activity it exerts on Fz4-FEVR. Fz4-FEVR bound to FzM1 is more resistant to receptor aggregation and is properly transported to PM.

Interestingly, we showed that FzM1.8, that is an analog of FzM1, is not able to act as pharmacological chaperone of Fz4-FEVR. However, FzM1.8 acts as a positive allosteric modulator of Fz4-wt. Thus, the PC-platform seems to select ligands binding the inactive state conformations (such as antagonist and negative allosteric modulator) rather than the active state conformations of GPCRs (such as agonist and positive allosteric modulators). The active-state conformations of GPCRs are higher in energy (less stable) than inactive-state conformations[50]. Moreover, it has been shown that agonist binding even does not stabilize a unique native conformation[51]. In contrast, it is associated with conformational heterogeneity that may be important to regulate several alternative pathways[51]. In our opinion, these could be the reasons why agonist and positive modulators fail in counterbalance the mutant protein tendency to aggregate and have been thus never identified with our PC platform.

PC-platform represents a useful alternative to traditional screenings. It might be especially useful in the case of protein targets difficult to purify in large quantities and proteins with unclear, unknown or even absent connections to signaling cascades, such as orphan GPCRs. Furthermore, we have already shown that the applicability of the platform can be extended also to the receptors for which mutations inducing misfolding and aggregations are not known. Indeed, we proved that appending the mutated tail of Fz4-FEVR to other TM proteins is sufficient to determine the unfolding of the chimeric proteins and their ER retention[65]. Finally, GPCRs would represent good candidates for PC-platform. Indeed, they are highly dynamic and can exist in a multitude of distinct conformations that allow the binding of different class of ligands.

MATERIALS AND METHODS

Reagents. Salts and organic solvents are from Sigma Aldrich (USA) and Applichem (Germany), are handled according to manufacturer instructions.

Cell culture. Hek293 cells and U87MG were cultured in RPMI medium and DMEM (Dulbecco's Modified Essential Medium), respectively, supplemented with 10% FBS (Fetal Bovine Serum), Glutamax and Pen/Strep antibiotics. Transfections were performed with polyethylenimine (PEI) according to the manufacturer's protocol. HA-Fz4-WT cell clones are obtained by transfecting Hek293 cells with the corresponding cDNA cloned in pcDNA 5.0, kind gift of B.T. Mac Donald (Harvard medical School).

Antibodies. The following antibodies were used: mouse monoclonal anti-HA peptide (HA-7), (code: #H3663 Sigma-Aldrich) was used at 1/2,000 dilution for both western blotting (WB) and immunofluorescence (IF) while a 1/200 dilution was used for immunoprecipitation (IP); mouse monoclonal anti- α-tubulin (10D8, code: sc-53646, Santa Cruz biotechnology, Santa Cruz, CA, USA) (WB: 1/200); rabbit polyclonal anti-β-catenin (H-102, sc-7199, Santa Cruz biotechnology) (WB: 1/1,000); rabbit polyclonal anti- HA peptide, (code: H6908Sigma-Aldrich) (WB: 1/2,000; IF : 1/500, IP: 1/500); rabbit anti- AKT (#9272, Cell Signaling Technology, USA) (WB: 1/1000); rabbit anti- Phospho-AKT (Thr308) (D25E6, #13038, Cell Signaling Technology, USA) (WB: 1/2000); rabbit anti- Phospho- AKT (Ser473) (D9E, #4060, Cell Signaling Technology, USA) (WB: 1/2000); goat anti-rabbit IgG (H&L), DyLight 594 conjugate, (code: GtxRb-003-D594NHSX, Immuno- Reagents, Inc. Raleigh, NC, USA) (IF:1/500); Donkey anti- Rabbit IgG (H&L), HRP conjugate, (code: DkxRb-003- DHRPX, ImmunoReagent, Inc.) (WB: 1/4,000), goat anti- Mouse IgG (H&L), DyLight 594 conjugate, (code: GtxMu- 003-D594NHSX, ImmunoReagent, Inc.) (IF: 1/500 dilution for); goat anti-Mouse IgG (H&L), HRP conjugate, (code : GtxMu-003-DHRPX ImmunoReagent, Inc.) (WB: 1/4,000).

DNA-All DNA constructs were synthesized at Gene Script (USA). The cDNA coding for N terminally HA-tagged FZD4wt (HA-FZD4-wt), and for the mutants HA-FZD4-T425A and HA-FZD4-T425D were all cloned in the expression vector pCDNA3.1 (Invitrogen). For the reporter construct WRE-GFP-wt: 8 repeats of the optimized TCF/LEF binding sequence [5'-AGATCAAAGGGG-3'] (interspaced by the triplet 5'-GTA-3') were positioned upstream to a minimal TATA box promoter [5'-tagagggtatataatggaagctcgaattccag-3'] and a KOZAC region [5'-

cttggcattccggtactgttggtaaaaagcttggcattccggtactgttggtaaagccacc-3']. The sequence was cloned in the vector pCDNA 3.1 (+) GFP between the restriction sites for NruI and HindIII. This replacement substitutes the CMV promoter of the original vector with the TCF/LEF responsive sequences. The correctness of the sequences was verified by DNA sequencing. The control reporter construct WREGFP-mut was obtained by mutagenesis of WRE-GFP-wt and presents the 8 repeats of the TCF/LEF binding sequence mutated to [5'-AGGCCAAAGGGG-3']. A reporter construct (cmv-GFP) presenting GFP under the control of the cmv promoter was used as control. The plasmid necessary for the expression of the human AKT mutant AKT-K179M was kindly provided by Prof. G. Condorelli (Naples, Italy).

Immunofluorescence. Cells were grown up on glass coverslips and then were fixed in 3,7% formaldehyde/PBS pH 7.4 freshly made for 30 minutes at room temperature (RT). Formaldehyde was quenched by incubating the coverslips for 30 minutes in 0.1 M glycine/PBS. Cells were permealized in 0.1% Triton/PBS pH 7.4 for 10 minutes at RT. Glass coverslips were then incubated with primary and secondary antibodies diluted in PBS for 1 h and 30 min at RT, respectively. After fixing coverslips with glycerol/PBS 1:1 on slides, cells were analyzed with a Leica TCS-SMD-SP5 confocal microscope.

Compound synthesis. FzM1 FzM1alk and FzM1.8 were synthesized following a procedure of urea derivatives formation by reacting amines and isocyanates. HNMR, CNMR, FTIR spectroscopy and HRMS were used to confirm of the compounds. The synthesis of the FzM1 analogues was realized in cooperation with Professor Romano Silvestri's research group at University "La Sapienza" in Rome.

Immunoisolation from cultured cells and trypsin digestion. Cells were lysed in B buffer (Hepes K-OH 50mM, 150 mM NaCl, 1% Triton X-100, supplemented with protease inhibitors). Lysates were centrifuged at 14,000 r.p.m. to remove cell debris and unbroken cells. Clarified lysates were incubated overnight with antibody at 4 °C, followed by treatment with Protein A–coupled Sepharose (45 min, 4 °C). Samples were extensively washed in B buffer and then resuspended in 40 μl of buffer containing trypsin and incubated overnight at RT. 20 μl of the tryptic digestion were processed for LC/MS.

HDX. Immunoisolated Fz4 (still bound to the Sepharose beads) was incubated in B buffer at 4 °C in the presence of or without ligand (10 μM concentration). After 30 min, beads were resuspended in 40 μl of D20 for 1 min in the presence of or without

ligand. Formic acid was then added (5% final concentration), and the samples were boiled four times at 95 °C for 15 min. DTT (50 mM) was added to reduce disulfide bonds. 20 μl of the tryptic digestion were processed for LC/MS.

HPLC/MS. All samples were analyzed by analytical HPLC/MS (Agilent 1200 series HPLC system, Agilent 1260 UV-vis detector Infinity and Agilent Quadrupole 6110 LC/MS) equipped with a C18-bounded analytical reverse-phase HPLC column (Vydac 218TP104, 4.6 × 250 mm) using a gradient elution (10–90% acetonitrile in water (0.1% TFA) over 20 min; flow rate = 1.0 ml/min. For HDX experiments, the gradient buffer was kept at a pH lower than 2.5 (to reduce back exchange) by adding formic acid.

LC/MS spectra analysis. LC/MS spectra were analyzed with MetAlign51 with the following settings: mass resolution/BIN (nominal mass mode, 0.65), peak slope factor (5× noise), peak threshold factor (5× noise), peak threshold abs. value (150), average peak width (three scans), autoscaling on total signal, tuning alignment (prealign processing iterative, mass peak selection set on min. factor (5× noise). Amplitudes of masses coming from treated and untreated samples were subtracted to identify mass exclusively present in each of the samples. Masses were assigned with Mascot (MatrixScience)52. Samples containing only trypsin or protein Sepharose or antibody were run as controls.

Statistical analysis. EC_{50} was calculated by fitting direct total binding data by nonlinear regression analysis of dose-response curve using Prism software (GraphPad). All data were analyzed using the two-tailed Student's *t*-test. P < 0,05 were considered statistically significant.

FZD4 receptor model. FZD4-wt three-dimensional model was built using the CCP4 suite. Crystal structure of the cysteine-rich domain of FZD8 (Protein Data Bank (PDB) code 4F0A) and the TMD of Smoothened receptor (PDB codes 4NAW and 4JKV) were used as templates for ECD and TMD and ILC of FZD4, respectively.

TCF/LEF activity measurement using the WRE-GFP constructs. Hek293 cells were seeded (5 x 103 per well) in 96-well black Optyplates (Perkin Elmer). After 24 hours, cells were co-transfected with both WRE-GFP-wt (or when indicated WRE-GFP-mut) and HA-FZD4-wt. Transfection mixtures were prepared as follow: for each well 0,25μg of PEI (pH 7.0) was supplemented with 0,08 μg of DNA(both diluted in 4 μl of DMEM). The mixture was incubated at room temperature for 30 min to be then

diluted in culture medium and added to the cells. 24 hours after transfection, the medium was replaced and cells incubated with FzM1, FzM1.8 at the indicated concentrations and for the indicated times. When reported, cells were supplemented with either CKI 7 dihydrochloride (5µM), NFAT inhibitor (10µM), Bisindolylmaleimide II (7,5µM), Ly294002 (10µM), Suramin Hexasodium salt (10µM), Gallein (10µM), I-CBP 112 (7,5µM) or Neomycin (500µM). At the end of the incubations, cells were fixed in 3.7% formaldehyde in PBS pH 7.4 for 30 min. Formaldehyde was quenched by incubating the cells for 30 min in 0.1 M Glycine in PBS 1X. The activity of the compounds was evaluated by measuring GFP expression.

SUPPLEMENTARY FIGURES

Supplementary Tab S1: ELEMENTS OF THE LIBRARY

A) Known Folding Chaperone and Correctors

Radicicol

17AAG

Geldanamycin

Glycerol

DMSO

Methylene Blue

Chloroquine

Acridine Orange
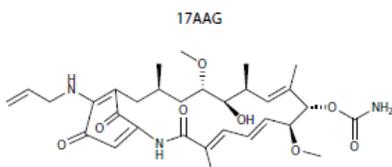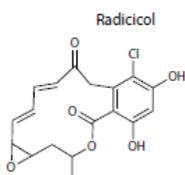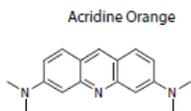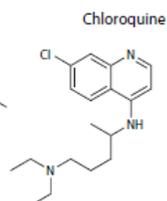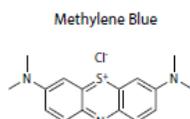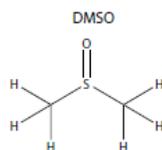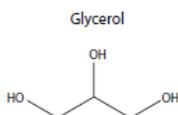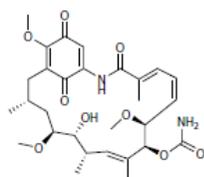
B) Other organic molecules of the library

FzM1

REFERENCES

(1) Shehu, A. (2010) Conformational Search for the Protein Native State, in *Introduction to Protein Structure Prediction: Methods and Algorithms*, pp 431–452.

(2) Gruebele, M. (2002) Protein folding: the free energy surface. *Curr Opin Struct Biol 12*, 161–168.

(3) Wong, K. B., Yu, T. H., and Chan, C. H. (2012) Energetics of protein folding, in *Comprehensive Biophysics*, pp 19–33.

(4) Ohgushi, M., and Wada, A. (1983) "Molten-globule state": a compact form of globular proteins with mobile side-chains. *FEBS Lett 164*, 21–24.

(5) Privalov, P. L. (1996) Intermediate states in protein folding. *J Mol Biol*.

(6) Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding "speed limit." *Curr Opin Struct Biol*.

(7) Skach, W. R. (2009) Cellular mechanisms of membrane protein folding. *Nat Struct Mol Biol 16*, 606–612.

(8) Onuchic, J. N., Nymeyer, H., García, A. E., Chahine, J., and Socci, N. D. (2000) The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. *Adv Protein Chem*.

(9) Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem 48*, 545–600.

(10) Miller, M. A., and Wales, D. J. (1999) Energy landscape of a model protein. *J Chem Phys 111*, 6610–6616.

(11) Dill, K. A., and Chan, H. S. (1997) From levinthal to pathways to funnels. *Nat Struct Biol 4*, 10–19.

(12) Kapon, R., Nevo, R., and Reich, Z. (2008) Protein energy landscape roughness. *Biochem Soc Trans 36*, 1404–1408.

(13) Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science (80-)*.

(14) Batey, S., Nickson, A. A., and Clarke, J. (2008) Studying the folding of multidomain proteins. *HFSP J*.

(15) Motojima, F. (2015) How do chaperonins fold protein? *Biophysics (Oxf) 11*, 93–102.

(16) Yang, A. S., and Honig, B. (1995) Free Energy Determinants of Secondary Structure Formation: I. A-Helices. *J Mol Biol 252*, 351–365.

(17) Yang, A. S., Hitz, B., and Honig, B. (1996) Free energy determinants of secondary structure formation: III. β-turns and their role in protein folding. *J Mol Biol 259*, 873–882.

(18) Rhee, Y. M., Sorin, E. J., Jayachandran, G., Lindahl, E., and Pande, V. S. (2004) Simulations of the role of water in the protein-folding mechanism. *Proc Natl Acad Sci 101*, 6456–6461.

(19) Tidor, B., and Karplus, M. (1993) The contribution of cross- links to protein stability: A normal mode analysis of the configurational entropy of the native state. *Proteins Struct Funct Bioinforma 15*, 71–79.

(20) Shandiz, A. T., Capraro, B. R., and Sosnick, T. R. (2007) Intramolecular cross-linking evaluated as a structural probe of the protein folding transition state. *Biochemistry 46*, 13711–13719.

(21) Anfinsen. (1973) Folding of Protein Chains. *Science (80-) 181*, 223–230.

(22) Bottomley, M. J., Batten, M. R., Lumb, R. A., and Bulleid, N. J. (2001) Quality control in the endoplasmic reticulum. *Curr Biol 11*, 1114–1118.

(23) Bu, G., Geuze, H. J., Strous, G. J., and Schwartz, A. L. (1995) 39 kDa receptor-associated protein is an ER resident protein and molecular chaperone for LDL receptor-related protein. *EMBO J 14*, 2269–80.

(24) Jessop, C. E., Chakravarthi, S., Watkins, R. H., and Bulleid, N. J. (2004) Oxidative protein folding in the mammalian endoplasmic reticulum. *Biochem Soc Trans 32*, 655–658.

(25) Neye, H. (2011) Peptidyl-prolyl isomerases, in *xPharm: The Comprehensive Pharmacology Reference*, pp 1–1.

(26) Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002) Proline cis-trans isomerization and protein folding. *Biochemistry*.

(27) Braakman, I., and Hebert, D. N. (2013) Protein folding in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol 5*, a013201.

(28) Mccaffrey, K., and Braakman, I. (2016) Protein quality control at the endoplasmic reticulum. *Essays Biochem 60*, 227–235.

(29) Nishikawa, S. I., Brodsky, J. L., and Nakatsukasa, K. (2005) Roles of molecular chaperones in endoplasmic reticulum (ER) quality control and ER-associated degradation (ERAD). *J Biochem*.

(30) Wilkinson, B., and Gilbert, H. F. (2004) Protein disulfide isomerase. *Biochim Biophys Acta 1699*, 35–44.

(31) Olzmann, J. A., Kopito, R. R., and Christianson, J. C. (2015) The Mammalian Endoplasmic Reticulum-Associated Degradation System 1–16.

(32) Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884–890.

(33) Winklhofer, K. F., Tatzelt, J., and Haass, C. (2008) The two faces of protein misfolding: Gain- and loss-of-function in neurodegenerative diseases. *EMBO J*.

(34) The Cystic Fibrosis Genotype-Phenotype Consortium. (1993) Correlation between genotype and phenotype in patients with cystic fibrosis. *N Engl J Med 329*, 1308–1313.

(35) Bobadilla, J. L., Macek, M., Fine, J. P., and Farrell, P. M. (2002) Cystic fibrosis: A worldwide analysis of CFTR mutations - Correlation with incidence data and application to screening. *Hum Mutat*.

(36) Skach, W. R. (2000) Defects in processing and trafficking of the cystic fibrosis transmembrane conductance regulator, in *Kidney International*, pp 825–831.

(37) Bichet, D. G. (2008) Vasopressin Receptor Mutations in Nephrogenic Diabetes Insipidus. *Semin Nephrol 28*, 245–251.

(38) Westergard, L., Christensen, H. M., and Harris, D. A. (2007) The cellular prion protein (PrP(C)): its physiological function and role in disease. *Biochim Biophys Acta 1772*, 629–44.

(39) Cordeiro, Y., Kraineva, J., Gomes, M. P. B., Lopes, M. H., Martins, V. R., Lima, L. M. T. R., Foguel, D., Winter, R., and Silva, J. L. (2005) The amino-terminal PrP domain is crucial to modulate prion misfolding and aggregation. *Biophys J 89*, 2667–2676.

(40) Scott, M. R., Groth, D., Tatzelt, J., Torchia, M., Tremblay, P., DeArmond, S. J., and Prusiner, S. B. (1997) Propagation of prion strains through specific conformers of the prion protein. *J Virol 71*, 9032–44.

(41) Mastrianni, J. A. (2010) The genetics of prion diseases. *Genet Med*.

(42) Ji, H.-F., and Zhang, H.-Y. (2010) Beta-Sheet Constitution of Prion Proteins. *Trends Biochem Sci 35*, 129–34.

(43) Eisenberg, D., and Jucker, M. (2012) The amyloid state of proteins in human diseases. *Cell*.

(44) Berg, J., Tymoczko, J., and Stryer, L. (2002) The Immunoglobulin Fold Consists of a Beta-Sandwich Framework with Hypervariable Loops, in *Biochemistry*.

(45) Ding, F., Borreguero, J. M., Buldyrey, S. V, Stanley, H. E., and Dokholyan, N. V. (2003) Mechanism for the alpha-helix to beta-hairpin transition. *Proteins 53*, 220–228.

(46) Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein-misfolding diseases. *Group 426*, 905–909.

(47) Leidenheimer, N. J., and Ryder, K. G. (2014) Pharmacological chaperoning: A primer on mechanism and pharmacology. *Pharmacol Res 83*, 10–19.

(48) Muntau, A. C., Leandro, J., Staudigl, M., Mayer, F., and Gersting, S. W. (2014) Innovative strategies to treat protein misfolding in inborn errors of metabolism: Pharmacological chaperones and proteostasis regulators. *J Inherit Metab Dis 37*, 505–523.

(49) Morello, J. P., Salahpour, A., Laperrière, A., Bernier, V., Arthus, M. F., Lonergan, M., Petäjä-Repo, U., Angers, S., Morin, D., Bichet, D. G., and Bouvier, M. (2000) Pharmacological chaperones rescue cell-surface expression and function of misfolded V2 vasopressin receptor mutants. *J Clin Invest 105*, 887–895.

