Cellular and Molecular Life Sciences (2019) 76:4461–4492
https://doi.org/10.1007/s00018-019-03276-1

REVIEW

Multi-functionality of proteins involved in GPCR and G protein signaling: making sense of structure–function continuum with intrinsic disorder-based proteoforms

Alexander V. Fonin[1] · April L. Darling[2] · Irina M. Kuznetsova[1] · Konstantin K. Turoverov[1,3] · Vladimir N. Uversky[2,4]

Received: 5 August 2019 / Revised: 5 August 2019 / Accepted: 12 August 2019 / Published online: 19 August 2019

Abstract

GPCR–G protein signaling system recognizes a multitude of extracellular ligands and triggers a variety of intracellular signaling cascades in response. In humans, this system includes more than 800 various GPCRs and a large set of heterotrimeric G proteins. Complexity of this system goes far beyond a multitude of pair-wise ligand–GPCR and GPCR–G protein interactions. In fact, one GPCR can recognize more than one extracellular signal and interact with more than one G protein. Furthermore, one ligand can activate more than one GPCR, and multiple GPCRs can couple to the same G protein. This defines an intricate multifunctionality of this important signaling system. Here, we show that the multifunctionality of GPCR–G protein system represents an illustrative example of the protein structure–function continuum, where structures of the involved proteins represent a complex mosaic of differently folded regions (foldons, non-foldons, unfoldons, semi-foldons, and inducible foldons). The functionality of resulting highly dynamic conformational ensembles is fine-tuned by various post-translational modifications and alternative splicing, and such ensembles can undergo dramatic changes at interaction with their specific partners. In other words, GPCRs and G proteins exist as sets of conformational/basic, inducible/modified, and functioning proteoforms characterized by a broad spectrum of structural features and possessing various functional potentials.

Keywords G proteins · G protein-coupled receptors · Intrinsically disordered protein · Proteoform · Protein–protein interaction · Post-translational modification · Alternative splicing

Introduction

G protein-coupled receptors (GPCRs) represent one of the largest families of protein receptors, which, in humans, includes over 800 members [1–4]. GPCRs can interact with (and be activated by) more than a 1000 natural and artificial extracellular ligands, ranging from photons to amines, lipids, nucleotides, organic odorants, peptides, and proteins [4, 5]. All these signals are used to initiate a variety of intracellular signaling cascades via interaction of an activated GPCR with one of four major Gα families (guanine nucleotide-binding proteins) encoded by 16 human genes [3, 5, 6] leading to the modulation of various downstream effector proteins (such as adenylate cyclase and phospholipase C) and key secondary messengers (e.g., cAMP, Ca^{2+}, and IP3) [5, 7, 8]. Interaction of activated GPCRs with Gα proteins is characterized by complex coupling selectivity, where several different GPCRs can couple to the same Gα protein and one GPCR can couple to more than one Gα protein. As a result, GPCRs mediate most cellular responses to hormones,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00018-019-03276-1) contains supplementary material, which is available to authorized users.

Alexander V. Fonin and April L. Darling contributed equally to this work.

✉ Vladimir N. Uversky
vuversky@health.usf.edu

[1] Laboratory of structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, St. Petersburg 194064, Russian Federation

[2] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, USA

[3] Department of Biophysics, Peter the Great St. Petersburg Polytechnic University, Polytechnicheskaya av. 29, St. Petersburg 195251, Russian Federation

[4] Institute for Biological Instrumentation, Russian Academy of Sciences, Pushchino, Moscow, Russian Federation

neurotransmitters, ions, photons, and other environmental stimuli, and are responsible for vision, olfaction, and taste. Therefore, these receptors are considered as a cellular "control panel" that is "able to detect the presence of a strikingly diverse array of molecules outside the cell and to initiate a variety of intracellular signaling cascades in response" [9]. Not surprisingly, GPCRs represent the largest class of drug targets, with more than half of all modern drugs being targeted at these proteins [10]. In addition to the complexity of coupling between the activated GPCRs and G proteins, it is likely that the presence of intrinsic disorder and associated with it high conformational flexibility plays an important role in regulation and controlling this GPCR–G protein signaling system and also contributes to the multifunctionality and binding promiscuity of proteins involved in the GPCR and G protein signaling.

As a matter of fact, it is well accepted now that the representation of interactions between a protein and its binding partners (substrates, ligands, nucleic acids, peptides or other proteins) in a form of the classical lock-and-key model [11] that dominated in molecular biology for more than a century and formed the foundation of modern protein science [12, 13] is an obvious oversimplification. Many biologically active proteins and protein regions, instead of following the protein structure–function paradigm, where protein functionality is directly related to its unique 3D-structure and which represented the major rule of the protein science and structural biology, have been shown to lack unique 3D structure in their native states under physiological conditions. Instead, the functional states of such intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) [14–20] represent highly dynamic conformational ensembles containing large number of very different and rapidly interconverting conformations [14, 16, 17, 21–25].