(50) Manglik, A., and Kobilka, B. (2014) The role of protein dynamics in GPCR function: Insights from the β2AR and rhodopsin. *Curr Opin Cell Biol 27*, 136–143.

(51) Nygaard, R., Zou, Y., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. J., Bokoch, M. P., Thian, F. S., Kobilka, T. S., Shaw, D. E., Mueller, L., Prosser, R. S., and Kobilka, B. K. (2013) The Dynamic Process of β2-Adrenergic Receptor Activation. *Cell 152*, 532–542.

(52) Rochdi, M. D., Vargas, G. a, Carpentier, E., Oligny-Longpré, G., Chen, S., Kovoor, A., Gitelman, S. E., Rosenthal, S. M., von Zastrow, M., and Bouvier, M. (2010) Functional characterization of vasopressin type 2 receptor substitutions (R137H/C/L) leading to nephrogenic diabetes insipidus and nephrogenic syndrome of inappropriate antidiuresis: implications for treatments. *Mol Pharmacol 77*, 836–845.

(53) Bernier, V., Lagacé, M., Lonergan, M., Arthus, M.-F., Bichet, D. G., and Bouvier, M. (2004) Functional rescue of the constitutively internalized V2 vasopressin receptor mutant R137H by the pharmacological chaperone action of SR49059. *Mol Endocrinol 18*, 2074–84.

(54) Bernier, V. (2005) Pharmacologic Chaperones as a Potential Treatment for X-Linked Nephrogenic Diabetes Insipidus. *J Am Soc Nephrol 17*, 232–243.

(55) Wang, X. H., Wang, H. M., Zhao, B. L., Yu, P., and Fan, Z. C. (2014) Rescue of defective MC4R cell-surface expression and signaling by a novel pharmacoperone Ipsen 17. *J Mol Endocrinol 53*, 17–29.

(56) Farooqi, I. S., and O'Rahilly, S. (2006) Genetics of obesity in humans. *Endocr Rev*.

(57) Janovick, J. A., Maya-Núñez, G., Ulloa-Aguirre, A., Huhtaniemi, I. T., Dias, J. A., Verbost, P., and Conn, P. M. (2009) Increased plasma membrane expression of human follicle-stimulating hormone receptor by a small molecule thienopyr(im)idine. *Mol Cell Endocrinol 298*.

(58) Layman, L. C., Cohen, D. P., Jin, M., Xie, J., Li, Z., Reindollar, R. H., Bolbolan, S., Bick, D. P., Sherins, R. R., Duck, L. W., Musgrove, L. C., Sellers, J. C., and Neill, J. D. (1998) Mutations in gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism. *Nat Genet*.

(59) Janovick, J. A., Stewart, M. D., Jacob, D., Martin, L. D., Deng, J. M., Stewart, C. A., Wang, Y., Cornea, A., Chavali, L., Lopez, S., Mitalipov, S., Kang, E., Lee, H.-S., Manna, P. R., Stocco, D. M., Behringer, R. R., and Conn, P. M. (2013) Restoration of testis function in hypogonadotropic hypogonadal mice harboring a misfolded GnRHR mutant by pharmacoperone drug therapy. *Proc Natl Acad Sci 110*, 21030–21035.

(60) Noorwez, S. M., Kuksa, V., Imanishi, Y., Zhu, L., Filipek, S., Palczewski, K., and Kaushal, S. (2003) Pharmacological chaperone-mediated in vivo folding and stabilization of the P23H-opsin mutant associated with autosomal dominant retinitis pigmentosa. *J Biol Chem 278*, 14442–14450.

(61) Kobayashi, H., Ogawa, K., Yao, R., Lichtarge, O., and Bouvier, M. (2009) Functional rescue of ??1-adrenoceptor dimerization and trafficking by pharmacological chaperones. *Traffic 10*, 1019–1033.

(62) Newton, C. L., Whay, A. M., McArdle, C. A., Zhang, M., van Koppen, C. J., van de Lagemaat, R., Segaloff, D. L., and Millar, R. P. (2011) Rescue of expression and signaling of human luteinizing hormone G protein-coupled receptor mutants with an allosterically binding small-molecule agonist. *Proc Natl Acad Sci U S A 108*, 7172–6.

(63) Generoso, S. F., Giustiniano, M., La Regina, G., Bottone, S., Passacantilli, S., Di Maro, S., Cassese, H., Bruno, A., Mallardo, M., Dentice, M., Silvestri, R., Marinelli, L., Sarnataro, D., Bonatti, S., Novellino, E., and Stornaiuolo, M. (2015) Pharmacological folding chaperones act as allosteric ligands of Frizzled4. *Nat Chem Biol 11*, 280–286.

(64) Robitaille, J., MacDonald, M. L. E., Kaykas, A., Sheldahl, L. C., Zeisler, J., Dubé, M. P., Zhang, L. H., Singaraja, R. R., Guernsey, D. L., Zheng, B., Siebert, L. F., Hoskin-Mott, A., Trese, M. T., Pimstone, S. N., Shastry, B. S., Moon, R. T., Hayden, M. R., Paul Goldberg, Y., and Samuels, M. E. (2002) Mutant frizzled-4 disrupts retinal angiogenesis in familial exudative vitreoretinopathy. *Nat Genet 32*, 326–330.

(65) Lemma, V., D'Agostino, M., Caporaso, M. G., Mallardo, M., Oliviero, G., Stornaiuolo, M., and Bonatti, S. (2013) A disorder-to-order structural transition in the COOH-tail of Fz4 determines misfolding of the L501fsX533-Fz4 mutant. *Sci Rep 3*, 1–10.

(66) Riccio, G., Bottone, S., La Regina, G., Badolati, N., Passacantilli, S., Rossi, G. B., Accardo, A., Dentice, M., Silvestri, R., Novellino, E., and Stornaiuolo, M. (2018) A Negative Allosteric Modulator of WNT Receptor Frizzled 4 Switches into an Allosteric Agonist. *Biochemistry 57*, 839–851.

(67) D'Agostino, M., Lemma, V., Chesi, G., Stornaiuolo, M., Cannata Serio, M., D'Ambrosio, C., Scaloni, A., Polishchuk, R., and Bonatti, S. (2013) The cytosolic chaperone -crystallin B rescues folding and compartmentalization of misfolded multispan transmembrane proteins. *J Cell Sci 126*, 4160–4172.

(68) Akiyama, T. (2000) Wnt/β-catenin signaling. *Cytokine Growth Factor Rev*.

(69) Ke, J., Harikumar, K. G., Erice, C., Chen, C., Gu, X., Wang, L., Parker, N., Cheng, Z., Xu, W., Williams, B. O., Melcher, K., Miller, L. J., and Eric Xu, H. (2013) Structure and function of Norrin in assembly and activation of a Frizzled 4-Lrp5/6 complex. *Genes Dev 27*, 2305–2319.

(70) Osakada, F., Jin, Z.-B., Hirami, Y., Ikeda, H., Danjyo, T., Watanabe, K., Sasai, Y., and Takahashi, M. (2009) In vitro differentiation of retinal cells from human pluripotent stem cells by small-molecule induction. *J Cell Sci 122*, 3169–3179.

(71) Cruciat, C. M. (2014) Casein kinase 1 and Wnt/β-catenin signaling. *Curr Opin Cell Biol*.

(72) Dijksterhuis, J. P., Petersen, J., and Schulte, G. (2014) WNT/Frizzled signalling: Receptor-ligand selectivity with focus on FZD-G protein signalling and its physiological relevance: IUPHAR Review 3. *Br J Pharmacol*.

(73) Von Maltzahn, J., Bentzinger, C. F., and Rudnicki, M. A. (2012) Wnt7a-Fzd7 signalling directly activates the Akt/mTOR anabolic growth pathway in skeletal muscle. *Nat Cell Biol 14*, 186–191.

(74) Kim, S. E., Lee, W. J., and Choi, K. Y. (2007) The PI3 kinase-Akt pathway mediates Wnt3a-induced proliferation. *Cell Signal 19*, 511–518.

(75) Li, J., Sutter, C., Parker, D. S., Blauwkamp, T., Fang, M., and Cadigan, K. M. (2007) CBP/p300 are bimodal regulators of Wnt signaling. *EMBO J 26*, 2284–2294.

(76) Thorpe, L. M., Yuzugullu, H., and Zhao, J. J. (2015) PI3K in cancer: Divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer*.

(77) Gallenkamp, D., Gelato, K. A., Haendler, B., and Weinmann, H. (2014) Bromodomains and their pharmacological inhibitors. *ChemMedChem*.

(78) MacDonald, B. T., and He, X. (2012) Frizzled and LRp5/6 receptors for wnt/β-catenin signaling. *Cold Spring Harb Perspect Biol 4*.

(79) Jin, X., Jeon, H. Y., Joo, K. M., Kim, J. K., Jin, J., Kim, S. H., Kang, B. G., Beck, S., Lee, S. J., Kim, J. K., Park, A. K., Park, W. Y., Choi, Y. J., Nam, D. H., and Kim, H. (2011) Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. *Cancer Res 71*, 3066–3075.

(80) Heathcote, J. (2011) Page 1 of 56 Hepatology. *Hepatology* 1–56.

Chapter 2

INTRODUCTION

Orphan GPCRs

The superfamily of GPCRs includes at least 800 seven-transmembrane receptors, all participating in the most diverse physiological and pathological functions[81]. Approximately 36% of marketed pharmaceuticals drugs target human GPCRs, the most successful targets of modern medicine[81].

While for a conspicuous number of GPCRs drugs have been identified, optimized and tested in clinical trials, for more than 90 GPCRs drug discovery is still at the stage of identifying their endogenous ligands and their biological functions[82]. These so-called "orphan" GPCRs (OrphGs) have been involved in many types of cancer and rank among the most interesting future pharmacological targets[81]. Identification of their endogenous ligands is an unquestionable missing milestone of science, necessary to start the characterization of these unexplored receptors. The worldwide scientific plan to identify OrphGs' endogenous ligands is referred to as "GPCRs deorphanization"[83]. Its anticipated outcome is to fully describe the structure, the biology and the physiology aspects of this interesting pool of pharmacological targets.

Orphan GPCRs and Glioblastoma

While science is still looking for OrphGs' endogenous function, their involvement in cancer is fully clear. Recently it has been shown that several OrphGs are specifically overexpressed in several forms of cancer, among which Glioblastomas[84]. Glioblastomas (World Health Organization (WHO) grade IV astrocytomas) are highly aggressive infiltrating brain tumors and represent more than 50% of all gliomas[85]. At the cellular level these tumors are constituted by a variety of cells in different states of differentiation and are enriched in cancer stem cells (CSCs)[86]. Topoisomerase and tyrosine kinase inhibitors, alkylating and intercalating agents are examples of compounds nowadays used to treat glioblastoma[87,88]. However, classical radiotherapies and chemotherapies act preferentially on highly proliferating cells and thus spare CSCs, which are able to enter in a quasi-quiescent state and escape treatments[86]. The severe morbidity and mortality of glioblastoma heightened the search for new drugs acting on new targets different from those explored so far. The search for new treatments has uncovered several G protein-coupled receptors (GPCRs) as candidates and generated a growing interest in this class of proteins as alternative therapeutic targets for the treatment of Glioblastoma[84,89].

CELSR2: an orphan GPCR overexpressed in Glioblastoma

CELSR2 [Cadherin-EGF-LAG-Seven Pass-G-type receptor 2 (uniprot:Q9HCU4)] is an orphan GPCR overexpressed in Glioblastoma. It belongs to class B GPCRs, also known as adhesion GPCRs (aGPCRs)[84]. This indicates the presence in their long extracellular domain (ECD) of sequences of repeated domains, each acquiring folds proper of adhesion molecules. Human CELSR2 has an extracellular domain of 2350 aa and contains 9 Cadherin repeats, 7 EGF-like domains, 2 Laminin G-like domains, one Laminin-EGF-like domain and one hormone receptor motif (HormR). The ECD is followed by the canonical GPCR transmembrane assemblement with 7 transmembrane (TM) helices linked by 3 intracellular and 3 extracellular loops. ECD and the first transmembrane are joint together by a GPCR self-proteolysis site (GPS)[90].

The CELSR2 cadherin domain presents 110 amino acid residues and contains the sequences responsible for extracellular calcium binding[91]. In most of the proteins in which are present, cadherin domains are involved in the formation of trans/cis dimers, with these interactions requiring calcium coordination. However, that this can be applied also to CELSR2 cadherin domains remains to be proved. *Shima et al.*[92] obtained important mechanistic insights into the molecular function of CELSR2 ectodomain. They expressed siRNA targeting CELSR2 in brain neurons and showed that CELSR2 loss of function caused dendritic simplification. Then, they expressed several deletion forms of CELSR2 to understand which domain of CELSR2 would have been able to rescue the phenotype and proved that the presence of the cadherin domain was essential for rescuing it [92]. Further experimental evidences showed that the isolated CELSR2 cadherin domain (CELSR2-CR) when applied to hippocampal primary neurons enhanced neurite growth. Interestingly, the neuronal phenotypic effects of CELSR2-CR were lost in CELSR2 knockdown cells, proving that homophilic interactions CELSR2-CR/CELSR2 were necessary for activation of endogenous CELSR2 proteins[93]. CELSR2-CR acts also as a ligand for CELSR2 proteins, evoking an increase of the intracellular calcium[93]. Further investigations on CELSR2 calcium-regulated cis/trans-dimerization are however necessary to explain in detail this potency of CELSR2 cadherin domains to activate receptor signaling.

Seven EGF-like domains (EGF) separated by two laminin G (LAG) domains follow the CELSR2 cadherin repeats, both potentially able to bind calcium. Each EGF-CA domain contains about 40 amino acids, and the LAG domain acquires a jelly roll structure[94]. Downstream to the EGF-CA and LAG domains, CELSR2 contains an

EGF-laminin domain of approximatively 60 amino acids with 8 conserved cysteines. At the C-terminus of these repeats there is an hormone receptor motif (HormR) which contains conserved tryptophans and cysteines[95]. The HormR domain is a common feature of adhesion GPCRs. Even if it is often regarded as a potential binding domain for hormones [96], the essential N-terminal α-helix responsible for hormone interaction is missing in all the adhesion GPCRs, suggesting a structural more than a functional activity for this domain.

At the membrane-proximal region of CELSR2, a GPCR Autoproteolysis-Inducing (GAIN) domain (320 residues) contains the GPCR proteolysis site (GPS) of approximatively 50 amino acids. The GAIN is the most highly conserved feature of aGPCRs[97]. The crystal structure of the aGPCR Latrophilin-1 GAIN domain revealed that it is organized in two subdomains: a α-helical and a β-sandwich subdomain [97]. The α-helical subdomain interacts with the HormR domain in a manner that would block the interaction of homologous HormR domains with hormones. The GPS motif comprises the last five β-strands of the GAIN domain, and the cleavage occurs between the last two β-strands[95,97] (**Fig. 28**). In vitro, proteolytic cleavage occurs through an intramolecular self-catalytic reaction between a conserved aliphatic residue (usually leucine) that acts as a general base and a threonine, serine or cysteine that acts as a nucleophile (L↓T/S/C)[98]. Interestingly, only human CELSR2 has the general base H2355 and the nucleophile T2357 suggesting that cleavage can be dispensable for the protein function. The cleavage occurs through a series of nucleophilic attacks and subsequent generation and hydrolysis of an ester intermediate. This process takes place in the Endoplasmic Reticulum (ER) and gives rise two fragments, an extracellular and a 7-TM subunits, that associate non covalently as a heterodimer. Although for several aGPCRs the proteolysis is an essential step for protein maturation and activity [99], recent studies have suggested that for some aGPCRs proteolysis is not necessary, and an intact GPS has structural meaning for the receptor [100]. The functional importance of the GPS is also underlined by the association between two GPS point mutations of an another class B GPCR, GPR56, with the bilateral front parietal polymicrogyria (BFPP), a congenital brain malformation. The GPS point mutations C346S and W349S impair cleavage and block protein trafficking to plasma membrane[101]. However, whether hsCELSR2 is regulated by the self-proteolysis process in physiologic contexts remains to be clarified.

A seven transmembrane helical region (7-TM) is located after the GAIN domain. Two conserved cysteines from the first and the second extracellular loop (C2439 and C2511) can form a disulfide bond that, stabilizing the extracellular structure might

ensure efficient ligand binding or conformational changes [102], as it has already been shown for the gonadotropin-releasing hormone receptor, a class A GPCR [103]. However, the TM bundle of CESLR2 is still orphan, undrugged and unexplored (IUPHAR/BPS).

The intracellular part of CELSR2, that includes intracellular loop 1-3, serves as a docking point for downstream effectors in the cytoplasm. The intracellular loop 3 contains about 20 amino acids and provide structural flexibility. Interestingly, CELSR2 has an arginine (Arg2413) in the first intracellular loop that is not shared by other family members. The latter, indeed, at the same position have a histidine (His2573). Site directed mutagenesis experiments showed that the Arg2413 in CELSR2 and the His2573 in CELSR3 are responsible for the opposite roles of these two proteins in neuron growth regulation[104], with CELSR2 enhancing and CELSR3 inhibiting neurite growth.

Finally, after the helix 7, CELSR2 has a C-tail of 304 amino acids enriched with Ser, Thr, Pro and acidic residues Glu and Asp. Moreover, CELSR2 has several predicted PEST regions (signal peptide for protein degradation) that may regulate receptor turnover in the cell[105].