Several computational studies revealed that intrinsic disorder is not a fluke, with IDPs/IDPRs being commonly found in all proteomes [18, 26–30], especially in eukaryotes [14]. While lacking stable 3D structures, IDPs/IDPRs play a number of crucial functional roles in living organisms [12, 13, 17, 31–34]. These proteins possess multiple unique features that allow them to be engaged in specific types of regulation and define the ability of IDPs/IDPRs to play crucial roles in vital biological processes, such as control, signaling, recognition, and regulation [14, 17, 22, 31, 32, 35, 36]. Functions of IDPs/IDPRs are believed to complement biological activities of ordered proteins and domains [37]. Since IDPs/IDPRs are capable of acting as promiscuous binders [31, 32, 38–40], high levels of intrinsic disorder represent an important feature of hub proteins found within the protein–protein interaction networks [41]. This high binding promiscuity combined with the ability of IDPs/IDPRs to be engaged in highly specific interactions with low affinity also explains why intrinsic disorder is frequently found in signaling proteins [42]. Many IDPs and IDPRs were shown to be related to the pathogenesis of various human diseases, such as amyloidoses, cancer, cardiovascular diseases, diabetes, and various neurodegenerative diseases [43, 44].

Phenomenon of intrinsic disorder defines exceptional spatiotemporal heterogeneity of proteins and regions. In fact, intrinsic disorder can be present in a whole protein or in protein regions, being able to penetrate into the different levels of protein organization (proteins or protein regions can be disordered at the tertiary structure or the secondary structure level). As a result, IDPs/IDPRs can exist in the extended or collapsed disordered forms, where extended IDPs/IDPRs behave as coil-like or pre-molten globule-like conformations characterized by the high level of solvent exposure, and where collapsed IDPs/IDPRs possess properties of native molten globule being characterized by a restricted range of motions [12, 13, 16, 24, 33, 45]. Since different protein regions can be disordered to a different degree, proteins are characterized by the mosaic structure containing a multitude of potentially foldable, partially foldable, differently foldable or not foldable at all segments playing different roles in protein functionality [24, 46], and even containing ordered regions that need to undergo order-to-disorder transition to make protein active [24, 47]. Furthermore, some IDPRs were shown to bind to multiple partners gaining very different structures in their bound states [40]. Therefore, IDPs/IDPRs can be described as a modular assembly of foldons (i.e., independently foldable regions), inducible foldons (IDPRs that undergo (partial) folding at interaction with specific partners), semi-foldons (IDPRs that are always semi-folded), non-foldons (IDPRs that never fold), and unfoldons (a part of the ordered protein that has to undergo order-to-disorder transition to make protein active), and this mosaic organization of these proteins defines extreme spatiotemporal structural heterogeneity related to their multifunctionality [24]. In fact, inclusion of intrinsic disorder into consideration of protein functionality resulted in the conversion of the "one sequence-one structure-one function" model (which represents a corollary of the famous lock-and-key model [11]) into the more realistic protein structure–function continuum representation of the correlation between protein structure and function, where a given protein exists as a dynamic conformational ensemble characterized by the diverse structural features and miscellaneous functions [48].

Finally, intrinsic disorder-based structure–function continuum, together with several other mechanisms (see below), provides important means for the dramatic expansion of the functional proteome relative to the encoding it genome, an important phenomenon behind the observation that the complexity of biological systems is mostly determined by their proteome sizes and not by the sizes of their genomes [49]. This can be illustrated by gene–protein relationship in *Homo*

sapiens [50–54], where the number of protein-coding genes is approaching 20,700 [55], but the number of functionally different proteins is in a range between a few million [56] and several billion [57]. The functional diversification of proteinaceous products of a gene is achieved by allelic variations (i.e., single or multiple point mutations (amino acid polymorphisms), indels, single nucleotide polymorphisms (SNPs)), alternative splicing (AS), mRNA editing and other pre-translational mechanisms affecting mRNA, as well as by a wide spectrum of various post-translational modifications (PTMs) of a polypeptide chain. As a result, a single gene encodes a set of distinct protein molecules, known as proteoforms [57]. Obviously, all these means lead to some changes in the physico-chemical structure of a polypeptide chain, and, therefore, the resulting proteoforms have induced or modified nature. Furthermore, intrinsic disorder and functionality can further increase protein structural diversity, giving rise to conformational or basic proteoforms and functioning proteoforms, respectively [48]. In fact, since many PTM sites are known to be preferentially found within the IDPRs [58, 59], since mRNA regions affected by AS predominantly encodes IDPRs [46], since IDPs/IDPRs are known to be highly promiscuous binders [14, 16, 17, 22–25, 31, 32, 35, 36, 38, 39, 45, 60–66], and since IDPs/IDPRs are characterized by the exceptional spatiotemporal heterogeneity (see above), proteins and protein regions without unique structures represent a very rich source of proteoforms [48].