The domain organization is strikingly conserved among CELSR2-like proteins in different species. The vertebrate genomes contain three homologs: CELSR1, CELSR2 and CELSR3 whereas *Drosophila melanogaster* and *Caenorhabditis elegans* genomes contain one ortholog of CELSR2, the proteins flamingo/starry night (*fmi/stan*) and Flamingo (*fmi-1*), respectively. The extracellular domains show a high degree of similarity between species, whereas the intracellular domains vary in length and are generally not conserved[106] (**Fig. 29**).
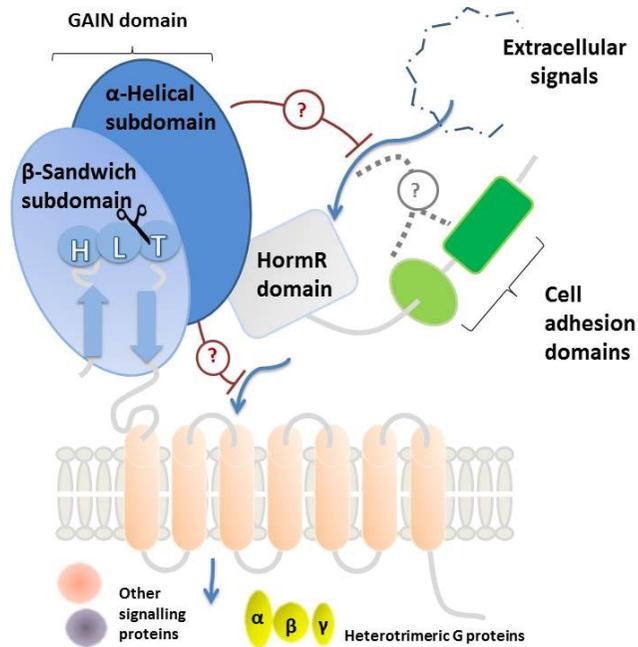
Figure 28. The conserved structural core of the extracellular regions of adhesion GPCRs, as revealed by the X-ray structure analysis of latrophilin. The GAIN domain, that is highly conserved among the 33 mammalian adhesion GPCRs, folds in an α-helical and a β-sandwich subdomain. The auto proteolytic processing occurs in a tight turn between the last two strands of the domain. On the contrary, the GAIN α-helical subdomain of latrophilin interacts with a HormR domain. This interdomain contact, and/or a direct interaction with the transmembrane helical bundle, may allow the GAIN domain to auto-inhibit transmembrane signaling to heterotrimeric G proteins or other factors inside the cell. After cleavage, the cleaved fragment appears firmly anchored within the GAIN domain by extensive backbone hydrogen bonds and hydrophobic interactions. Image adapted from The EMBO Journal (2012) 31, 1334–1335.
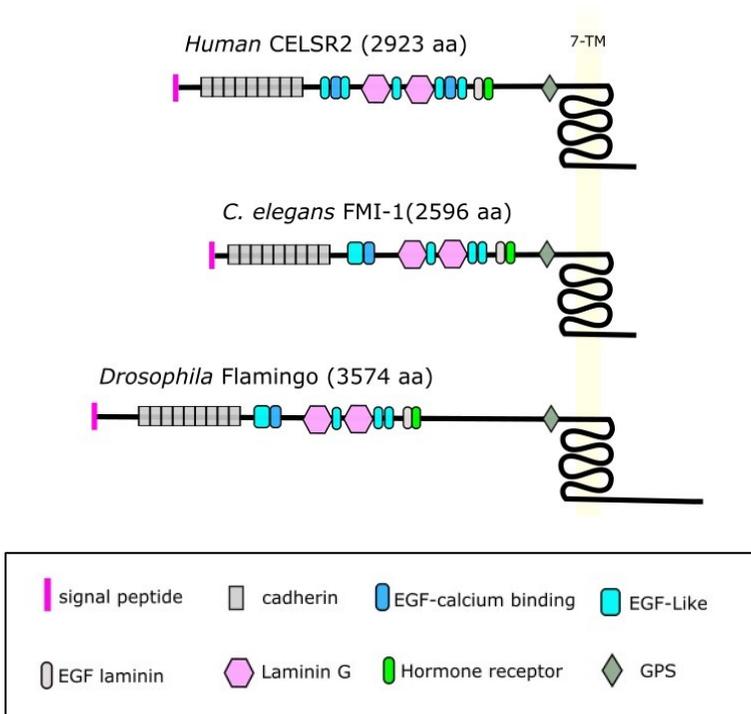
Figure 29. Protein domain organization of CELSR2. CELSR2 ectodomain contains 9 Cadherin repeats, 7 EGF-like domains, 2 Laminin G-like domains, one Laminin-EGF-like domain and one hormone receptor motif (HormR). The ECD is followed by the canonical GPCR transmembrane assemblement with 7 transmembrane (TM) helices. ECD and the first transmembrane are joint together by a GPCR self-proteolysis site (GPS). The protein domain organization of *Drosophila* Flamingo and *C.elegans* FM-1 are shown for comparison. Image adapted from Development 137, 3663-3673 (2010).

CELSR2 cellular functions

CELSR2 is widely expressed in nervous system and its expression is spatially and temporally regulated[107,108]. It is mainly localized in dendrites and axons of different types of embryonic and post-natal neurons[109]. It is also expressed in non-neuronal epithelia, such as kidney, lungs, skin, endothelial cells, gastrointestinal and reproductive systems. Phenotypic analysis of CELSR2 mutants showed its involvement in both epithelial and neuronal morphogenesis.

There are several evidences of the role of CELSR2 in cilia development and organization. In mouse nervous tissue cells, CELSR2 controls function of cilia that, normally, beat in a coordinated manner to facilitate circulation of the cerebrospinal fluid. Indeed, CELSR2 mutants affect this process causing defective beats at individual cell level[110,111] (**Fig. 30 a**). Mice with homozygous null mutations in CELSR2 gene develop hydrocephalus with accumulation of cerebrospinal fluid within the brain, due to decreased numbers of cilia that are also short and abnormally oriented[110]. CELSR2 regulates ciliogenesis via Planar cell polarity[112], the signalling pathway involved in the orientation of cells and cellular structures along an axis within the plane of an epithelial surface. Indeed, in CELSR2 mutant cells, the membrane localization of VANGL2 and FZD3, two key regulators of planar cell polarity (PCP), is disrupted[112]. The role of CELSR2 in human ciliopathies was recently emphasized with the demonstration that mutated CELSR2 is implicated in Joubert syndrome (JS)[113]. This genetic disease has a prevalence of 1-9/100,000 and it is characterized by central nervous system malformations with neurological features (ataxia, cognitive impairment, abnormal eye and tongue movements) and multisystem organ involvement (fibrocystic kidney and liver disease, retinal dystrophy and polydattilia)[111]. Consequently, defects of CELSR2 function might be associated also to other pathological conditions characterized by cilia abnormalities such as cystic pathology in liver, kidney and pancreas.

CELSR2 is also involved in neuronal migration. In zebrafish embryos, loss of function of CELSR2 results in migration defects of the neurons controlling the muscles of facial expression also known as facial brachiomotor neurons (FBMN)[114]. Another evidence of CELSR2 involvement in neuronal migration is given by CELSR2 mutant mice, where the directionality of the neuron migration is compromised[115]. Furthermore, in cells from *Drosophila* embryos the C-terminal truncated form of *Flamingo* (*fmi,* homolog of CELSR2) affects cell cohesion that primarily leads to the

in vivo cell movement defects[116]. Less is known about CELSR2 molecular function in cell migration. However, the phenotype of CELSR2 mutants is similar to that of FZD3 mutants and other PCP related genes also implicated in FBMN migration[117].

Other structure-phenotype analysis showed that CELSR2 plays an important role also in dendrite development and axon guidance, from Drosophila to mammals.

In Drosophila, *fmi* mutants affect both the dendritic morphogenesis and the competition between dendrites of homologous neurons in embryonic peripheral nervous system (PNS)[118] (**Fig. 30 b**). Structure-function analysis showed that the expression of a *fmi* truncated form lacking the entire cadherin repeat sequence, rescues overgrowth of branches in *fmi*-mutants in the early stage of dendritic development. In *fmi*-mutants, another phenotype observed in the later stage, is the branches overlapping that is reproduced by the overexpression of the ECD truncated form of CELSR2 in the wild-type background[119]. These studies suggest that *fmi* could play two different roles during dentritic development: act as a receptor for an unknown ligand to limit branch elongation in the initial phase and establish homophilic interactions to elicit avoidance between dendritic terminals in the terminal phase.

CELSR2 regulation of dendritic maintenance and growth exists also in mammals. Normally, dendrites are highly dynamic and undergo rapid and continuous extensions and retractions[93]. In cultures of murine brain slices, loss of function of CELSR2 by siRNA causes severe simplification of dendritic arbores probably due to a more frequent dendritic retraction (**Fig. 30 c**). Indeed, the siRNA induced phenotype is rescued by co-expressing a siRNA resistant form of CELSR2, containing only the cadherin domain in the ECD, that is not able to promote elongation when expressed by itself [92]. Finally, CELSR2 cadherin domain is necessary for the rescue of the retraction phenotype and homophilic CELSR2-CELSR2 interactions might have a central role in maintaining dendritic arbores.

CELSR2 plays a pivotal role also in axon growth. *Fmi* gene mutations are lethal in Droshophila embryos and the mutant embryos showed local disconnection of longitudinal axon tracts in the central nervous system[106] (**Fig. 30 d**). CELSR2 might regulate axon development by PCP-independent pathways since mutants of PCP genes do not display the axon phenotype seen in *fmi* mutant neurons[106]. Axonal defects and correlated lethality can be rescued by *fmi* expression in the nervous system of the mutants indicating that the role of this receptor is essential for viability.

Finally, CELSR2 regulates cellular functions such as cell adhesion, migration, polarity and guidance, which are all highly relevant to tumour cell biology. Nevertheless many questions remain unanswered, such as how CELSR2 is activated, who are CELSR2 endogenous ligands and what signal they transduce and whether CELSR2 can be coupled to G proteins.
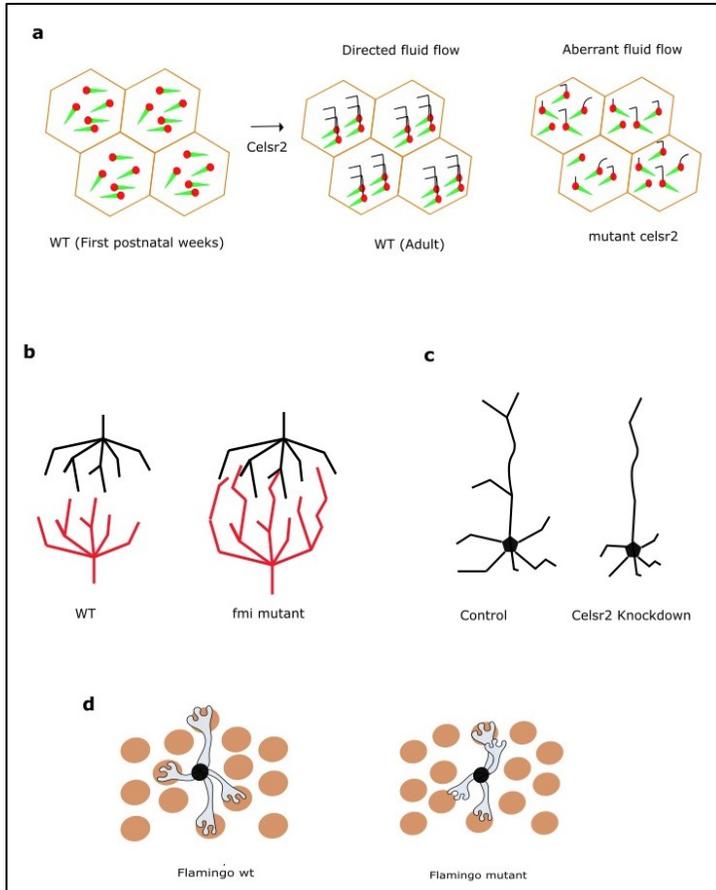
Figure 30. CELSR2/Flamingo roles in physiological processes. (a) During the first postnatal week, cilia basal body (green arrowheads) point in different directions. In the adult, in response to planar cell polarity signals, they rotate and adopt a more uniform orientation that enables cilia to beat in a coordinated manner and generate a directed fluid flow. In CELSR2 mutant cells, basal bodies exhibit divergent orientation and fail to generate a directed fluid flow. **(b)** In *Drosophila*, peripheral neuron dendrites avoid dendrites of homologous neurons in normal conditions. In *fmi* mutants, on the contrary, dendrites establish competitive interactions. **(c)** In mouse neuronal cultures, knocking down CELSR2 suppresses neurite growth. **(d)** In flamingo mutants, axons make directional errors and innervate inappropriate targets.

74

CELSR2 molecular partners

There are several lines of evidence, both in *Drosophila* and mammals, of the CELSR2 involvement in the planar cell polarity signalling, that is the polarity axis that organizes cells in the plane of the tissue. It seems that *fmi*/CELSR2 works together with PCP genes in development processes that require cell polarity, including hair cell orientation, convergence and extension and neuronal migration.

CELSR2 ortholog *fmi* has been found to functionally interact with key PCP proteins, including the 7-TM receptor Frizzled, tetraspanin Van Gogh, Dishevelled, Prickle and probably other less clearly identified members. Also in mammals, CELSR2 works together with Frizzled3 and Vangl-2[120]. In the context of PCP signalling pathway, CELSR2 may regulate the C-JUN N-terminal kinase (JNK) pathway, that is known to be activated by PCP in multiple organs[121]. Phosphorylation of c-Jun during pancreatic development relies on CELSR2 activity, indeed, CELSR2 KO mice embryos have a downregulation of phospho-Jun in pancreatic buds[122].

In contrast, *fmi/*CELSR2 may act by PCP-indipendent pathway in other kind of physiologic processes such as dendritogenesis and axon advance.

In mammals, CELSR2 role in dendritogenesis has been proved to be dependent on homophilic interactions of CELSR2 cadherin domains. In this way, CELSR2 may activate Phospholipase C to produce inositol-3-phosphate that is responsible of a calcium release from internal stores[90]. These evidences suggest that CELSR2 have a GPCR activity, however, it would be necessary to carry out biochemical analysis to identify relevant G proteins. Moreover, several lines of evidences suggest that CELSR2 is linked to calcium/calmodulin-dependent protein kinase II (CaMKII). Pharmacological inhibition experiments, for example, revealed that CaMKII mediate CELSR2 cadherin domain effects on neurite growth[93].

Also in Drosophila, *fmi* mediates homophilic interactions in neuro epithelial cells to regulate axon guidance. In Drosophila, several transmembrane molecules have been shown to regulate axon guidance and among these, the transmembrane receptor Golden Goal (*gogo*) have phenotypic similarities and genetically interacts with *fmi* [123]. *gogo* and *fmi* can mutually influence their localization and this suggests that they are molecular partners. Moreover, *fmi* seems to trigger intracellular signaling via *gogo*

since the synergistic effect of *gogo* and *fmi* overexpression in recovering phenotype of *fmi* mutant neurons depends on Gogo cytoplasmic domain[123].

Finally, to date, less is known about the signaling events triggered by CELSR2 activation or inactivation. One potential cause is the lack of efficient *in vitro* research tools, such as selective compounds that specifically activate or inactivate CELSR2.

Spatiotemporal expression profiles of CELSR proteins in zebrafish

Recently, zebrafish (*Danio rerio*) has proven to be a useful model to study adhesion GPCRs from a functional prospective, especially in development context. The utility of this model has been further underlined by a genomic study that identified the adhesion GPCRs encoded by the zebrafish genome. Moreover, in this study the expression profiles of each adhesion GPCRs identified in zebrafish was provided by using high-throughput quantitative real-time PCR in a collection of 12 developmental time-points and 10 adult tissues. The following tissues were selected to represent the major organ systems: brain, eye, heart, intestine, kidney, liver, skeletal muscle, skin, ovaries and testes. On the contrary, the developmental time-points that were chosen to represent major milestones are the following: 1 hour post-fertilization (hpf; cleavage period, ~4 cells), 3 hpf (blastula period; ~1000 cells), 5.3 hpf (early gastrulation), 10 hpf (late gastrulation, early segmentation), 14 hpf (segmentation), 24 hpf (most organ systems have formed), 3 days post-fertilization (dpf, larvae have hatched from chorions), 5 dpf (swimming), 7 dpf, 11 dpf, 14 dpf, and 21 dpf (juvenile stages defined by active hunting and rapid body growth) [124].

Interestingly, the study revealed that there are at least 59 adhesion GPCRs in zebrafish representing 24 of the 33 human aGPCRs. Moreover, zebrafish adhesion GPCRs have similar expression profiles as their homologs in mouse, rat and human. CELSR1 and CELSR2 are highly expressed during gastrulation and segmentation stages and then expressed at lower levels during later development and in adult tissues. In particular, they are significantly enriched between 10 and 15 hours post fertilization, in contrast, CELSR3 is lowly expressed until 3 dpf, and then highly enriched in the brain, eye, and skeletal muscle in adults.

Phylogenetic analysis suggest that zebrafish aGPCRs cluster closely with their mammalian homologs and are organized in the same nine groups as previously described for the human aGPCR repertoire. Also structure prediction studies show that most zebrafish aGPCRs share relevant protein domains found in the N-termini of their mammalian counterparts. Importantly, the high level of homology between zebrafish and human adhesion GPCRs found in this study is consistent with global characteristics of the zebrafish genome relative to the human genome. For example, 71.4% of all human genes have at least one ortholog in zebrafish, which is consistent with the data that 72.7% of human aGPCRs have at least one zebrafish ortholog. Thus, studying adhesion GPCRs by using zebrafish as model could have important

implications also for understanding some functional aspects of adhesion GPCRs in humans.

Inapplicability of traditional GPCR screening technologies to CELSR2 deorphanization

In the last decades, several GPCR assays have been designed to identify novel drug candidates. Historically, radio ligand binding assays have been used to screen GPCRs for lead compounds. This assay is also known as "radio-ligand displacement assay" since compounds are characterized for their ability to displace the binding of a radio labelled ligand to the target[125]. Although this assay has been widely used to identify new ligands of GPCRs, it is not useful for the deorphanization of CELSR2 as well as of the other orphan GPCRs. Indeed, the lack of information about orphan GPCR endogenous ligands and the absence of any traceable radioactive ligand binding to CELSR2 makes this assay inapplicable to the deorphanization of this receptor.

New binding assays, based on the fluorescence resonance energy transfer (FRET) technology, and "functional assays" have been developed to study both the compound activity and its effect as potential novel drug candidate. Originally, functional assays were G-protein dependent since they were based on the measurement of G-protein-mediated second messenger generation. For example, *Guanine nucleotide binding assay* measures an early event after GPCR activation that is the release of GDP from the G-protein and the binding of GTP. On the other hand, a functional assay that measures a downstream effect of GPCR activation is the *cAMP assay*. This assay, indeed, is sensitive to variations in cAMP intracellular levels. Taking advantage of the new insights into GPCR signal transduction, many G-protein-independent assays have been developed to provide additional information on the functional selectivity of the candidate compounds. Among these assays, *receptor internalization assay* allows to visualize GPCR internalization by labelling GPCR with fluorescent proteins (GFP or RFP). When internalized, the tagged-GPCR forms intracellular spots that are analysed by a detector[126]. This assay was proved to be useful for GPCR deorphanization. Indeed, with the internalization assay it has been showed that unsaturated long-chain free fatty acids activate the so far, orphan receptor GPR120[127].

However, most of the functional assays are not useful to identify ligands of orphan receptors like CELSR2. Indeed, they rely on monitoring changes at the second messenger level. Unfortunately, to date, the knowledge about orphan GPCRs role in the transmission of signal pathways is limited.

.

Deorphanization strategies

In the last decades, the rate of GPCR deorphanization decreased drastically because of the inapplicability of the traditional screening assays.

Some attempts to develop appropriate assays for GPCRs deorphanization led to alternative functional assays. For example, an alternative calcium assay approach takes advantage of the propriety of promiscuous Gα subunits to enforce $Gα_i$, $Gα_s$ coupled-receptor to signal via $PLC-IP_3-Ca^{2+}$. Overexpression of promiscuous G-proteins in reporter cells, containing the receptor and a calcium sensitive probe, allows to perform the assay also without preventive knowledge of G-protein coupling conditions or signalling pathways[128]. Recently, an elegant way of measuring GPCR activation by using β-arrestin translocation has been developed. Before starting screening, this system ensures, in a ligand independent way, that the receptor can recruit arrestin proteins after stimulation. This approach uses a constitutively active kinase (GRK2) to phosphorylate the receptor and initiate GPCR-β-arrestin interaction[129]. Another way to study signalling proprieties of orphan GPCRs with unknown G-protein coupling is to overexpress them in a cell line and then measure activation of a specific signalling pathway. The overexpression, indeed, cause a constitutive receptor activation that might be successful to identify some intracellular signalling characteristics[130]. This method has been used to identify GPR101 activities in the central nervous system via modulation of cAMP levels[130]. However, the successful application of these methods depends on sufficient orphan receptor expression. It is well known that the expression and membrane translocation of some GPCRs, in reporter cell lines such as Hek293 or COS-1, can fail because of a lack of a specific protein maturation machinery. Other limitations are due to a different pharmacology and/or biological role of certain receptors. Some of the orphan receptors may simply have ligand-independent functions. Several examples include $GABA_B$, GPR50 and the constitutively active receptors.

To avoid these difficulties, different in silico strategies are being attempted to identify candidate endogenous ligands. Bioinformatics provides tools to predict cognate ligands based on sequence homology. For example, the orphan receptor GPR105 (known as P2Y14) has been deorphanized in virtue of its high level of homology with other purinergic receptors. Indeed, the orphan receptor was screened against a library containing nucleotide di- and triphosphate conjugates to identify UDP-glucose as ligand. Also the deorphanization of EDG-3 and EDG-5 receptors was successfully

conduced on the basis of a sequence homology analysis that showed an high level of homology with EDG-1, the receptor binding sphingosine-1-phosphate (SP-1)[131]. Moreover, the deposition of high resolution X-ray and NMR structures represents an opportunity to give a boost to structure-based design (SBDD), an effective tool to design potential agonists or antagonists. SBDD in conjunction with computer modelling allowed the identification of allosteric modulators and agonists of the proton receptor GPR68[132]. Notwithstanding, the specificity that some ligands have revealed for GPCRs has mined a full confidence in such strategy. CD97 and EMR2 receptors, for example, share 97% sequence identity while, in contrast, their ligand CD55 binds with high affinity only to CD97 and not to EMR2. On the other hand, alkyl-imidazoles function as dual Histamine H3/H4 receptors, while these receptors share very little sequence identity. These examples suggest the "similarity" based approaches must be taken with caution[133,134].