Therefore, instead of the influential, but oversimplified "one gene—one enzyme" model, according to which each gene is responsible for producing a single enzyme that in turn affects a single step in a metabolic pathway [67], the actual gene–protein relationship is better described by the "one-gene—many-proteins" or "one-gene—many-functions" models. In other words, "one gene—many proteins—many functions" paradigm and "protein structure–function continuum" model provide a global description of a link between protein structure and function, where a given protein exists as a dynamic conformational ensemble containing multiple proteoforms (conformational/basic, inducible/modified, and functioning) characterized by a broad spectrum of structural features and possessing various functional potentials. Figure 1 illustrates this idea by showing that the differently folded pieces that can be found in a protein molecule (i.e., all these foldons, inducible foldons, semi-foldons, non-foldons, and unfoldons) might have well-defined and specific functions, indicating that protein multifunctionality is naturally encoded in this heterogeneous structural mosaic [68, 69].

Here, we utilize examples of human GPCRs and G proteins to illustrate the versatility of intrinsic disorder and proteoform concepts for the description of protein multifunctionality. In line with an outstanding previous study of Venkatakrishnan et al. [70], we are showing here that extracellular and cytoplasmic regions of GPCRs invariantly contain numerous IDPRs, multiple disorder-based binding sites, abundant post-translational modification sites, and typically have multiple isoforms generated by alternative splicing, providing strong support to the idea that intrinsic disorder plays a crucial role in functionality of GPCRs. Similarly, we provide compelling evidence that all human G proteins contain noticeable levels of functional intrinsic disorder, with IDPRs possessing numerous sites of various post-translational modifications and often including disorder-based interactions sites. Many G proteins also contain multiple alternative splicing-generated isoforms. Furthermore, both GPCRs and G proteins often undergo functional conformational changes that range from domain motion to binding-induced disorder-to-order transitions. In other words, multifunctionality of these major players of the GPCR–G protein system is determined by the fact that all these proteins exist as numerous highly dynamic conformational/basic, inducible/modified, and functioning proteoforms.

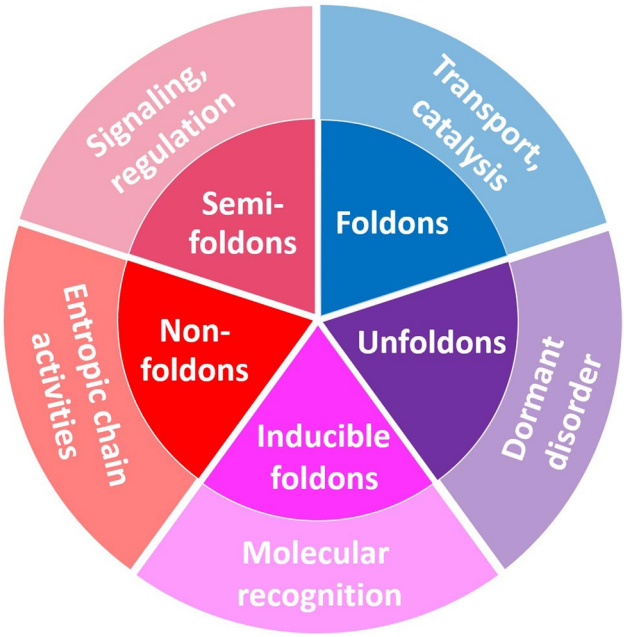
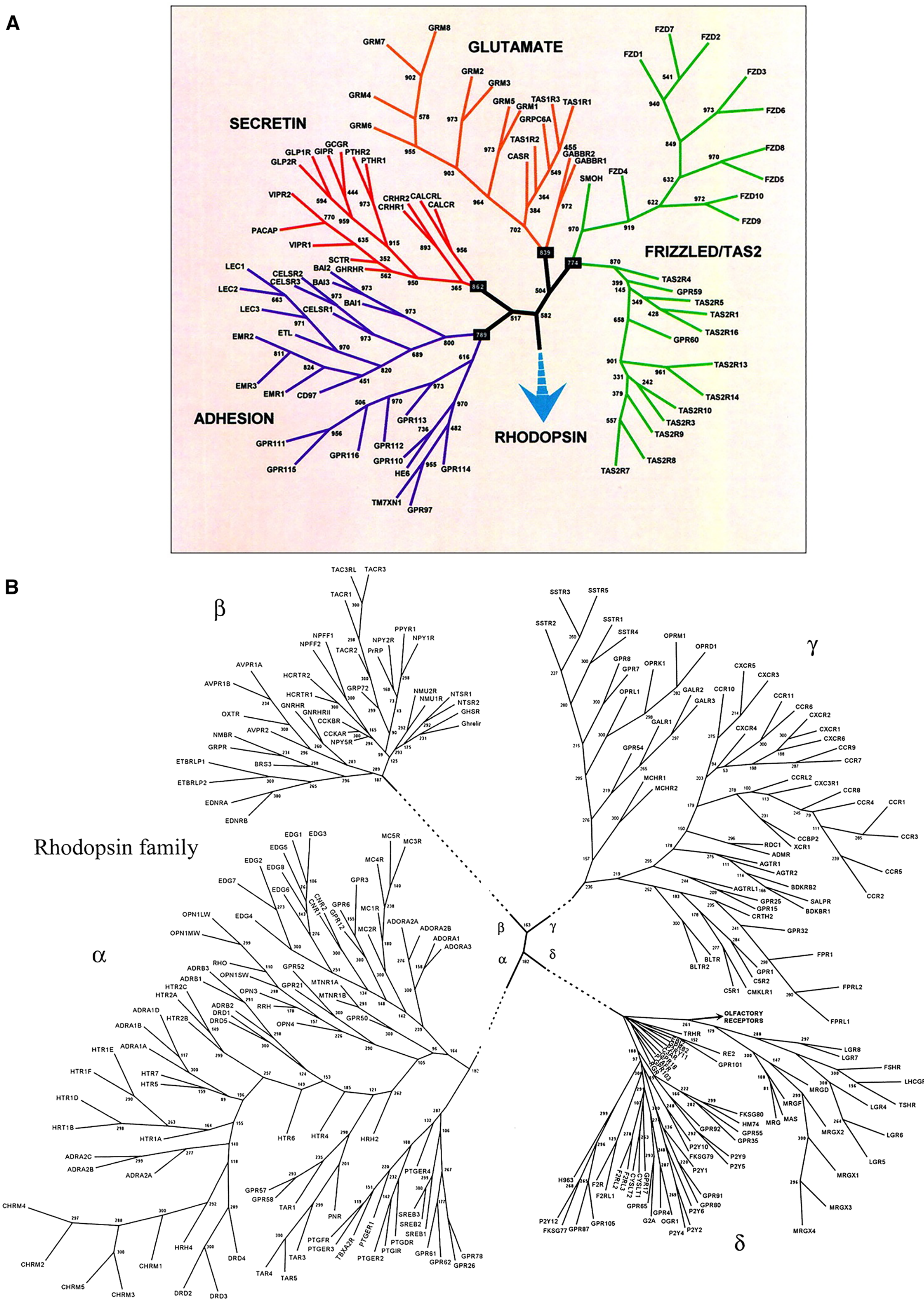


Fig. 1 Schematic representation of the mosaic nature of the protein structure–function space, where differently folded segments of a protein (foldons, semi-foldons, non-foldons, inducible foldons, and unfoldons) might possess different biological functions. One should keep in mind that 'Dormant disorder' is different from the other functional groups located in the 'outer-ring', as the corresponding segment does not describe a particular functional group but rather represents the means by which the functionality is achieved. Reproduced with permission from [68]


A
GLUTAMATE
SECRETIN
FRIZZLED/TAS2
ADHESION
RHODOPSIN
GRM8
GRM7
GRM2
GRM3
GRM4
GRM6
GRM5
GRM1
GRPC6A
TAS1R3
TAS1R1
TAS1R2
CASR
GABBR2
GABBR1
SMOH
FZD7
FZD2
FZD1
FZD3
FZD6
FZD8
FZD5
FZD4
FZD10
FZD9
GLP1R
GLP2R
GCGR
GIPR
PTHR2
PTHR1
VIPR2
PACAP
VIPR1
CRHR2
CRHR1
CALCRL
CALCR
SCTR
GHRHR
LEC1
LEC2
LEC3
CELSR2
CELSR3
BAI2
BAI3
BAI1
CELSR1
ETL
EMR2
EMR3
EMR1
CD97
GPR111
GPR115
GPR116
GPR113
GPR112
GPR110
HE6
GPR114
TM7XN1
GPR97
TAS2R4
GPR59
TAS2R5
TAS2R1
TAS2R16
GPR60
TAS2R13
TAS2R14
TAS2R10
TAS2R3
TAS2R9
TAS2R8
TAS2R7
B
Rhodopsin family
α
β
γ
δ
OLFACTORY RECEPTORS