Thus, the development of a universal (valid for all the remaining OrphG), improved (designed to overcome problems nowadays faced with traditional approaches) and sensitive (needing low amount of expressed/purified OrphG) platform for OrphG deorphanization is urgently needed.

RESULTS

Development of a specific PC-platform to use for the identification of CELSR2 modulators

In CELSR2 two conserved cysteines from the first and the second extracellular loop (C2439 and C2511) can form a disulfide bond that, stabilizing the extracellular structure might ensure efficient ligand binding or conformational changes[102], as has already been shown for the gonadotropin-releasing hormone receptor, a class A GPCR[103]. Thus, we supported the idea that CELSR2 TM-bundle can potentially bind ligands to modulate receptor activity.

We attempted to deorphanize the TM bundle of CELSR2 by using the novel PC-platform recently ideated in our research group. As already described, PC-Platform identifies new ligands in virtue of their ability to act as pharmacological chaperones (**Fig. 31**). Therefore, the target receptor must be as first made misfolded in a way to trap it intracellularly so that a library of molecules and compounds can be evaluated for their ability to interact with the target protein and correct its folding and cellular localization.

To generate misfolded versions of the CELSR2 TMD, we created as first, two mutant receptors fully or partially deleted of their extracellular domain. The CELSR2 ECD consists of 2350 aa: the juxta-membrane region (containing the EGF-laminin G and the Hormone Receptor domains) and the more distal region containing EGF-like, Laminin G and a string of tandemly repeated cadherin domains. We generated a cDNA coding for CELSR2(1917-2923), missing only the distal region, and CELSR2(2358-2617) a CELSR2 variant missing the entire ECD (**Fig. 32**).

We also tried to affect folding impairing the proteolytic processing of CELSR2 at the GPS. Several lines of evidences suggest that mutations in the GAIN domain and more specifically of the GPS consensus sequence might lead to protein folding defects, which would in turn impair the receptor trafficking. Hence, we introduced the single point mutations, H2355S and T2357G in the GPS consensus sequence of CELSR2(1917-2923) and CELSR2(2358-2617) to generate CELSR2(1917-2923)ΔGPS and CELSR2(2358-2617)ΔGPS (**Fig. 32**).

In third line, to increase the chance of obtaining a misfolded variant of CELSR2, we decided to replace its tail with the carboxyl-terminal tail of Frizzled-4 L501fsz533. This latter is a mutant of the class F GPCR receptor Frizzled-4 which is responsible *in vivo* for the Familial Exudative Vitreoretinopathy (FEVR), a pathology of the retina. This amino-acidic sequence, henceforth referred to as FEVR, is a disposable module inducing misfolding and aggregation. The FEVR tail causes the trapping of the misfolded receptor in the ER, in contrast to the wt receptor that is localized at steady state on the PM of the cells. We have already shown that when the FEVR tail is appended to the C-terminal tail of other transmembrane proteins (even Type I transmembrane protein like human CD8 and the viral VSV-G) it induces misfolding, aggregation and trapping of the chimeric proteins in the ER[65]. We decided to append the FEVR tail at the C-terminal of CELSR2(2358-2617), CELSR2(1917-2923)ΔGPS and CELSR2(2358-2617)ΔGPS to obtain the following mutants: CELSR2(1917-2617)FEVR, CELSR2(1917-2617)ΔGPS_FEVR and CELSR22(2358-2617)ΔGPS_FEVR (**Fig. 32**).

Wt targets localize on
Plasma Membrane

Ligand binding rescues folding and
localization of mutant targets to Plasma
Membrane
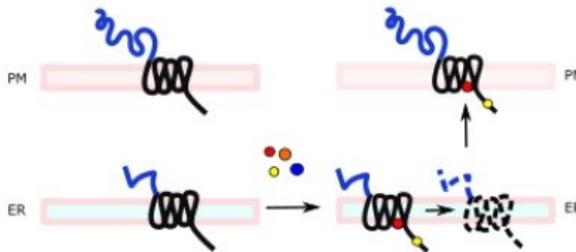
Mutant targets localize
intracellularly trapped in the
Endoplasmic Reticulum

Figure 31. Schematic representation of the PC-platform showing wt receptors localized at PM and mutant receptors being trapped in the ER. Molecules are screened for their ability to interact with the mutant receptor and correct its folding and localization.
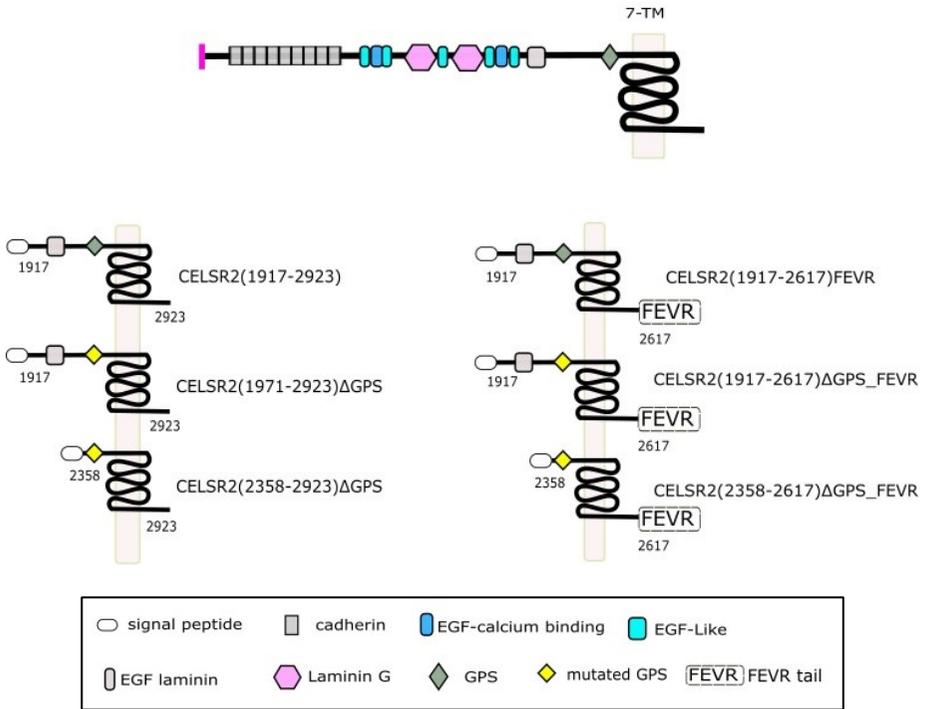
Figure 32. Schematic representation of CELSR2 wt and CELSR2 mutants generated to obtain a misfolded version of this receptor.

CELSR2 mutants expression in mammalian cells

CELSR2 constructs were expressed in mammalian cells (Hek293 and HUH7.5 cells). Their cellular localization was analyzed by using immunofluorescence and western blot analysis, 48 hours post transfection. Hek293 cells were used for Western Blot analysis for the easiness of their transfectability. HUH7.5 cells, on the other hand, were used to perform immunofluorescence analysis, since for their flat shape suit better immunofluorescence study of PM-localized proteins.

As shown by the immunofluorescence data, CELSR2(1917-2923) is correctly localized on Plasma Membrane (PM). In this case, the lack of the EGF-like, Laminin G and Cadherin domains did not affect the intracellular protein trafficking. Moreover, the auto proteolytic processing gives rise to the expected cleaved protein at 48KDa (**fig. 33 a**, **d**).

Mutations affecting CELSR2 folding should trap CELSR2 mutants in the Endoplasmic Reticulum (ER). CELSR2(1917-2923)ΔGPS presenting the two points mutations abolishing autocatalytic cleavage activity appears, as expected, mainly as full length protein of 109 KDa. However, this un-cleaved CELSR2 variant is still correctly localized at PM (**Fig. 33 a**, **d**). In our cellular system, the receptor cleavage is thus not essential for folding and trafficking to PM. Contrary to our expectations, also the membrane localization of the CELSR2(1917-2617)FEVR variant was not affected by the presence of the FEVR tail. Also for this construct, the mutant is normally cleaved and visible as a band at 48 KDa (**Fig. 33 b**, **d**). Similarly, the membrane localization of CELSR2(1917-2617)ΔGPS_FEVR is only partially affected but, as expected the mutant protein is resistant to the autoproteolytic cleavage (full length at 81 KDa) (**Fig. 33 b**, **d**). Finally, as predicted, the lack of the entire extracellular domain causes misfolding of CELSR2(2358-2617)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR variants. Indeed, the two mutants are uncleaved (full lengths at 61 KDa and 32.8 KDa respectively) unprocessed and totally insoluble in non-reducing conditions. Moreover, the IF data show that they are totally trapped in the ER (**Fig. 33 c, d**).

a

Celsr2(1917-2923)
Celsr2(1971-2923)ΔGPS

100KDa
uncleaved
75Da
cleaved
37Da

+DTT

Celsr2(1917-2923)
Celsr2(1971-2923)ΔGPS

100KDa
uncleaved
75Da
cleaved
37Da

-DTT

CELSR2(1917-2923)

CELSR2(1971-2923)ΔGPS

b

Celsr2(1917-2617)FEVR
Celsr2(1917-2617)ΔGPS_FEVR

100KDa
uncleaved
cleaved
75Da
37Da

+DTT

Celsr2(1917-2617)FEVR
Celsr2(1917-2617)ΔGPS_FEVR

100KDa
uncleaved
cleaved
75Da
37Da

-DTT

CELSR2(1917-2617)FEVR

CELSR2(1917-2617)ΔGPS_FEVR

d

Intracellular localization of CELSR2 mutants

100

50

0

Celsr2(1917-2923)
Celsr2(1971-2923)ΔGPS
Celsr2(1917-2617)FEVR
Celsr2(1917-2617)ΔGPS_FEVR
Celsr2(2358-2923)ΔGPS
Celsr2(2358-2617)ΔGPS_FEVR

%PM
%ER

c

Celsr2(2358-2923)ΔGPS
Celsr2(2358-2617)ΔGPS_FEVR

insolubule

75Da
uncleaved

uncleaved

+DTT

Celsr2(2358-2923)ΔGPS
Celsr2(2358-2617)ΔGPS_FEVR

insolubule

-DTT

CELSR2(2358-2923)ΔGPS

CELSR2(2358-2617)ΔGPS_FEVR

Figure 33. The deletion of the whole extracellular domain of CELSR2 causes the misfolding and trapping of the protein in the ER. (**a**) Hek293 and HUH7.5 cells were transfected with vectors coding CELSR2 mutants. Cellular localization of each mutants was analyzed by Western Blot and by immunofluorescence (**a,b,c**). Western Blot results are shown on the left side. 48 h after transfection Hek293 cells were lysed in TritonX-100. Proteins were denatured in SDS at 37 °C for 20 min, and run on either reducing or non-reducing SDS-PAGE and processed for W.B. Filters were decorated with an anti-V5 antibody to reveal CELSR2 (only relevant part of the gels are shown). GPS Point mutations abolish cleavage activity. Full length and cleaved proteins are shown. Deletion of the whole extracellular domain makes the proteins insoluble in non-reducing conditions. For immunofluorescence analysis, CELSR2 mutants were expressed in HUH7.5 cells (results are showed on the right side) immunofluorescence of permeabilized (intracellular) cells showing PM and ER localization of CELSR2 mutants (red). Deletion of the full extracellular domain causes the trapping of the proteins in the ER. (**d**) Percentage of PM and ER localization of CELSR2 mutants 48 h after transfection in Hek293 cells (n=3 independent experiments, co-localization was calculated in 30 cells for each replicate).

We thus obtained two CELSR2 variants misfolded and trapped in the ER. They are designed as CELSR2(2358-2617)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR and are characterized by the deletion of the extracellular domain.

Then we proved that CELSR2(2358-2617)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR folding and plasma membrane localization could be rescued. We attempted to rescue the folding of the CELSR2 misfolded variants by co-expressing *in trans* their extracellular domains. We generated constructs expressing both the string of CELSR2 cadherin domains and the EGF-like-laminin G domains. We proved that when CELSR2(2358-2617)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR constructs were co-expressed with the corresponding isolated CELSR2(Cadherin domain), the plasma membrane localization of these constructs was rescued in 80% and 60% of the cell population, respectively (**Fig. 34 a, b, c**). In contrast, CELSR2 mutants, when co-expressed with the isolated CELSR2(EGF-like Laminin G domains), are still misfolded and trapped in the ER (**Fig. 34 a, b**).

Finally, we decided to use extracellular deletion mutants of CELSR2 to generate the biological platform for the screening campaign. The construct CELSR2 (Cadherin domain) was used as PC-Platform positive control.
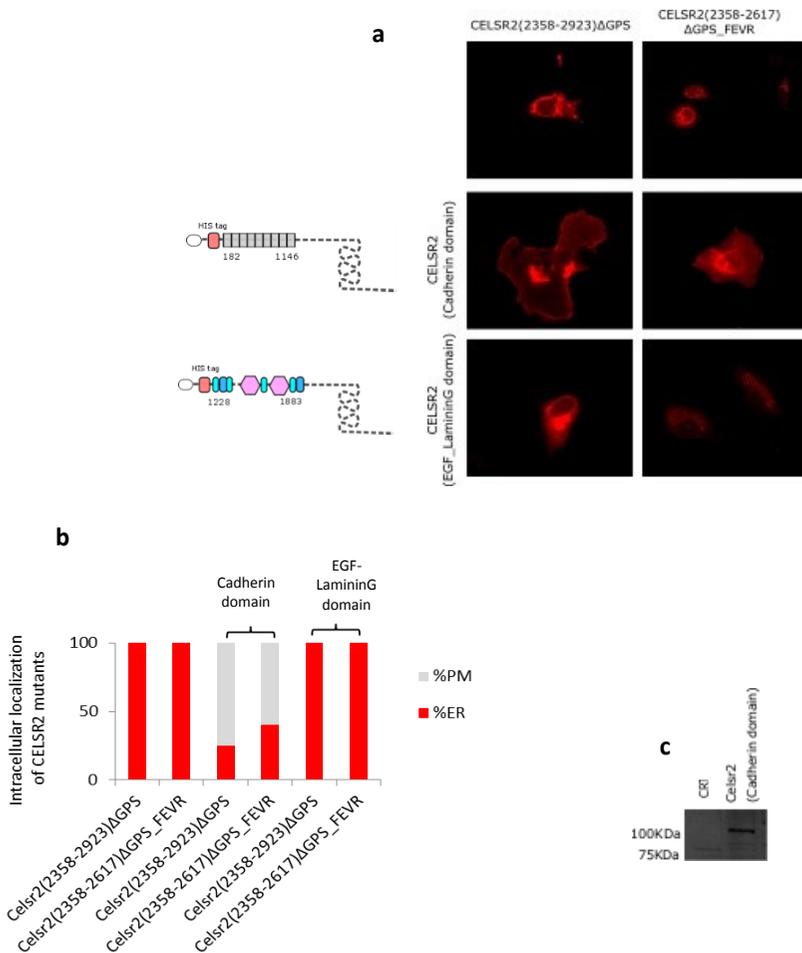
Figure 34. Rescue of CELSR2(2358-2923)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR upon co-transfection with CELSR2 (Cadherin domain) in HUH7.5 cells. (a) Immunofluorescence of permeabilized cells showing ER localization of CELSR2 mutants and rescue of CELSR2 mutants localization at PM after co-transfection with the CELSR2(Cadherin domain). CELSR2 (EGF_Laminin G domain) failed in rescuing CELSR2 mutants PM localization. (b) Percentage of PM and ER localization of CELSR2(2358-2923)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR before and after co-transfection with CELSR2 (Cadherin domain) and CELSR2 (EGF_Laminin G domain) (n=3 independent experiments, co-localization was calculated in 30 cells for each replicate). (c) W.B analysis of CELSR2(Cadherin domain) expression in Hek293 cells 48h after transfection.

89

Generation of CELSR2 endogenous metabolite library

We decided to create a metabolite library by using Zebrafish as biomass. Indeed, there is an high level of homology between aGPCRs found in human and zebrafish genome. Expression profiles of zebrafish adhesion GPCRs during development[124] shows CELSR2 enrichment during gastrulation and segmentation stages and specifically between 10 and 15 hours post fertilization. For this reason, we decided to collect Zebrafish eggs (more then 100.000) at 15 hours post fertilization and then, we performed a three phase extraction of the metabolites. We used the organic solvents N-hexane, butanol and ethyl acetate to obtain hexane (25 mg), ethyl acetate (7 mg), butanol (26 mg) and residual aqueous fractions (242 mg) respectively. Ethyl acetate, butanol and aqueous fractions, were re-suspended in dimethyl sulfoxide (DMSO) at a concentration of 10 mg/ml. Hexane fraction was insoluble in DMSO (**Fig. 35**). Fractions obtained were screened for pharmacological chaperone efficiency.
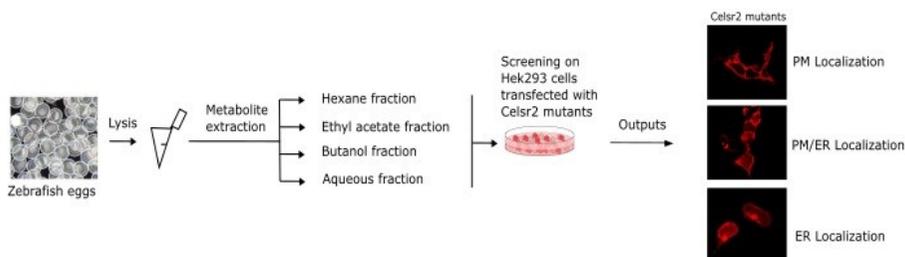
Figure 35. Generation of a metabolite library to be screened by using PC-platform. Zebrafish eggs were used as biomass to obtain a metabolite library. After a three phase metabolite extraction, fractions obtained were screened for pharmacological chaperone efficiency on Hek293 cells expressing CELSR2(2358-2923)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR.

Rescue of ER-trapped CELSR2 mutants as readout to identify CELSR2 modulators

Hek293 cells were transfected with vectors coding for CELSR2(2358-2617)ΔGPS and CELSR2(2358-2617)ΔGPS_FEVR mutants; 4 hours after transfection cells were treated with ethyl acetate, butanol or aqueous fractions at 3 different final concentrations: 100 μM, 10 μM and 1 μM. The dilutions were performed in 10% FBS supplemented DMEM and after centrifugation at 112 RCF, fractions were divided in two parts: insoluble and soluble (named fraction 1 and 2 respectively). 72 hours after transfection, chaperone activity was assayed by measuring the recovery of CELSR2 mutants localization at the PM. We used DMSO as negative control. Soluble ethyl acetate fraction at the concentration of 100μM rescued CELSR2(2358-2923)ΔGPS localization at the PM in the 40% of the cell population (**Fig. 36**). Also insoluble ethyl acetate and butanol fractions (100 μM) rescued CELSR2(2358-2923)ΔGPS at the PM in the 20% of the cell population (**Fig. 36**). On the contrary, butanol and aqueous fractions either at 100 μM or 10 μM were able to rescue CELSR2(2358-2617)ΔGPS_FEVR at PM in than the 20% of the cell population (**Fig. 37**).