◂**Fig. 2** **a** Phylogenetic relationship between the GPCRs (TMI–TMVII) in the human genome. The tree was calculated using the maximum parsimony method on 1000 replicas of the data set terminally truncated GPCR. The position of the rhodopsin family was established by including twenty random receptors from the rhodopsin family. These branches were removed from the final figure and replaced by an arrow toward the rhodopsin family. **b** The phylogenetic relationship between GPCRs (TMI–TMVII) in the human rhodopsin family. The tree was calculated using the maximum parsimony method on 300 replicas. The position of the olfactory cluster was established by including 17 diverse random receptors from the olfactory cluster. These branches were removed from the final figure and replaced by an arrow toward the olfactory receptor cluster. Reproduced with permission from [4]

Intrinsic disorder and multifunctionality of human GPCRs

Classification of human GPCRs

Globally, all GPCRs are characterized by similar structural organization always containing seven transmembrane helices, extracellularly located N-terminal domain and three extracellular loops (ECLs), and intracellularly located three intracellular loops (ICLs) and a C-terminal domain. The 7TM domain, which is a characteristic feature of the GPCR family, is structurally conserved between different GPCRs, whereas N-terminal and C-terminal domains are rather diversified and characterized by various lengths in different GPCRs. In different GPCRs, there are also significant differences in the lengths and amino acid sequences of ECLs and ICLs.

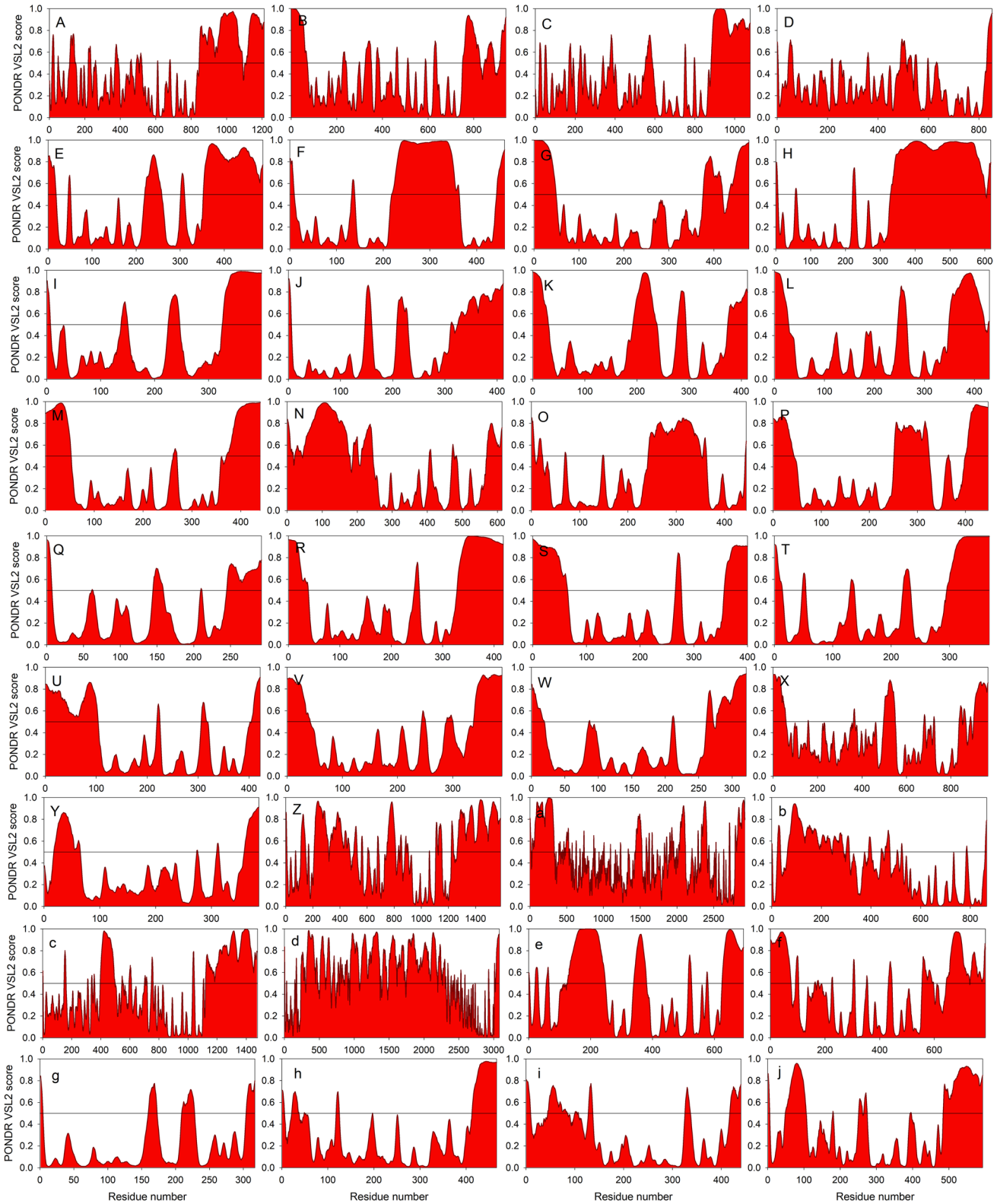
Based on their sequence homology and functional similarity, GPCRs from the *Chordata*, *Echinodermata*, *Arthropoda*, *Nematoda*, *Cnidaria*, *Placozoa*, and *Amoebozoa* phyla (i.e., from both vertebrates and invertebrates) were originally classified into six clans (A, B, C, D, E, and F) [71], with subclans named using roman numbers [72]. It was also pointed out that there is a large difference between the sequences of mammalian and invertebrate GPCRs, and that several of the A–F classes (e.g., clans D, E, and F, as well as family IV in clan A) are not found among human GPCRs [4]. Based on the systematic sequence analysis of human proteome, 802 human proteins were identified as GPCRs, and based on the outputs of the multiple phylogenetic analyses, this superfamily of seven transmembrane (TM) receptors was clustered into five main families, such as Glutamate (G, with 15 members), Rhodopsin (R, 701 members), Adhesion (A, 24 members), Frizzled/taste2 or Frizzled (F, 24 members), and Secretin (S, 15 members) [4]. These five families form the basis for the GRAFS (Glutamate (Class C), Rhodopsin (Class A), Adhesion (Class B), Frizzled (Class F), and Secretin (Class B)) classification of human GPCRs, with the corresponding phylogenetic relationship between different GRAFS proteins shown in Fig. 2 [4]. In addition to these 779 GRAFS proteins, a repertoire of human GPCRs includes a set of 23 other 7TM receptors that could not be assigned to any of these five families [4]. To illustrate highly diversified distribution of intrinsic disorder predisposition within the amino acid sequences of human GPCRs, Fig. 3 represents illustrative examples of disorder profiles of some of the most disordered members of the major groups of these proteins.

Sections below consider major characteristics of some of the members of the human GPCR families included into the GRAFS classification.

Functional intrinsic disorder of the members of glutamate receptor family

According to the GRAFS classification, the glutamate receptor family in human contains 15 members, including eight metabotropic glutamate receptors (GRM1, GRM2, GRM3, GRM4, GRM5, GRM6, GRM7, and GRM8), two GABA receptors (GABBR1 and GABBR2), a single calcium-sensing receptor (CASR), a single G protein-coupled receptor family C group 6 member A (GPRC6A), and three taste receptors (TAS1R1, TAS1R2, and TAS1R3) [4].

GRM1 (UniProt ID: Q13255) is a 1194-residue-long glutamate receptor containing an N-terminally located long extracellular ligand-recognition domain (residues 19–580), 7TM GPCR domain (residues 581–860), and a long intracellular C-terminal domain (residues 861–1194) engaged in interaction with G proteins. X-ray structures were solved for a significant part of the N-terminal extracellular domain (residues 28–523, PDB ID: 3KS9) and for the 7TM GPCR domain (residues 581-860, PDB ID: 4OR2, [73]). Although almost the entire N-terminal extracellular domain (residues 28–523) was used in the crystallization experiment, no structural information was obtained for several regions (residues 28–34, 128–153, 383–387, 489–490, and 513–523), indicating that these regions of missing electron density are characterized by high conformational flexibility. Furthermore, no structural information is currently available for the 524–580 region connecting N-terminal and 7TM domains and for the 334-residue-long C-terminal domain. More than 30% of GRM1 residues are predicted to be disordered, with the most of disorder being concentrated at the C-terminal cytoplasmic domain of GRM1 (residues 930–1,194, see Supplementary Materials, Figure S1). Furthermore, this protein contains multiple phosphorylation, methylation, ubiquitination, and glycosylation sites, and also includes several disorder-based interaction sites, molecular recognition features (MoRFs); i.e., disordered regions that are expected to fold at interaction with specific binding partners. Finally, GRM1 has three forms produced by AS, a canonical, full-length form, and two isoforms, a 906-residue-long isoform beta with region 887–906 changed from NSNGKSVSWSEPGGGQVPKG to KKRQPEFSPTSQCPSAHVQL



and missing region 907–1194, and isoform 3, with missing region 908–1194 and with region 887–907 changed from NSNGKSVSWSEPGGGQVPKGQH to KWRTGAQGTAYVAPPLCAREDQ.