We obtained different fractions that are active on rescuing the folding of the two CELSR2 variants. Moreover, the mutations characterizing these constructs induce deep changes in the protein conformation so that their folding can be modulated by compounds with different chemical properties.
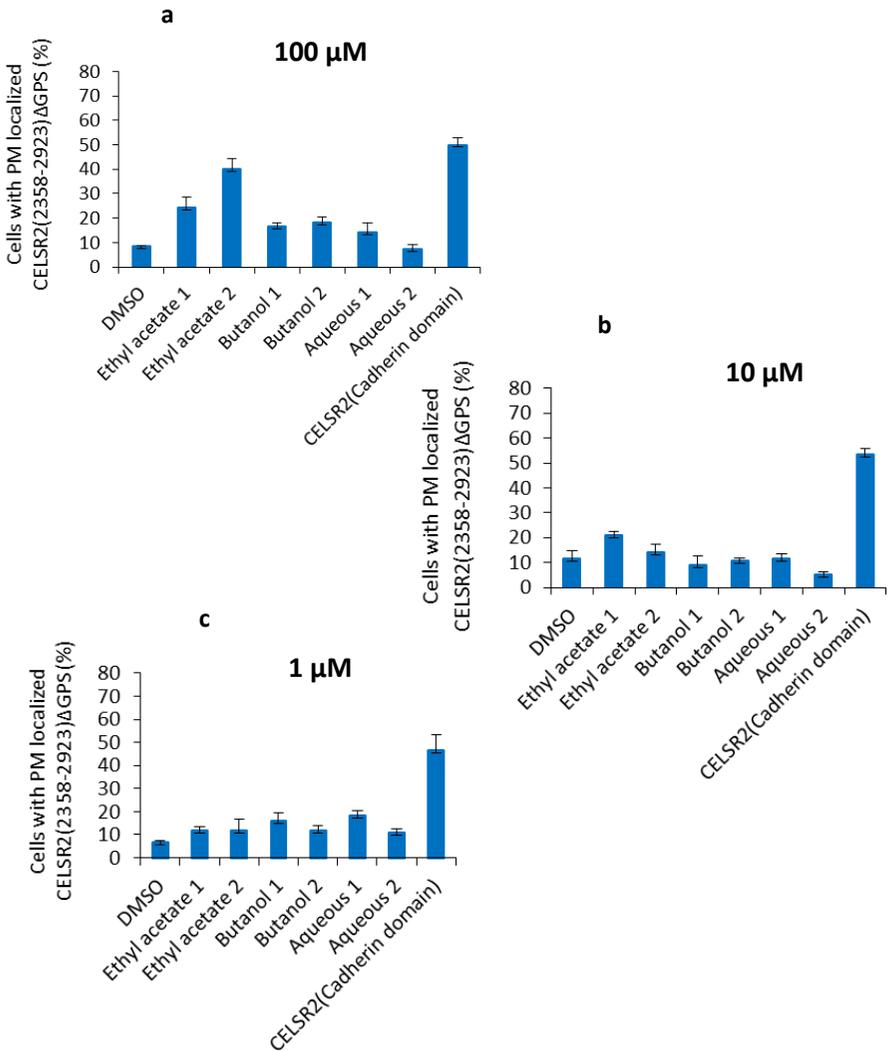
Figure 36. Rescue of CELSR2(2358-2923)ΔGPS PM localization as biological platform toward the identification of CELSR2-wt modulators. Bar graphs representation of metabolite fractions that rescue of CELSR2(2358-2923)ΔGPS PM localization (n = 3, ten random fields containing at least 50 cells counted for each replicate; s.d. are indicated). Each fraction was tested at three final concentrations: 100 μM (**a**), 10 μM (**b**) and 1 μM (**c**). The fraction dilutions were performed in 10% FBS supplemented DMEM and after centrifugation at 112 RCF, each fraction was divided in two parts: insoluble and soluble (fraction 1 and 2, respectively).
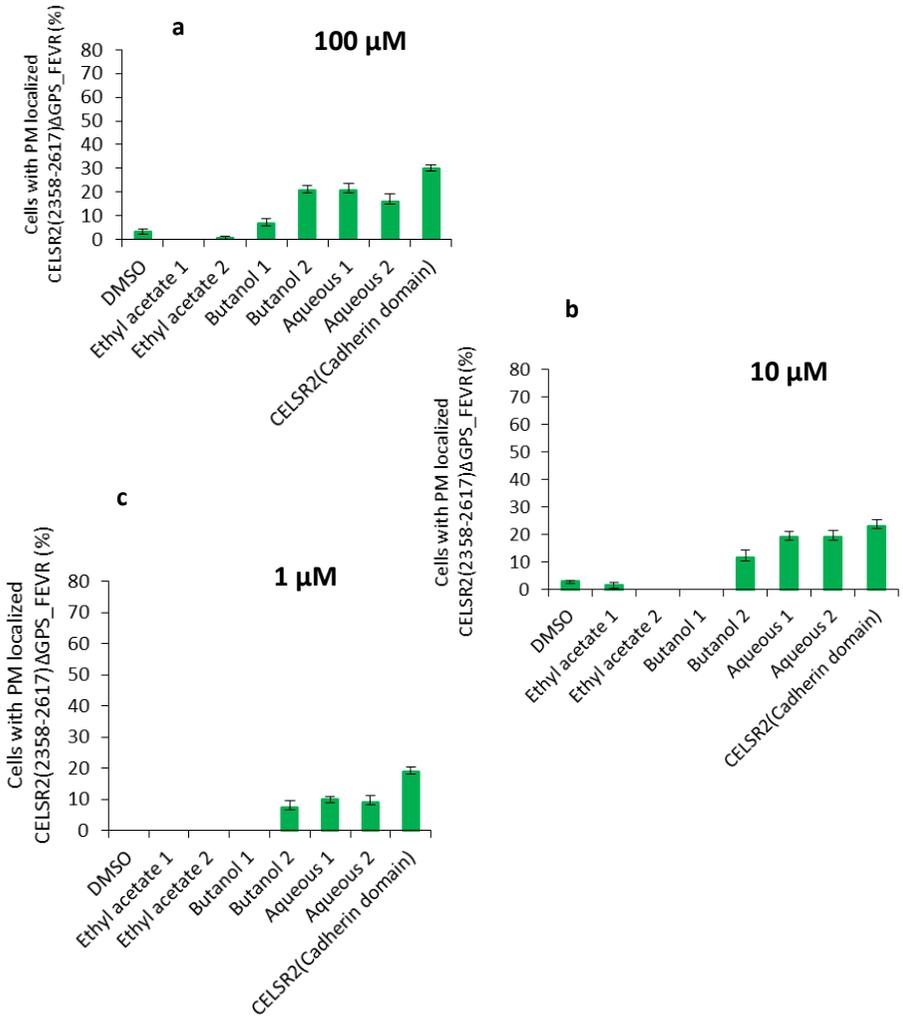
Figure 37. Rescue of CELSR2(2358-2923)ΔGPS_FEVR PM localization as biological platform toward the identification of CELSR2-wt modulators. Bar graphs of metabolite fractions rescue of CELSR2(2358-2923)ΔGPS_FEVR PM localization (n = 3, ten random fields containing at least 50 cells counted for each replicate; s.d. are indicated). Each fraction was tested at three final concentration: 100 μM (**a**), 10 μM (**b**) and 1 μM (**c**). The fraction dilutions were performed in 10% FBS supplemented DMEM and after centrifugation at 112 RCF, each fraction was divided in two parts: insoluble and soluble (fraction 1 and 2, respectively).

Three molecules from Ethyl acetate fraction act as CELSR2 modulators

We collected information about the chemical composition of the most active fraction ethyl acetate. We used Direct Infusion Fourier Transform-ion cyclotron resonance mass spectrometry (DI-FT-ICR-MS), which is characterized by unmatched ultra-high mass accuracy and resolution. FT-ICR analysis revealed that this fraction contained carotenoids, prostaglandinE2 (PGE2), bile acids, and cholesterol (**Supplementary Fig. 1, Fig. 2, Fig. 3 and Fig. 4**). For this reason, we decided to evaluate the pharmacological chaperone efficiency of these molecules. We transfected Hek293 cells with CELSR2(22358-2923)ΔGPS and 4 hours after transfection we treated them with the above mentioned compounds. Chaperone activity was assayed by measuring the recovery of the CELSR2 mutant localization at the PM, 72 hours after transfection. Among the tested candidates, β-carotene rescued CELSR2(2358-2617)ΔGPS PM localization with an efficiency higher than 40%. In contrast, Taurocholic acid failed in rescuing CELSR2(2358-2617)ΔGPS PM localization. PGE2 and cholesterol had a similar effect with a maximum of 15% of cells with the mutant protein at PM. However, PGE2 at the highest concentration was toxic for cells (**Fig. 38**). We showed that the EC50 of β-carotene was ~ 1 μM, on the contrary PGE2 and cholesterol EC50 was ~ 2 μM and 1 μM, respectively (**Fig. 38**).

We also proved that β-carotene, PGE2 and cholesterol did not rescue either Fz4-FEVR or the mutant CFTRΔF508, indicating that the molecules did not act as promiscuous molecular chaperones (**Fig. 39**).
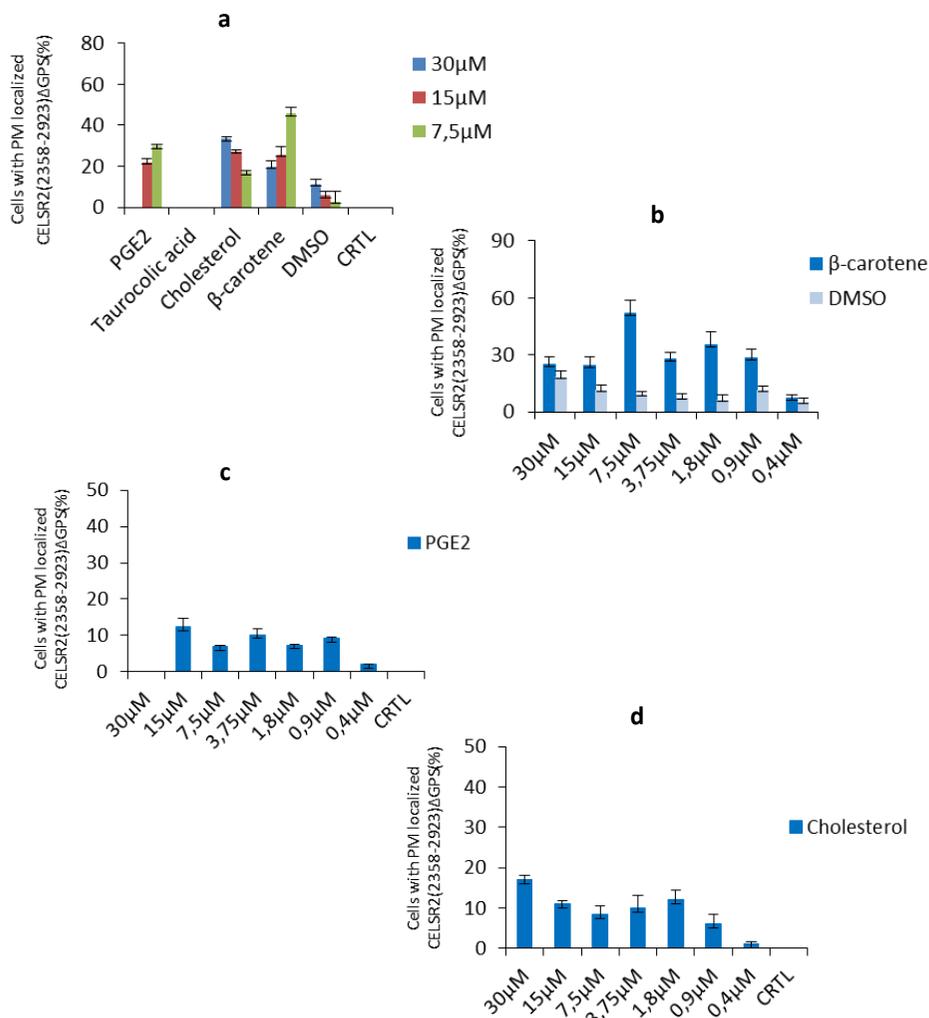
Figure 38. Screening of compounds composing the ethyl acetate fraction for rescuing CELSR2(2358-2923)ΔGPS PM localization. Bar graphs of β-carotene, PGE2 and cholesterol rescue of CELSR2(2358-2923)ΔGPS PM localization (n = 3, ten random fields containing at least 50 cells counted for each replicate; s.d. are indicated). In a preliminary experiment (**a**) compounds composing the ethyl acetate fraction were tested at three final concentrations: 30, 15 and 7,5 μM. Taurocholic acid failed in rescuing CELSR2 mutant localization and was excluded from further tests. On the contrary, β-carotene (**b**), PGE2 (**c**) and cholesterol (**d**) were tested in a wider range of concentrations (30 μM-0,4 μM).

95
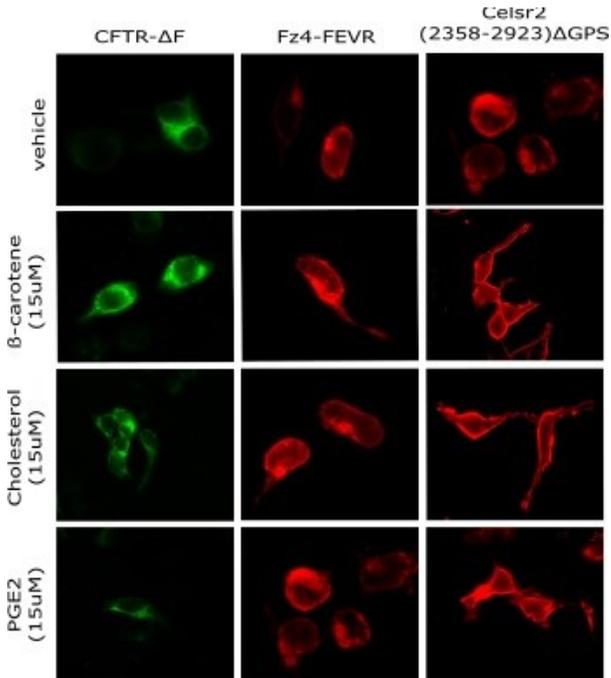
Figure 39. β-carotene, PGE2 and cholesterol specifically rescue CELSR2(2358-2923)ΔGPS PM localization. β-carotene, PGE2 and cholesterol failed in rescuing the PM localization of the mutants CFTR-ΔF and Fz4-FEVR. Intracellular localization of GFP tagged CFTR-ΔF (ER, on the left) and HA-tagged Fz4-FEVR (ER, in the middle) transiently transfected in Hek293 cells after treatment with (15 μM β-carotene, 15 μM PGE2 and 15 μM cholesterol for 48 hours) or with vehicle alone. The right column shows the effect of the compounds on cells expressing CELSR2(2358-2923)ΔGPS.

Binding site prediction

According to our binding site prediction, cholesterol allocates at the interface between the ICL1 and the juxta-membrane region of TM2, TM3 and TM4 of CELSR2 (**Fig. 40 a**). Cholesterol hydroxyl group forms ionic bonds with the amino group of Asp2414 or Arg2413 of the ICL1. The four linked hydrocarbon rings are located between the apolar residues Ala2417 and Leu2421 of the TM2 and the polar residues Trp2487 and Tyr2483 of TM4. The aliphatic chain protrudes in an apolar cavity formed by Ala2490 of TM4 and by Leu2445 and Leu2449 of the TM3. According to our model, two molecules of cholesterol can be localized at the same binding site with their hydroxyl groups interacting with the hydrophilic region of the ICL1 and the four hydrocarbon rings with the aliphatic chain enclosed in the hydrophobic region formed by the transmembrane helices (**Fig. 40 b**).

Similarly to cholesterol, β-carotene derivative compound retinoic acid, allocates at the interface between the ICL2 and the juxta-membrane region of TM3, TM4 and TM5 of CELSR2 **(Fig. 41)**. The carboxylic group forms an ionic interaction with Arg2469 of the ICL2. The aliphatic chain lodges in a cavity formed by apolar residues of the transmembrane helixes with the cyclohexenyl ring interacting with Phe2532 and Pro2529.

According to our *in silico* model, the PGE2 carboxylic group interacts with Leu2519 of the ECL1. The hydroxyl group at the C9 of the cyclopentanone ring interacts with the Ser2522 of the TM5. The two aliphatic side chains mostly contact apolar residues of the TM3, TM4 and TM5 (**Fig. 42)**.
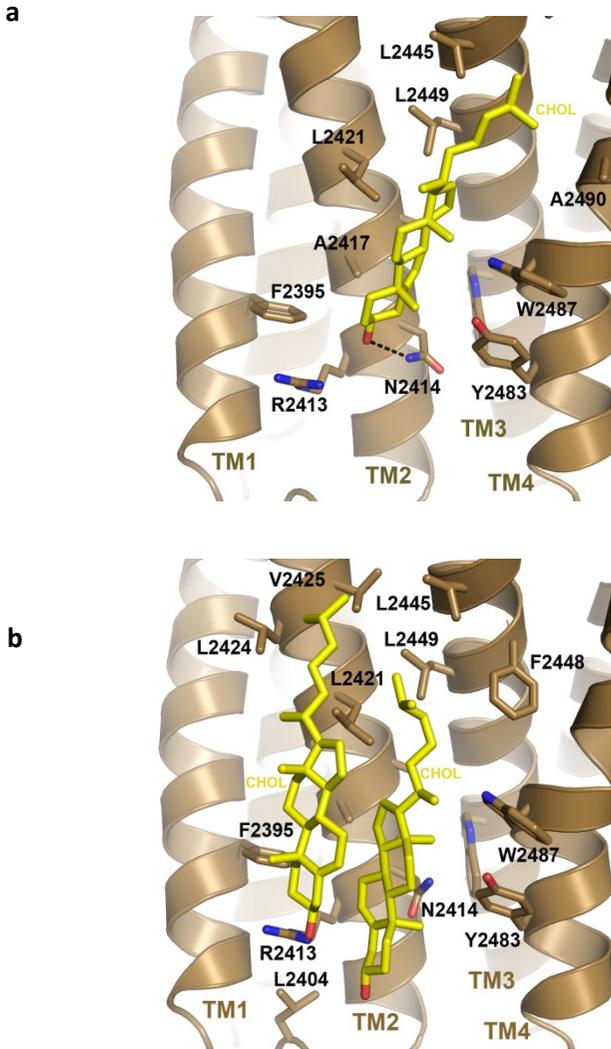
Figure 40. Cholesterol ligand binding site prediction. (a) Cholesterol allocates at the interface between the ICL1 and the juxta-membrane region of TM2, TM3 and TM4 of CELSR2. (**b**) Two molecules of cholesterol can be localized in the same binding site.
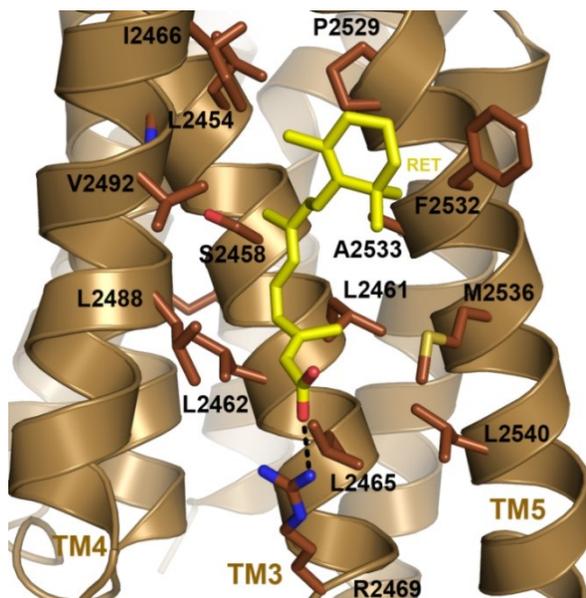
Figure 41. Retinoic acid ligand binding site prediction. The β-carotene derivative compound retinoic acid, allocates at the interface between the ICL2 and the juxta-membrane region of TM3, TM4 and TM5 of CELSR2.
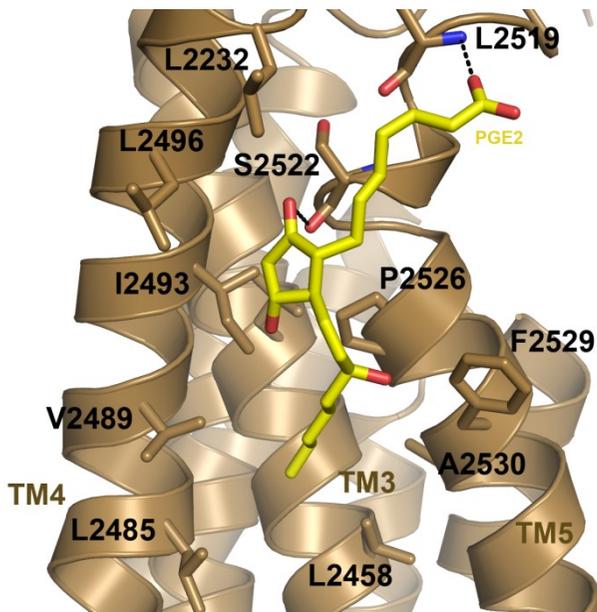
Figure 42. PGE2 ligand binding site prediction. PGE2 is located at the interface between the ECL1 and the juxta-membrane region of the TM3, TM4 and TM5 of CELSR2.