Peculiarities of disorder distribution and some structural information available for the glutamate receptors GRM2, GRM3, GRM4, GRM5, GRM6, GRM7, and GRM8 are presented in Supplementary Materials, and their corresponding

◂**Fig. 3** Intrinsic disorder predisposition of human GPCRs. Here, PONDR® VSL2-generated disorder profiles of the most disordered members of different groups of GPCRs in five families are shown. The glutamate receptor family: **a** GRM5 (UniProt ID: P41594; *N* = 1212; PDR = 38.8%); **b** GABBR2 (UniProt ID: O75899; *N* = 941; PDR = 32.6%); **c** CASR (UniProt ID: P41180; *N* = 1078; PDR = 28.6%); **d** TAS1R3 (UniProt ID: Q7RTX0; *N* = 852; PDR = 13.3%). The rhodopsin family, α-group: **e** PTGER4 (UniProt ID: P35408; *N* = 488; PDR = 42.6%); **f** CHRM1 (UniProt ID: P11229; *N* = 460; PDR = 38.5%); **g** OPN4 (UniProt ID: Q9UHM6, *N* = 478; PDR = 28.9%); **h** GPR50 (UniProt ID: Q13585; *N* = 617; PDR = 49.3%); **i** EDG8 (UniProt ID: Q9H228; *N* = 398; PDR = 28.6%); **j** ADORA2A (UniProt ID: P29274; *N* = 412; PDR = 32.0%). The rhodopsin family, β-group: **k** MLNR (UniProt ID: O43193; *N* = 412; PDR = 31.6%); **l** NPFF1 (UniProt ID: Q9GZQ6; *N* = 430; PDR = 28.1%); **m** TAC3RL (NP_006670.1; *N* = 440; PDR = 29.6%); **n** ETBRLP1 (UniProt ID: O15354; *N* = 613; PDR = 45.2%); **o** NPY5R (UniProt ID: Q15761; *N* = 445; PDR = 31.9%); **p** CCKBR (UniProt ID: P32239; *N* = 447; PDR = 37.1%); **q** GNRHRII (NP_476504.1; *N* = 292; PDR = 24.7%). The rhodopsin family, γ-group: **r** SSTR3 (UniProt ID: P32745; *N* = 418; PDR = 32.8%); **s** OPRM1 (UniProt ID: P35372; *N* = 400; PDR = 30%); **t** GALR3 (UniProt ID: O60755; *N* = 368; PDR = 30.4%); **u** MCHR1 (UniProt ID: Q99705; *N* = 422; PDR = 32.9%); **v** BLTR2 (UniProt ID: Q9NPC1; *N* = 389; PDR = 28.8%). The rhodopsin family, δ -group: **w** MRGD (UniProt ID: Q8TDS7; *N* = 321; PDR = 24.3%); **x** LGR6 (UniProt ID: Q9HBX8; *N* = 967; PDR = 23.3%); **y** F2RL3 (UniProt ID: Q96RI0; *N* = 385; PDR = 21.0%). The adhesion receptor family: **z** ADGRB2 (UniProt ID: O60241; *N* = 1,585; PDR = 54.4%); **a** CELSR3 (UniProt ID: Q9NYQ7, *N* = 2,923; PDR = 40.8%); **b** EMR1 (UniProt ID: Q14246; *N* = 866; PDR = 36.0%); **c** LEC2 (UniProt ID: O94910; *N* = 1,478; PDR = 37.5%); **d** GPR112 (UniProt ID: Q8IZF6; *N* = 3,080; PDR = 64.4%). The frizzled/taste2 receptor family: **e** FZD8 (UniProt ID: Q9H461; *N* = 694; PDR = 39.2%); **f** SMOH (UniProt ID: Q99835; *N* = 787; PDR = 38.6%); **g** TAS2R14 (UniProt ID: Q9NYV8; *N* = 317; PDR = 15.5%). The secretin receptor family: **h** GIPR (UniProt ID: P48546; *N* = 466; PDR = 17.8%); **i** SCTR (UniProt ID: P47872; *N* = 440; PDR = 22.5%); **j** PTHR1 (UniProt ID: Q03431; *N* = 593; PDR = 33.6%)

D^2P^2 disorder profiles are shown in Supplementary Figures S2–S8.