DISCUSSION

CELSR2 in an adhesion GPCR overexpressed in Glioblastoma. To date, ligands binding CELSR2 have not yet been identified, thus it is classified as an orphan GPCR. In the last decade, the deorphanization rate has drastically slowed down because of the inapplicability of the traditional ligand screening platforms. Indeed, these latter identify ligands in virtue of their ability to either displace reference ligands or to affect signaling pathway linked to the target. However, we do not have any information about reference ligands and signaling pathways of orphan GPCRs. Recently we have ideated a new screening platform (PC-platform) that allows to identify new ligands in virtue of their ability to act as pharmacological chaperones. Here, we applied the PC-platform for CELSR2 deorphanization. We generated misfolded and ER-trapped versions of CELSR2 and then different metabolite fractions extracted from zebrafish eggs were screened for their ability to correct folding and localization of CELSR2 variants. β-carotene, cholesterol and ProglandinE2, which belongs to the most active fraction ethyl acetate, were able to rescue CELSR2 mutant folding and localization.

We identified the hypothetical binding site of these molecules on CELSR2. They do not target the canonical orthosteric binding site of CELSR2 but seem to occupy hydrophobic cavities at the interface between the TM bundle and the lipid bilayer. By interacting with these regions, they probably hinder the receptor protecting it from aggregation.

The identification of cholesterol and PGE2 as putative allosteric modulators of CELSR2 is consistent with the recent theory about the existence of "orphan" allosteric sites for, until now, unappreciated endogenous modulators[135]. Indeed, there are several examples of endogenous substances that can act as allosteric modulators of GPCRs. These include a variety of ions, lipids, amino acids, peptides, and accessory proteins that display different degrees of receptor specific modulatory effects[135]. In particular, it has been widely proved the cholesterol modulatory role on the function and ligand-binding properties of several G-protein-coupled receptors (GPCRs). Indeed, crystal structures of prototypical GPCRs such as the adenosine A_{2A} receptor ($A_{2A}R$) have confirmed that cholesterol finds stable binding sites at the receptor interface with the lipid bilayer[136]. Moreover, it has been proved that the cholesterol inside the receptor allosterically modulate $A_{2A}R$-binding affinity [136]. Substantial interest in the potential for cholesterol to directly interact with GPCRs gained further impetus upon the solution of a 2.8-Å structure of the β2-AR simultaneously bound to two molecules of

cholesterol, and the identification in this binding site of a cholesterol consensus motif (CCM)[137]. Moreover, cholesterol depletion has been shown to affect receptor signaling. Indeed in this condition, the extracellular signal-regulated kinases 1/2 (ERK1/2) phosphorylation and *c-Fos* activation for the gonadotrophin-releasing hormone receptor are lost, but can be restored upon replenishment[138].

Interestingly, the increasing identification of lipids or inflammatory molecules as GPCRs allosteric modulators indicates that disease context plays an important role in the generation of putative endogenous GPCR modulators. In a disease context, the presence of receptor-specific endogenous modulators likely points to a need for particularly tight regulation of certain GPCRs. For example, the ability of cholesterol to influence GPCR biology could depend on cholesterol levels that can vary dramatically in disease[139]. Similarly, the increase of the inflammatory mediators as a consequence of inflammatory processes likely introduce a plethora of substances, among these PGE2, that could otherwise go undetected as potential allosteric modulators. It is possible that the study of inflammation represents fertile ground for further discoveries in GPCR allostery.

Some experimental evidences showed that cholesterol is required for dendrite differentiation, that is the rate limiting step for glia-induced synaptogenesis[140]. Specifically, *Goritz et al.*[140] showed that cholesterol enhances directly dendritic differentiation and that it is essential for continuous synaptogenesis. So far, however, it remained unclear how cholesterol enhances the formation and function of synapses. Interestingly, structure-phenotype analysis showed that CELSR2 plays an important role in dendrite development, from Drosophila to mammals[92]. Kimura et al.[119] experimental evidences suggested that during *Drosophila* dendritic development, CELSR2 could act as a receptor for an unknown ligand to control branch elongation. Thus, according to our data, cholesterol binding on CELSR2 could regulate this process either by direct modulation of CELSR2 activity or by modulating receptor clustering. Such an effect has been showed for Glutamate receptors, that are other receptors involved in synaptogenesis, whose postsynaptic clusters are sensitive to the neuronal cholesterol content. The hypothesis of cholesterol as an allosteric modulator of CELSR2 is also supported by the identification of a specific amino acid residue in the cholesterol ligand binding site on CELSR2, that is responsible for the receptor function in enhancing dendritic growth, as showed by site directed mutagenesis experiments[104]. This residue, the Arg2413, is localized in the first intracellular loop of CELSR2 and can form an ionic interaction with the cholesterol hydroxyl group.

Similarly, retinoic acid as potential allosteric modulator of CELSR2 could play an important role in neuronal differentiation. Indeed, it has been widely proved the involvement in the neuronal development either of CELSR2 [113] and retinoic acid [141]. Moreover, it has been shown that in an *in vitro* model of the CNS differentiation process, the retinoic acid requires the presence of ciliary proteins such as Meckelin and Jouberin to induce the neural differentiation of mouse embryonic stem cells[142]. Interestingly, CELSR2 has been proved to regulate ciligionesis and its role in human ciliopathies was recently emphasized with the demonstration that mutated CELSR2 is implicated in Joubert syndrome[113]. Thus, CELSR2 function as a ciliary protein could play an important role in retinoic acid-dependent neuronal differentiation.

ProstagladinE2 has been shown to play an important role in physiological processes such as cell migration[143]. These processes are involved in gastrulation movements and it has been shown that decreased prostaglandinE2 synthesis halts gastrulation[143]. However, to date, how PGE2 regulates cell motility remains unclear. *Speirs et al.*[143] showed that PGE2-deficient zebrafish embryos, impaired in gastrulation movements, exhibit markedly increased cell-cell adhesion, which contributes to defective cell movements in the gastrula[143]. Similarly, in zebrafish embryos, loss of function of CELSR2 results in migration neuronal defects. *Barbosa et al.* showed that a C-terminal truncated form of CELSR2 affects cell cohesion that primarily leads to the *in vivo* cell movement defects[116]. These data suggest that PGE2 could regulate cell migration by modulating CELSR2 function in cell-to-cell contacts and that PGE2 binding to CELSR2 could require the receptor C-terminal tail to initiate a signalling pathway. One of the signalling pathways that PGE2 could modulate through CELSR2 is PCP[120]. Indeed, phenotype of CELSR2 mutants is similar to that of FZD3 mutants, and other PCP proteins also implicated in cell migration[120].

The potential ability of PGE2 to allosterically modulate CELSR2 to regulate cell adhesion and cell migration, could have important implications also in pathological processes such as cancer metastasis[143]. Indeed, both CELSR2 and PGE2 expression levels have been proved to be associated to cancer cell invasiveness. This is also consistent with the identification of single amino acid substitutions (L2496I, P2526Q and A2530V) in the PGE2 ligand binding site on CELSR2 that have been associated to the highly invasive tumors esophageal cancer, melanoma and urinary bladder cancer, respectively.

Moreover, also mutations of some amino acid residues involved in the ligand binding sites of cholesterol and retinoic acid have been associated to cancer. In the cholesterol binding site, the single amino acid substitutions R2413L and W2487C A2417T have been associated to liver cancer, stomach cancer and lung cancer, respectively . On the contrary, in the retinoic binding site, the single amino acid substitutions F2454V and M2533I have been associated to uterine and colorectal cancer, respectively. Our data suggest that, in some cases, loss of the GPCR allosteric modulation could be implicated in tumorigenesis. This is consistent with the current idea of using allosteric modulators as a new pharmacological strategy in the cancer therapy. Moreover, allosteric modulators as new GPCR-targeted drugs provide several mechanistic advantages, including the ability to distinguish between closely related receptor subtypes. Recently, the discovery of allosteric ligands that confer bias and modulate some, but not all, of a given receptor's downstream signaling pathways can provide pharmacological modulation of cancer and other pathological conditions with remarkable precision[144].

None of the molecules till here identified as CELSR2 modulators, is predicted to target the orthosteric binding site of CELSR2. This could suggest that, even in the absence of the ECD, the CELSR2 TMD does not use its orthosteric binding site to allocate any molecule. This is in a way supported by the fact that while the ECD of CELSR2 is highly conserved among species, the TM bundle is less conserved.

Finally, the allosteric modulators we identified derive from the apolar ethyl acetate fraction. Although less effective than the ethyl acetate fraction, also the aqueous fraction was active on rescuing CELSR2 mutant. This polar fraction represent a new source of molecules that could target the orthosteric binding site of CELSR2.

MATERIALS AND METHODS

Reagents. Salts and organic solvents are from Sigma Aldrich (USA) and Applichem (Germany), are handled according to manufacturer instructions.

Cell culture. Hek293 cells and HuH7.5 cells were cultured in DMEM (Dulbecco's Modified Essential Medium) supplemented with 10% FBS (Fetal Bovine Serum), Glutamax and Pen/Strep antibiotics. Hek293 transfections were performed with polyethylenimine (PEI) according to the manufacturer's protocol. On the contrary, Huh7.5 and U87MG transfections were performed with Lipofectamine 2000 according to manufacturer's protocol. OrphG cell clones are obtained by transfecting cells with the corresponding cDNA cloned in pcDNA 3.1(Invitrogen).

Mutagenesis. All of the mutants of CELSR2 were obtained by site-directed mutagenesis using CELSR2 cDNA (Origene) as a template.

Antibodies Rabbit polyclonal anti-V5 (Sigma-Aldrich) was used at 1:1000 dilution for immunofluorescence (IF) and 1:2000 dilution for western blot (WB); Mouse monoclonal anti-HA (HA-7, Sigma-Aldrich) was used at 1:2000 dilution for immunofluorescence (IF); Rabbit anti-CELSR2 abcam was used at 1:500 dilution for Western blot (WB)**;** Other antibodies used: rabbit polyclonal anti-HA (H6908, Sigma-Aldrich; dilution of 1:200 for IF); Peroxidase-conjugated anti- rabbit IgG (A0545, Sigma-Aldrich WB 1:5,000); Texas Red–conjugated anti-rabbit IgG (111-025-003, Jackson ImmunoResearch Laboratories; 1:500 dilution for IF); Texas Red-conjugated anti-mouse (115-025-003, Jackson ImmunoResearch Laboratories) at 1:500 dilution for IF;

Immunofluorescence. Cells were grown up on glass coverslips and then were fixed in 3,7% formaldehyde/PBS pH 7.4 freshly made for 30 minutes at room temperature (RT). Formaldehyde was quenched by incubating the coverslips for 30 minutes in 0.1 M glycine/PBS. Cells were permealized in 0.1% Triton/PBS pH 7.4 for 10 minutes at RT. Glass coverslips were then incubated with primary and secondary antibodies diluted in PBS for 1 h and 30 min at RT, respectively. After fixing coverslips with glycerol/PBS 1:1 on slides, cells were analyzed with a Leica TCS-SMD-SP5 confocal microscope.

Western blotting. Samples were diluted in 20 mM Tris HCl (pH 6.8), 50 mM DTT, 1% SDS, 5% glycerol and bromophenol blue and then incubated for 20 min at 37° C. Samples were loaded on a SDS-PAGE gel. The run was performed at 100 V at 25° C.

After the run, proteins were transferred onto a nitrocellulose filter (Schleicher-Schuel) at 80 V for 1 h at 4° C. Filters were blocked in PBS supplemented with 3% nonfat dry milk (Bio-Rad) for 2 h at 25° C and incubated with primary and secondary antibody diluted in PBS supplemented with 0.3% nonfat dry milk in PBS. The ECL reaction was performed using the Lumi Light ECL Kit (Roche) according to the manufacturer's procedures.

Metabolite three phase extraction from Zebrafish eggs. More than 100.000 zebrafish eggs were lysed and the solution was evaporated to dryness on a rotatory vacuum evaporator set at a maximum temperature of 40° C. This constituted the crude extract, which was dissolved in distilled water to be partitioned subsequently with n-hexane, ethyl acetate and n-butanol to afford non-polar, semi-polar and polar fractions respectively. The extracts were weighed and stored at -20° C. The hexane (25 mg), ethyl acetate (7 mg), butanol (26 mg) and residual aqueous fractions (242 mg) were dissolved in DMSO (10mg/ml) and further separated after centrifugation at 112 RCF in an insoluble and soluble fraction.

Mass spectrometry analysis. Analyses were performed in direct infusion following a previous protocol [31] employing a Hamilton syringe (250 µl) at a flow rate of 2 µL/min. Data were acquired on a SolariX XR 7T (Bruker Daltonics, Bremen, Germany). The instrument was tuned with a standard solution of sodium trifluoracetate. Mass Spectra were recorded in broadband mode in the range 100–1500 m/z, with an ion accumulation of 20 ms, with 32 scans using 2 million data points (2M). Nebulizing (N2) and drying gases (air) were set at 1 and 4 ml/min, respectively, with a drying temperature of 200 °C. Both positive and negative ESI ionization were employed. Five replicates of each injection were carried out. The instrument was controlled by Bruker FTMS Control, MS spectra were elaborated with Compass Data Analysis version 4.2 (Bruker), identification of compounds based on accurate MS measurements was performed by Compound Crawler ver. 3.0 and Metaboscape 3.0 (Bruker). Metabolites signals were normalized using internal standards. Comparisons and differences were analyzed for statistical significance by two-way Anova test and Bonferroni posttests analysis. All graphs, bars or lines indicate mean and error bars indicate standard error of the mean (s.e.m.). Furthermore, Partial Least Square Discriminant Analysis (PLS-DA) was used as classification model. All graphs, bars or lines indicate mean and error bars indicate standard error of the mean (s.e.m.). Statistical analysis was performed using Statistica software (StatSoft, Oklahoma, USA) and Minitab (Minitab Inc, Pennsylvania, USA).

SUPPLEMENTARY FIGURES
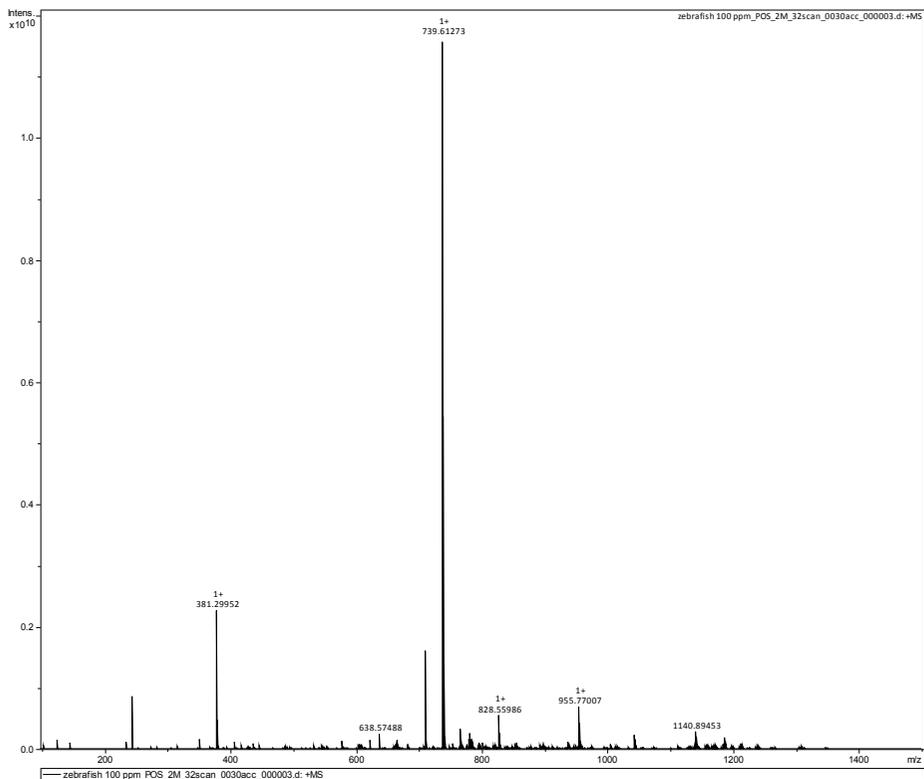
Supplementary Figure 1. Positive electrospray ionization (ESI)
Fourier-transform ion cyclotron resonance (FT-ICR) of the Ethyl
acetate fraction.
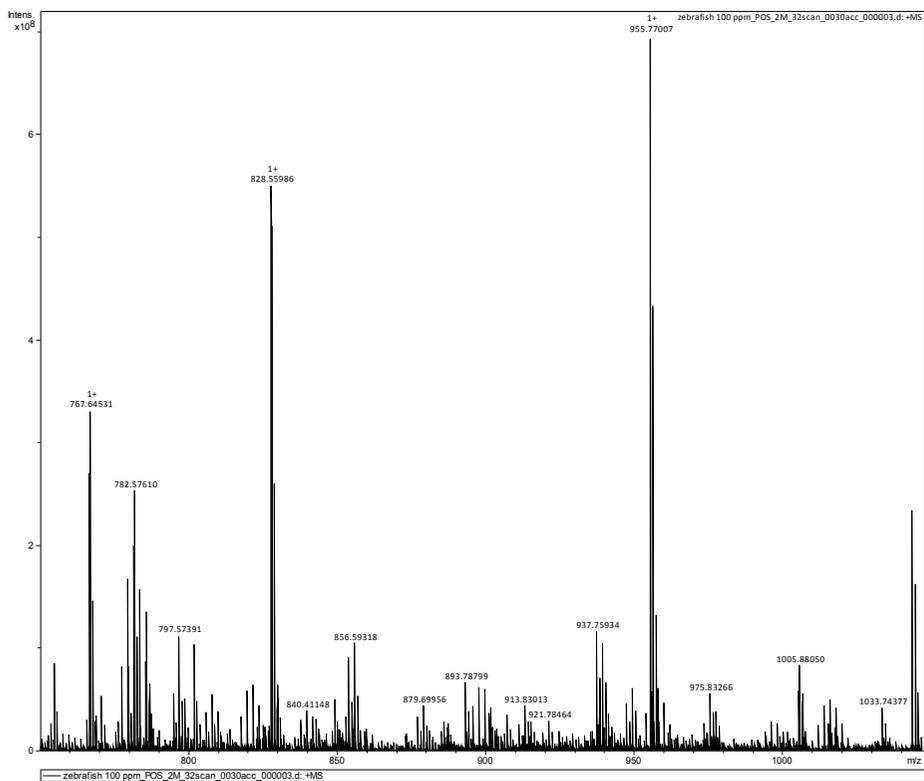
Supplementary Figure 2. Expansion of the Positive electrospray ionization (ESI) Fourier-transform ion cyclotron resonance (FT-ICR) of the Ethyl acetate fraction.
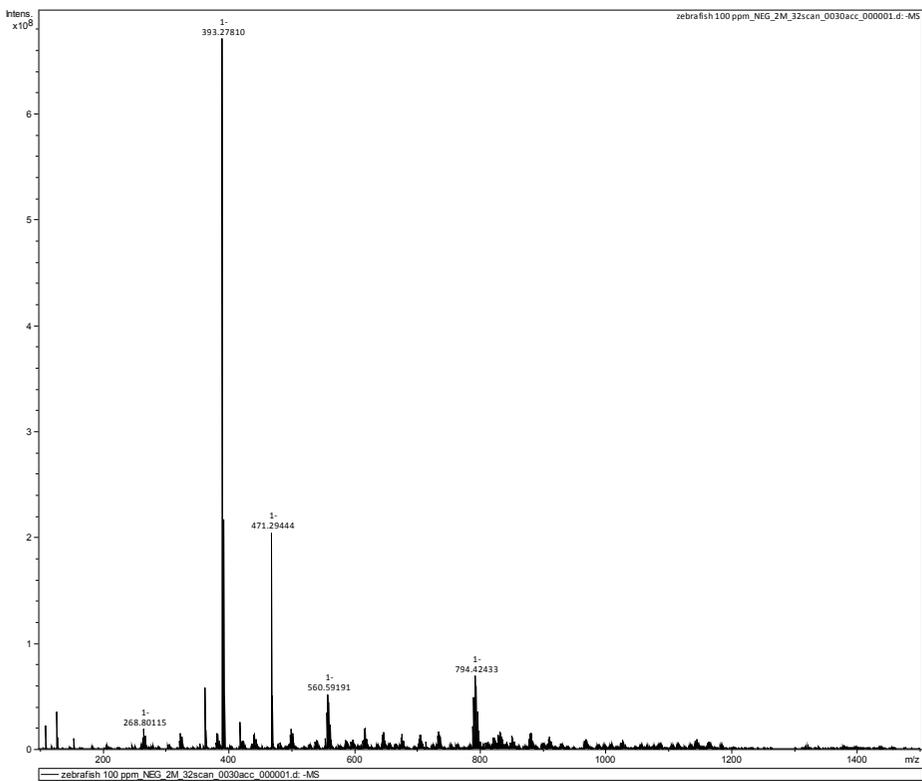
Supplementary Figure 3. Negative electrospray ionization (ESI) Fourier-transform ion cyclotron resonance (FT-ICR) of the Ethyl acetate fraction.
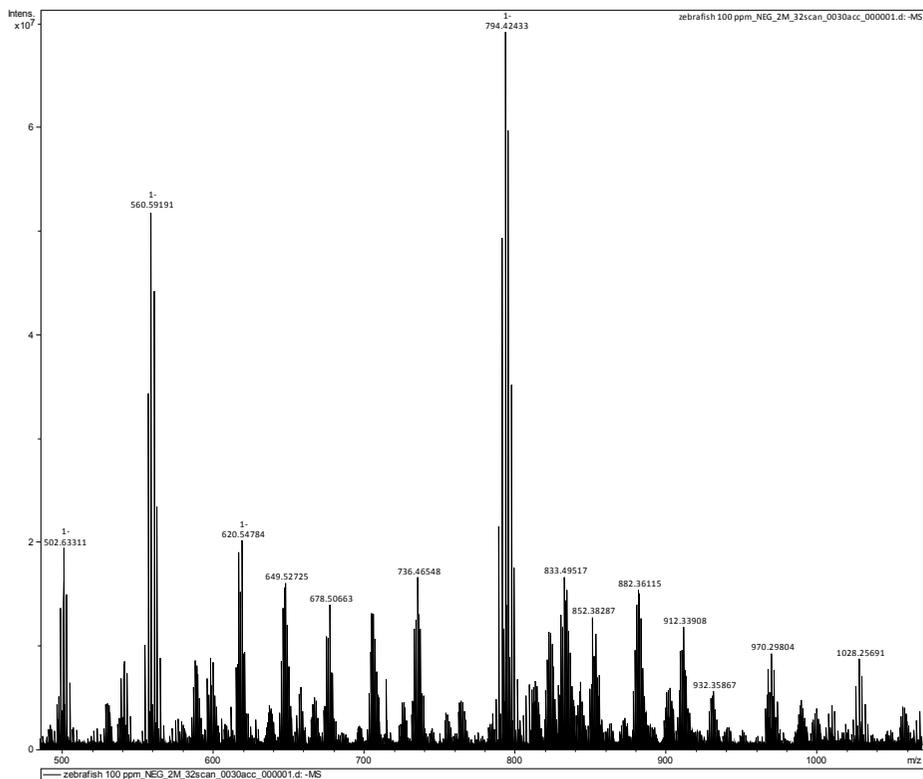
Supplementary Figure 4. Expansion of the negative electrospray ionization (ESI) Fourier-transform ion cyclotron resonance (FT-ICR) of the Ethyl acetate fraction.

ACKNOWLEDGMENT

REFERENCES

(1) Shehu, A. (2010) Conformational Search for the Protein Native State, in *Introduction to Protein Structure Prediction: Methods and Algorithms*, pp 431–452.

(2) Gruebele, M. (2002) Protein folding: the free energy surface. *Curr Opin Struct Biol 12*, 161–168.

(3) Wong, K. B., Yu, T. H., and Chan, C. H. (2012) Energetics of protein folding, in *Comprehensive Biophysics*, pp 19–33.

(4) Ohgushi, M., and Wada, A. (1983) "Molten-globule state": a compact form of globular proteins with mobile side-chains. *FEBS Lett 164*, 21–24.

(5) Privalov, P. L. (1996) Intermediate states in protein folding. *J Mol Biol*.

(6) Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding "speed limit." *Curr Opin Struct Biol*.

(7) Skach, W. R. (2009) Cellular mechanisms of membrane protein folding. *Nat Struct Mol Biol 16*, 606–612.

(8) Onuchic, J. N., Nymeyer, H., García, A. E., Chahine, J., and Socci, N. D. (2000) The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. *Adv Protein Chem*.

(9) Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem 48*, 545–600.

(10) Miller, M. A., and Wales, D. J. (1999) Energy landscape of a model protein. *J Chem Phys 111*, 6610–6616.

(11) Dill, K. A., and Chan, H. S. (1997) From levinthal to pathways to funnels. *Nat Struct Biol 4*, 10–19.

(12) Kapon, R., Nevo, R., and Reich, Z. (2008) Protein energy landscape roughness. *Biochem Soc Trans 36*, 1404–1408.

(13) Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science (80-)*.

(14) Batey, S., Nickson, A. A., and Clarke, J. (2008) Studying the folding of multidomain proteins. *HFSP J*.

(15) Motojima, F. (2015) How do chaperonins fold protein? *Biophysics (Oxf) 11*, 93–

102.

(16) Yang, A. S., and Honig, B. (1995) Free Energy Determinants of Secondary Structure Formation: I. A-Helices. *J Mol Biol 252*, 351–365.

(17) Yang, A. S., Hitz, B., and Honig, B. (1996) Free energy determinants of secondary structure formation: III. β-turns and their role in protein folding. *J Mol Biol 259*, 873–882.

(18) Rhee, Y. M., Sorin, E. J., Jayachandran, G., Lindahl, E., and Pande, V. S. (2004) Simulations of the role of water in the protein-folding mechanism. *Proc Natl Acad Sci 101*, 6456–6461.

(19) Tidor, B., and Karplus, M. (1993) The contribution of cross-links to protein stability: A normal mode analysis of the configurational entropy of the native state. *Proteins Struct Funct Bioinforma 15*, 71–79.

(20) Shandiz, A. T., Capraro, B. R., and Sosnick, T. R. (2007) Intramolecular cross-linking evaluated as a structural probe of the protein folding transition state. *Biochemistry 46*, 13711–13719.

(21) Anfinsen. (1973) Folding of Protein Chains. *Science (80-) 181*, 223–230.

(22) Bottomley, M. J., Batten, M. R., Lumb, R. A., and Bulleid, N. J. (2001) Quality control in the endoplasmic reticulum. *Curr Biol 11*, 1114–1118.

(23) Bu, G., Geuze, H. J., Strous, G. J., and Schwartz, A. L. (1995) 39 kDa receptor-associated protein is an ER resident protein and molecular chaperone for LDL receptor-related protein. *EMBO J 14*, 2269–80.

(24) Jessop, C. E., Chakravarthi, S., Watkins, R. H., and Bulleid, N. J. (2004) Oxidative protein folding in the mammalian endoplasmic reticulum. *Biochem Soc Trans 32*, 655–658.

(25) Neye, H. (2011) Peptidyl-prolyl isomerases, in *xPharm: The Comprehensive Pharmacology Reference*, pp 1–1.

(26) Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002) Proline cis-trans isomerization and protein folding. *Biochemistry*.

(27) Braakman, I., and Hebert, D. N. (2013) Protein folding in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol 5*, a013201.

(28) Mccaffrey, K., and Braakman, I. (2016) Protein quality control at the endoplasmic

reticulum. *Essays Biochem 60*, 227–235.

(29) Nishikawa, S. I., Brodsky, J. L., and Nakatsukasa, K. (2005) Roles of molecular chaperones in endoplasmic reticulum (ER) quality control and ER-associated degradation (ERAD). *J Biochem*.

(30) Wilkinson, B., and Gilbert, H. F. (2004) Protein disulfide isomerase. *Biochim Biophys Acta 1699*, 35–44.

(31) Olzmann, J. A., Kopito, R. R., and Christianson, J. C. (2015) The Mammalian Endoplasmic Reticulum-Associated Degradation System 1–16.

(32) Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884–890.

(33) Winklhofer, K. F., Tatzelt, J., and Haass, C. (2008) The two faces of protein misfolding: Gain- and loss-of-function in neurodegenerative diseases. *EMBO J*.

(34) The Cystic Fibrosis Genotype-Phenotype Consortium. (1993) Correlation between genotype and phenotype in patients with cystic fibrosis. *N Engl J Med 329*, 1308–1313.

(35) Bobadilla, J. L., Macek, M., Fine, J. P., and Farrell, P. M. (2002) Cystic fibrosis: A worldwide analysis of CFTR mutations - Correlation with incidence data and application to screening. *Hum Mutat*.

(36) Skach, W. R. (2000) Defects in processing and trafficking of the cystic fibrosis transmembrane conductance regulator, in *Kidney International*, pp 825–831.

(37) Bichet, D. G. (2008) Vasopressin Receptor Mutations in Nephrogenic Diabetes Insipidus. *Semin Nephrol 28*, 245–251.

(38) Westergard, L., Christensen, H. M., and Harris, D. A. (2007) The cellular prion protein (PrP(C)): its physiological function and role in disease. *Biochim Biophys Acta 1772*, 629–44.

(39) Cordeiro, Y., Kraineva, J., Gomes, M. P. B., Lopes, M. H., Martins, V. R., Lima, L. M. T. R., Foguel, D., Winter, R., and Silva, J. L. (2005) The amino-terminal PrP domain is crucial to modulate prion misfolding and aggregation. *Biophys J 89*, 2667–2676.

(40) Scott, M. R., Groth, D., Tatzelt, J., Torchia, M., Tremblay, P., DeArmond, S. J., and Prusiner, S. B. (1997) Propagation of prion strains through specific conformers of the prion protein. *J Virol 71*, 9032–44.

(41) Mastrianni, J. A. (2010) The genetics of prion diseases. *Genet Med*.

(42) Ji, H.-F., and Zhang, H.-Y. (2010) Beta-Sheet Constitution of Prion Proteins. *Trends Biochem Sci 35*, 129–34.

(43) Eisenberg, D., and Jucker, M. (2012) The amyloid state of proteins in human diseases. *Cell*.

(44) Berg, J., Tymoczko, J., and Stryer, L. (2002) The Immunoglobulin Fold Consists of a Beta-Sandwich Framework with Hypervariable Loops, in *Biochemistry*.

(45) Ding, F., Borreguero, J. M., Buldyrey, S. V, Stanley, H. E., and Dokholyan, N. V. (2003) Mechanism for the alpha-helix to beta-hairpin transition. *Proteins 53*, 220–228.

(46) Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein-misfolding diseases. *Group 426*, 905–909.

(47) Leidenheimer, N. J., and Ryder, K. G. (2014) Pharmacological chaperoning: A primer on mechanism and pharmacology. *Pharmacol Res 83*, 10–19.

(48) Muntau, A. C., Leandro, J., Staudigl, M., Mayer, F., and Gersting, S. W. (2014) Innovative strategies to treat protein misfolding in inborn errors of metabolism: Pharmacological chaperones and proteostasis regulators. *J Inherit Metab Dis 37*, 505–523.

(49) Morello, J. P., Salahpour, A., Laperrière, A., Bernier, V., Arthus, M. F., Lonergan, M., Petäjä-Repo, U., Angers, S., Morin, D., Bichet, D. G., and Bouvier, M. (2000) Pharmacological chaperones rescue cell-surface expression and function of misfolded V2 vasopressin receptor mutants. *J Clin Invest 105*, 887–895.

(50) Manglik, A., and Kobilka, B. (2014) The role of protein dynamics in GPCR function: Insights from the β2AR and rhodopsin. *Curr Opin Cell Biol 27*, 136–143.

(51) Nygaard, R., Zou, Y., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. J., Bokoch, M. P., Thian, F. S., Kobilka, T. S., Shaw, D. E., Mueller, L., Prosser, R. S., and Kobilka, B. K. (2013) The Dynamic Process of β2-Adrenergic Receptor Activation. *Cell 152*, 532–542.

(52) Rochdi, M. D., Vargas, G. a, Carpentier, E., Oligny-Longpré, G., Chen, S., Kovoor, A., Gitelman, S. E., Rosenthal, S. M., von Zastrow, M., and Bouvier, M. (2010) Functional characterization of vasopressin type 2 receptor substitutions (R137H/C/L) leading to nephrogenic diabetes insipidus and nephrogenic syndrome of inappropriate antidiuresis: implications for treatments. *Mol Pharmacol 77*, 836–845.

(53) Bernier, V., Lagacé, M., Lonergan, M., Arthus, M.-F., Bichet, D. G., and Bouvier, M. (2004) Functional rescue of the constitutively internalized V2 vasopressin receptor mutant R137H by the pharmacological chaperone action of SR49059. *Mol Endocrinol 18*, 2074–84.

(54) Bernier, V. (2005) Pharmacologic Chaperones as a Potential Treatment for X-Linked Nephrogenic Diabetes Insipidus. *J Am Soc Nephrol 17*, 232–243.

(55) Wang, X. H., Wang, H. M., Zhao, B. L., Yu, P., and Fan, Z. C. (2014) Rescue of defective MC4R cell-surface expression and signaling by a novel pharmacoperone Ipsen 17. *J Mol Endocrinol 53*, 17–29.

(56) Farooqi, I. S., and O'Rahilly, S. (2006) Genetics of obesity in humans. *Endocr Rev*.

(57) Janovick, J. A., Maya-Núñez, G., Ulloa-Aguirre, A., Huhtaniemi, I. T., Dias, J. A., Verbost, P., and Conn, P. M. (2009) Increased plasma membrane expression of human follicle-stimulating hormone receptor by a small molecule thienopyr(im)idine. *Mol Cell Endocrinol 298*.

(58) Layman, L. C., Cohen, D. P., Jin, M., Xie, J., Li, Z., Reindollar, R. H., Bolbolan, S., Bick, D. P., Sherins, R. R., Duck, L. W., Musgrove, L. C., Sellers, J. C., and Neill, J. D. (1998) Mutations in gonadotropin-releasing hormone receptor gene cause hypogonadotropic hypogonadism. *Nat Genet*.

(59) Janovick, J. A., Stewart, M. D., Jacob, D., Martin, L. D., Deng, J. M., Stewart, C. A., Wang, Y., Cornea, A., Chavali, L., Lopez, S., Mitalipov, S., Kang, E., Lee, H.-S., Manna, P. R., Stocco, D. M., Behringer, R. R., and Conn, P. M. (2013) Restoration of testis function in hypogonadotropic hypogonadal mice harboring a misfolded GnRHR mutant by pharmacoperone drug therapy. *Proc Natl Acad Sci 110*, 21030–21035.

(60) Noorwez, S. M., Kuksa, V., Imanishi, Y., Zhu, L., Filipek, S., Palczewski, K., and Kaushal, S. (2003) Pharmacological chaperone-mediated in vivo folding and stabilization of the P23H-opsin mutant associated with autosomal dominant retinitis pigmentosa. *J Biol Chem 278*, 14442–14450.

(61) Kobayashi, H., Ogawa, K., Yao, R., Lichtarge, O., and Bouvier, M. (2009) Functional rescue of ??1-adrenoceptor dimerization and trafficking by pharmacological chaperones. *Traffic 10*, 1019–1033.

(62) Newton, C. L., Whay, A. M., McArdle, C. A., Zhang, M., van Koppen, C. J., van de Lagemaat, R., Segaloff, D. L., and Millar, R. P. (2011) Rescue of expression and signaling of human luteinizing hormone G protein-coupled receptor mutants with an allosterically binding small-molecule agonist. *Proc Natl Acad Sci U S A 108*, 7172–6.

(63) Generoso, S. F., Giustiniano, M., La Regina, G., Bottone, S., Passacantilli, S., Di Maro, S., Cassese, H., Bruno, A., Mallardo, M., Dentice, M., Silvestri, R., Marinelli, L., Sarnataro, D., Bonatti, S., Novellino, E., and Stornaiuolo, M. (2015) Pharmacological folding chaperones act as allosteric ligands of Frizzled4. *Nat Chem Biol 11*, 280–286.

(64) Robitaille, J., MacDonald, M. L. E., Kaykas, A., Sheldahl, L. C., Zeisler, J., Dubé, M. P., Zhang, L. H., Singaraja, R. R., Guernsey, D. L., Zheng, B., Siebert, L. F., Hoskin-Mott, A., Trese, M. T., Pimstone, S. N., Shastry, B. S., Moon, R. T., Hayden, M. R., Paul Goldberg, Y., and Samuels, M. E. (2002) Mutant frizzled-4 disrupts retinal angiogenesis in familial exudative vitreoretinopathy. *Nat Genet 32*, 326–330.

(65) Lemma, V., D'Agostino, M., Caporaso, M. G., Mallardo, M., Oliviero, G., Stornaiuolo, M., and Bonatti, S. (2013) A disorder-to-order structural transition in the COOH-tail of Fz4 determines misfolding of the L501fsX533-Fz4 mutant. *Sci Rep 3*, 1–10.

(66) Riccio, G., Bottone, S., La Regina, G., Badolati, N., Passacantilli, S., Rossi, G. B., Accardo, A., Dentice, M., Silvestri, R., Novellino, E., and Stornaiuolo, M. (2018) A Negative Allosteric Modulator of WNT Receptor Frizzled 4 Switches into an Allosteric Agonist. *Biochemistry 57*, 839–851.

(67) D'Agostino, M., Lemma, V., Chesi, G., Stornaiuolo, M., Cannata Serio, M., D'Ambrosio, C., Scaloni, A., Polishchuk, R., and Bonatti, S. (2013) The cytosolic chaperone -crystallin B rescues folding and compartmentalization of misfolded multispan transmembrane proteins. *J Cell Sci 126*, 4160–4172.

(68) Akiyama, T. (2000) Wnt/β-catenin signaling. *Cytokine Growth Factor Rev*.

(69) Ke, J., Harikumar, K. G., Erice, C., Chen, C., Gu, X., Wang, L., Parker, N., Cheng, Z., Xu, W., Williams, B. O., Melcher, K., Miller, L. J., and Eric Xu, H. (2013) Structure and function of Norrin in assembly and activation of a Frizzled 4-Lrp5/6 complex. *Genes Dev 27*, 2305–2319.

(70) Osakada, F., Jin, Z.-B., Hirami, Y., Ikeda, H., Danjyo, T., Watanabe, K., Sasai, Y., and Takahashi, M. (2009) In vitro differentiation of retinal cells from human pluripotent stem cells by small-molecule induction. *J Cell Sci 122*, 3169–3179.

(71) Cruciat, C. M. (2014) Casein kinase 1 and Wnt/β-catenin signaling. *Curr Opin Cell Biol*.

(72) Dijksterhuis, J. P., Petersen, J., and Schulte, G. (2014) WNT/Frizzled signalling: Receptor-ligand selectivity with focus on FZD-G protein signalling and its physiological relevance: IUPHAR Review 3. *Br J Pharmacol*.

(73) Von Maltzahn, J., Bentzinger, C. F., and Rudnicki, M. A. (2012) Wnt7a-Fzd7 signalling directly activates the Akt/mTOR anabolic growth pathway in skeletal muscle. *Nat Cell Biol 14*, 186–191.

(74) Kim, S. E., Lee, W. J., and Choi, K. Y. (2007) The PI3 kinase-Akt pathway mediates Wnt3a-induced proliferation. *Cell Signal 19*, 511–518.

(75) Li, J., Sutter, C., Parker, D. S., Blauwkamp, T., Fang, M., and Cadigan, K. M. (2007) CBP/p300 are bimodal regulators of Wnt signaling. *EMBO J 26*, 2284–2294.

(76) Thorpe, L. M., Yuzugullu, H., and Zhao, J. J. (2015) PI3K in cancer: Divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer*.

(77) Gallenkamp, D., Gelato, K. A., Haendler, B., and Weinmann, H. (2014) Bromodomains and their pharmacological inhibitors. *ChemMedChem*.

(78) MacDonald, B. T., and He, X. (2012) Frizzled and LRp5/6 receptors for wnt/β-catenin signaling. *Cold Spring Harb Perspect Biol 4*.

(79) Jin, X., Jeon, H. Y., Joo, K. M., Kim, J. K., Jin, J., Kim, S. H., Kang, B. G., Beck, S., Lee, S. J., Kim, J. K., Park, A. K., Park, W. Y., Choi, Y. J., Nam, D. H., and Kim, H. (2011) Frizzled 4 regulates stemness and invasiveness of migrating glioma cells established by serial intracranial transplantation. *Cancer Res 71*, 3066–3075.

(80) Heathcote, J. (2011) Page 1 of 56 Hepatology. *Hepatology* 1–56.

(81) Nambi, P., and Aiyar, N. (2003) G protein-coupled receptors in drug discovery. *Assay Drug Dev Technol 1*, 305–10.

(82) Tang, X., Wang, Y., Li, D., Luo, J., and Liu, M. (2012) Orphan G protein-coupled receptors (GPCRs): biological functions and potential drug targets. *Acta Pharmacol Sin 33*, 363–371.

(83) Civelli, O., Reinscheid, R. K., Zhang, Y., Wang, Z., Fredriksson, R., and Schiöth, H. B. (2013) G protein-coupled receptor deorphanizations. *Annu Rev Pharmacol Toxicol*.

(84) Fève, M., Saliou, J. M., Zeniou, M., Lennon, S., Carapito, C., Dong, J., Van Dorsselaer, A., Junier, M. P., Chneiweiss, H., Cianférani, S., Haiech, J., and Kilhoffer, M. C. (2014) Comparative expression study of the endo-G protein coupled receptor (GPCR) repertoire in human glioblastoma cancer stem-like cells, U87-MG cells and non malignant cells of neural origin unveils new potential therapeutic targets. *PLoS One 9*.

(85) Tamimi, A. F., and Juweid, M. (2018) Epidemiology and Outcome of Glioblastoma Incidence of Glioblastoma. *NCBI Bookshelf* 1–10.

(86) Singh, Sheila K.; Clarke, Ian D.; Terasaki, Mizuhiko; Bonn, Victoria E.; Hawkins, Cynthia; Squire, Jeremy; Dirka, P. B. (2003) Identification of a cancer stem cell in human brain tumors. *Cancer Res 63*, 5821–5828.

(87) Stupp, R., Mason, W., Bent, M. van den, Weller, M., Fisher, B., Taphoorn, M., and Jo. (2005) EORTC 22981: Radiotherapy plus concomitant and adjuvant temozolomide for globlastoma. *N Engl J Med 352*, 987–96.

(88) Hsieh, J. C., and Lesniak, M. S. (2005) Surgical management of high-grade gliomas. *Expert Rev Neurother*.

(89) Cherry, A. E., and Stella, N. (2014) G protein-coupled receptors as oncogenic signals in glioma: Emerging therapeutic avenues. *Neuroscience 278*, 222–236.

(90) Wang, X. J., Zhang, D. L., Xu, Z. G., Ma, M. L., Wang, W. B., Li, L. L., Han, X. L., Huo, Y., Yu, X., and Sun, J. P. (2015) Understanding cadherin EGF LAG seven-pass G-type receptors. *J Neurochem 131*, 699–711.

(91) Takeichi, M. (2007) The cadherin superfamily in neuronal connections and interactions. *Nat Rev Neurosci 8*, 11–20.

(92) Shima, Y., Kengaku, M., Hirano, T., Takeichi, M., and Uemura, T. (2004) Regulation of dendritic maintenance and growth by a mammalian 7-pass transmembrane cadherin. *Dev Cell 7*, 205–216.

(93) Shima, Y., Kawaguchi, S. Y., Kosaka, K., Nakayama, M., Hoshino, M., Nabeshima, Y., Hirano, T., and Uemura, T. (2007) Opposing roles in neurite growth control by two seven-pass transmembrane cadherins. *Nat Neurosci 10*, 963–969.

(94) Hohenester, E., Tisi, D., Talts, J. F., and Timpl, R. (1999) The crystal structure of a laminin G-like module reveals the molecular basis of α-dystroglycan binding to laminins, perlecan, and agrin. *Mol Cell 4*, 783–792.

(95) Araç, D., Boucard, A. a, Bolliger, M. F., Nguyen, J., Soltis, S. M., Südhof, T. C., and Brunger, A. T. (2012) A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *EMBO J 31*, 1364–1378.

(96) Runge, S., Thøgersen, H., Madsen, K., Lau, J., and Rudolph, R. (2008) Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. *J Biol Chem 283*, 11340–11347.

(97) Tesmer, J. J. G. (2012) A GAIN in understanding autoproteolytic G protein-coupled receptors and polycystic kidney disease proteins. *EMBO J 31*, 1334–1335.

(98) Lin, H. H., Chang, G. W., Davies, J. Q., Stacey, M., Harris, J., and Gordon, S. (2004) Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-coupled receptor proteolytic site motif. *J Biol Chem 279*, 31823–31832.

(99) Krasnoperov, V., Lu, Y., Buryanovsky, L., Neubert, T. A., Ichtehenko, K., and Petrenko, A. G. (2002) Post-translational proteolytic processing of the calcium-independent receptor of α-latrotoxin (CIRL), a natural chimera of the cell adhesion protein and the G protein-coupled receptor: Role of the G protein-coupled receptor proteolysis site (GPS) motif. *J Biol Chem 277*, 46518–46526.

(100) Prömel, S., Frickenhaus, M., Hughes, S., Mestek, L., Staunton, D., Woollard, A., Vakonakis, I., Schöneberg, T., Schnabel, R., Russ, A. P., and Langenhan, T. (2012) The GPS Motif Is a Molecular Switch for Bimodal Activities of Adhesion Class G Protein-Coupled Receptors. *Cell Rep 2*, 321–331.

(101) Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., and Araki, T. (2018) American Association for the Advancement of Science *301*, 1515–1519.

(102) Venkatakrishnan, A. J., Deupi, X., Lebon, G., Tate, C. G., Schertler, G. F., and Madan Babu, M. (2013) Molecular signatures of G-protein-coupled receptors. *Nature 494*, 185–194.

(103) Gross, J., Grimm, O., Ortega, G., Teuber, I., Lesch, K. P., and Meyer, J. (2001) Mutational analysis of the neuronal cadherin gene CELSR1 and exclusion as a candidate for catatonic schizophrenia in a large family. *Psychiatr Genet 11*, 197–200.

(104) Tissir, F., and Goffinet, A. M. (2013) Atypical cadherins Celsr1-3 and planar cell polarity in vertebrates. *Prog Mol Biol Transl Sci* 1st ed. Elsevier Inc.

(105) Meyer, R. D., Srinivasan, S., Singh, A. J., Mahoney, J. E., Gharahassanlou, K. R., and Rahimi, N. (2011) PEST Motif Serine and Tyrosine Phosphorylation Controls Vascular Endothelial Growth Factor Receptor 2 Stability and Downregulation. *Mol Cell Biol 31*, 2010–2025.

(106) Usui, T., Shima, Y., Shimada, Y., Hirano, S., Burgess, R. W., Schwarz, T. L., Takeichi, M., and Uemura, T. (1999) Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell 98*, 585–595.

(107) Shima, Y., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Chisaka, O., Takeichi, M., and Uemura, T. (2002) Differential expression of the seven-pass transmembrane

cadherin genes Celsr1-3 and distribution of the Celsr2 protein during mouse development. *Dev Dyn 223*, 321–332.

(108) Formstone, C. J., and Little, P. F. R. (2001) The flamingo-related mouse Celsr family (Celsr1-3) genes exhibit distinct patterns of expression during embryonic development. *Mech Dev 109*, 91–94.

(109) Tissir, F., De-Backer, O., Goffinet, A. M., and Lambert de Rouvroit, C. (2002) Developmental expression profiles of Celsr (Flamingo) genes in the mouse. *Mech Dev 112*, 157–160.

(110) Tissir, F., Qu, Y., Montcouquiol, M., Zhou, L., Komatsu, K., Shi, D., Fujimori, T., Labeau, J., Tyteca, D., Courtoy, P., Poumay, Y., Uemura, T., and Goffinet, A. M. (2010) Lack of cadherins Celsr2 and Celsr3 impairs ependymal ciliogenesis, leading to fatal hydrocephalus. *Nat Neurosci 13*, 700–707.

(111) Vilboux, T., Doherty, D. A., Glass, I. A., Parisi, M. A., Phelps, I. G., Cullinane, A. R., Zein, W., Brooks, B. P., Heller, T., Soldatos, A., Oden, N. L., Yildirimli, D., Vemulapalli, M., Mullikin, J. C., Malicdan, M. C. V., Gahl, W. A., and Gunay-Aygun, M. (2017) Molecular genetic findings and clinical correlations in 100 patients with Joubert syndrome and related disorders prospectively evaluated at a single center. *Genet Med 19*, 875–882.

(112) Park, T. J., Haigo, S. L., and Wallingford, J. B. (2006) Ciliogenesis defects in embryos lacking inturned or fuzzy function are associated with failure of planar cell polarity and Hedgehog signaling. *Nat Genet 38*, 303–311.

(113) Vilboux, T., Malicdan, M. C. V., Roney, J. C., Cullinane, A. R., Stephen, J., Yildirimli, D., Bryant, J., Fischer, R., Vemulapalli, M., Mullikin, J. C., Steinbach, P. J., Gahl, W. A., and Gunay-Aygun, M. (2017) CELSR2, encoding a planar cell polarity protein, is a putative gene in Joubert syndrome with cortical heterotopia, microophthalmia, and growth hormone deficiency. *Am J Med Genet Part A 173*, 661–666.

(114) Wada, H., Tanaka, H., Nakayama, S., Iwasaki, M., and Okamoto, H. (2006) Frizzled3a and Celsr2 function in the neuroepithelium to regulate migration of facial motor neurons in the developing zebrafish hindbrain. *Development 133*, 4749–4759.

(115) Qu, Y., Glasco, D. M., Zhou, L., Sawant, A., Ravni, A., Fritzsch, B., Damrau, C., Murdoch, J. N., Evans, S., Pfaff, S. L., Formstone, C., Goffinet, A. M., Chandrasekhar, A., and Tissir, F. (2010) Atypical Cadherins Celsr1-3 Differentially Regulate Migration of Facial Branchiomotor Neurons in Mice. *J Neurosci 30*, 9392–9401.

(116) Carreira-Barbosa, F., Kajita, M., Morel, V., Wada, H., Okamoto, H., Arias, A. M., Fujita, Y., Wilson, S. W., and Tada, M. (2009) Flamingo regulates epiboly and convergence/extension movements through cell cohesive and signalling functions during zebrafish gastrulation. *Development 136*, 877–877.

(117) Vivancos, V., Chen, P., Spassky, N., Qian, D., Dabdoub, A., Kelley, M., Studer, M., and Guthrie, S. (2009) Wnt activity guides facial branchiomotor neuron migration, and involves the PCP pathway and JNK and ROCK kinases. *Neural Dev 4*, 1–16.

(118) Gao, F. B., Kohwi, M., Brenman, J. E., Jan, L. Y., and Jan, Y. N. (2000) Control of dendritic field formation in Drosophila: The roles of flamingo and competition between homologous neurons. *Neuron 28*, 91–101.

(119) Kimura, H. (2006) Potential dual molecular interaction of the Drosophila 7-pass transmembrane cadherin Flamingo in dendritic morphogenesis. *J Cell Sci 119*, 1118–1129.

(120) Boutin, C., Goffinet, A. M., and Tissir, F. (2012) Celsr1-3 Cadherins in PCP and Brain Development. *Curr Top Dev Biol 101*, 161–183.

(121) Geetha-Loganathan, P., Nimmagadda, S., Fu, K., and Richman, J. M. (2014) Avian facial morphogenesis is regulated by c-Jun N-terminal Kinase/Planar Cell Polarity (JNK/PCP) wingless-related (WNT) signaling. *J Biol Chem 289*, 24153–24167.

(122) Cortijo, C., Gouzi, M., Tissir, F., and Grapin-Botton, A. (2012) Planar Cell Polarity Controls Pancreatic Beta Cell Differentiation and Glucose Homeostasis. *Cell Rep 2*, 1593–1606.

(123) Hakeda-Suzuki, S., Berger-Müller, S., Tomasi, T., Usui, T., Horiuchi, S. Y., Uemura, T., and Suzuki, T. (2011) Golden goal collaborates with flamingo in conferring synaptic-layer specificity in the visual system. *Nat Neurosci 14*, 314–323.

(124) Harty, B. L., Krishnan, A., Sanchez, N. E., Schiöth, H. B., and Monk, K. R. (2015) Defining the gene repertoire and spatiotemporal expression profiles of adhesion G protein-coupled receptors in zebrafish. *BMC Genomics 16*.

(125) Tränkle, C., Weyand, O., Schröter, a, and Mohr, K. (1999) Using a radioalloster to test predictions of the cooperativity model for gallamine binding to the allosteric site of muscarinic acetylcholine M(2) receptors. *Mol Pharmacol 56*, 962–965.

(126) Haasen, D., Schnapp, A., Valler, M. J., and Heilker, R. (2006) G Protein-Coupled Receptor Internalization Assays in the High-Content Screening Format. *Methods Enzymol 414*, 121–139.

(127) Hirasawa, A., Tsumaya, K., Awaji, T., Katsuma, S., Adachi, T., Yamada, M., Sugimoto, Y., Miyazaki, S., and Tsujimoto, G. (2005) Free fatty acids regulate gut incretin glucagon-like peptide-1 secretion through GPR120. *Nat Med 11*, 90–94.

(128) Offermanns, S., and Simon, M. I. (1995) G 15 and G 16 Couple a Wide Variety of Receptors to Phospholipase C. *J Biol Chem 270*, 15175–15180.

(129) Sato, S., Huang, X.-P., Kroeze, W. K., and Roth, B. L. (2016) Discovery and Characterization of Novel GPR39 Agonists Allosterically Modulated by Zinc. *Mol Pharmacol 90*, 726–737.

(130) Bates, B., Zhang, L., Nawoschik, S., Kodangattil, S., Tseng, E., Kopsco, D., Kramer, A., Shan, Q., Taylor, N., Johnson, J., Sun, Y., Chen, H. M., Blatcher, M., Paulsen, J. E., and Pausch, M. H. (2006) Characterization of Gpr101 expression and G-protein coupling selectivity. *Brain Res 1087*, 1–14.

(131) Okamoto, H., Takuwa, N., Yatomi, Y., Gonda, K., Shigematsu, H., and Takuwa, Y. (1999) EDG3 is a functional receptor specific for sphingosine 1-phosphate and sphingosylphosphorylcholine with signaling characteristics distinct from EDG1 and AGR16. *Biochem Biophys Res Commun 260*, 203–208.

(132) Huang, X. P., Karpiak, J., Kroeze, W. K., Zhu, H., Chen, X., Moy, S. S., Saddoris, K. A., Nikolova, V. D., Farrell, M. S., Wang, S., Mangano, T. J., Deshpande, D. A., Jiang, A., Penn, R. B., Jin, J., Koller, B. H., Kenakin, T., Shoichet, B. K., and Roth, B. L. (2015) Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. *Nature 527*, 477–483.

(133) Libert, F., Schiffmann, S. N., Lefort, A., Parmentier, M., Gérard, C., Dumont, J. E., Vanderhaeghen, J. J., and Vassart, G. (1991) The orphan receptor cDNA RDC7 encodes an A1 adenosine receptor. *EMBO J 10*, 1677–1682.

(134) Civelli, O., Reinscheid, R. K., and Nothacker, H. P. (1999) Orphan receptors, novel neuropeptides and reverse pharmaceutical research. *Brain Res 848*, 63–65.

(135) May, L. T., and Christopoulos, A. (2003) Allosteric modulators of G-protein-coupled receptors. *Curr Opin Pharmacol*.

(136) Guixà-González, R., Albasanz, J. L., Rodriguez-Espigares, I., Pastor, M., Sanz, F., Martí-Solano, M., Manna, M., Martinez-Seara, H., Hildebrand, P. W., Martín, M., and Selent, J. (2017) Membrane cholesterol access into a G-protein-coupled receptor. *Nat Commun 8*.

(137) Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V. P., Chien, E. Y. T., Velasquez, J., Kuhn, P., and Stevens, R. C. (2008) A Specific Cholesterol

Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor. *Structure 16*, 897–905.

(138) Potter, R. M., Harikumar, K. G., Wu, S. V., and Miller, L. J. (2012) Differential sensitivity of types 1 and 2 cholecystokinin receptors to membrane cholesterol. *J Lipid Res 53*, 137–148.

(139) Vance, J. E. (2012) Dysregulation of cholesterol balance in the brain: contribution to neurodegenerative diseases. *Dis Model Mech 5*, 746–755.

(140) Goritz, C., Mauch, D. H., and Pfrieger, F. W. (2005) Multiple mechanisms mediate cholesterol-induced synaptogenesis in a CNS neuron. *Mol Cell Neurosci 29*, 190–201.

(141) Rhinn, M., and Dolle, P. (2012) Retinoic acid signalling during development. *Development 139*, 843–858.

(142) Romani, S., Illi, B., De Mori, R., Savino, M., Gleeson, J. G., and Valente, E. M. (2014) The ciliary proteins Meckelin and Jouberin are required for retinoic acid-dependent neural differentiation of mouse embryonic stem cells. *Differentiation 87*, 134–146.

(143) Speirs, C. K., Jernigan, K. K., Kim, S.-H., Cha, Y. I., Lin, F., Sepich, D. S., DuBois, R. N., Lee, E., and Solnica-Krezel, L. (2010) Prostaglandin G signaling stimulates gastrulation movements by limiting cell adhesion through Snai1a stabilization. *Development 137*, 1327–1337.

(144) Van Der Westhuizen, E. T., Valant, C., Sexton, P. M., and Christopoulos, A. (2015) Minireview Endogenous Allosteric Modulators of G Protein–Coupled Receptors. *J Pharmacol Exp Ther J Pharmacol Exp Ther 353*, 246–260.