GABBR1 (UniProt ID: Q9UBS5) serves as an agonist-binding component of the heterodimeric GPCR for γ-aminobutyric acid (GABA) [74] and has a domain organization typical of the members of GPCR family; i.e., an N-terminal domain (residues 15–591) containing Sushi-1 and Sushi-2 domains (residues 30–96 and 98–159, respectively) and ligand-binding domain (residues 160–591) with known 3D structure (residues 165–576, e.g., PDB ID: 4MS4, [75]), 7TM domain (residues 592–854), a cytoplasmic domain (residues 855–961) with known structure for the coiled-coil region responsible for hererodimerization (residues 879–919, PDB ID: 4PAS, [76]). GABBR1 is predicted to have 24% disordered residues, most of which are organized into two long IDPRs (residues 414–465 and 864–961). There are multiple glycosylation and phosphorylation sites in this protein, which is also predicted to possess 4 MoRFs, one located within the Sushi-2 domain (residues 171–176) and others (residues 898–904, 912–920, and 948–961) in the cytoplasmic C-terminal domain (see Figure S9). In addition to the canonical form of 961 residues, GABBR1 has multiple AS-generated isoforms ranging in length from 578 to 931 residues and characterized by changes in the N-terminal-most or C-terminal-most (residues 1–164 and 570–961) elements.

GABBR2 (UniProt ID: O75899) is a second subunit of the heterodimeric GABA receptor that serves as a mediator of coupling to G proteins [74] and that facilitates cell surface expression of GABBR1 by masking an its endoplasmic reticulum retention signal [77, 78]. Although GABBR2 does not bind any ligands, its N-terminally located ectodomain directly interacts with the GABBR1 ligand-binding ectodomain to enhance agonist affinity [79, 80] and is needed for the receptor activation [81]. GABBR2 has an N-terminal ectodomain (residues 42–483) stabilized by 3 disulfide bonds, 7TM transmembrane domain (residues 484–741), and 200-residue-long C-terminal cytoplasmic domain (residues 742–941). Crystal structure of a significant part (residues 42–466) of the N-terminal ectodomain was solved (PDB ID: 4MS4, [75]). This ectodomain was co-crystallized with the GABBR1 ectodomain in the absence of ligand and in the presence of various agonists and antagonists [75]. Furthermore, the coiled-coil region (residues 779–819) of the C-terminal domain GABBR2 was co-crystallized with the coiled-coil region of the GABBR1 subunit (PDB ID: 4PAS, [76]). GABBR2 is shown to contain multiple phosphorylation and glycosylation sites. Intrinsically disordered residues in this protein are amounting to 32.6%, being very abundantly present in its N- and C-termini (residues 1–71 and 751–941). The protein also has several acetylation sites and possesses 3 C-terminally located MoRFs (Figure S10).

CASR (UniProt ID: P41180) is a 1078-residue-long GPCR sensing changes in the extracellular concentration of calcium ions and playing a crucial role in the calcium homeostasis maintenance [82]. Besides being a major regulator of the extracellular Ca^{2+} homeostasis, CASR also plays important roles in embryonic development [83], regulation of neuronal excitability [84, 85], and nutrient sensing [86]. Functional form of CASR is a disulfide-crosslinked homodimer, each monomer containing four main domains, the N-terminal ligand-binding domain (residues 20–612) that includes a Venus Flytrap module containing two ligand-binding (residues 22–188 and 189–324) and two anion-binding regions (residues 66–70 and 415–417) and a cysteine-rich domain (residues 542–612), a 7TM domain (residues 613–862), and a cytoplasmic domain (residues 863–1078). In addition to two interchain disulfide bonds, each monomer of CASR is stabilized by eight intrachain disulfide bonds. Crystal structure of the entire extracellular domain of CASR was solved (residues 20–607, PDB ID: 5K5S). This protein is predicted to contain 28.6% disordered residues assembled

into several short and two long IDPRs (residues 548–592 and 878–1078), with C-terminal IDPR containing three MoRFs (see Figure S11).

Peculiarities of disorder distribution and some structural information available for the glutamate receptors GPRC6A, TAS1R1, TAS1R2, and TAS1R3 are presented in Supplementary Materials, and their corresponding D^2P^2 disorder profiles are shown in Supplementary Figures S12–S15.

Functionality and intrinsic disorder of the members of rhodopsin family

Among the 779 human GRAFS proteins, rhodopsin-like receptors represent the largest subfamily, which was subdivided into four main groups (α, β, γ, and δ) and 13 branches (five, one, three, and four main branches in α-, β-, γ-, and δ-groups, respectively). The α-group contains 89 receptors, which are subdivided into the prostaglandin receptor cluster of 15 members, amine receptor cluster of 40 members, opsin receptors cluster of 9 members, melatonin receptor cluster of 3 members, and MECA receptor cluster of 22 members. The β-group includes 35 receptors, whose ligands are various peptides. There are 59 receptors in the γ-group, that are further subdivided into the SOG receptor cluster with 15 members, MCH receptor cluster with 2 members, and the chemokine receptors cluster containing 42 receptors. Finally, 58 non-olfactory and 460 olfactory receptors constitute the δ-group containing four main branches, MAS-related receptor cluster with 8 members, glycoprotein receptor cluster of 8 members, purin receptor cluster with 42 members, and the olfactory receptor cluster estimated at 460 members [4].

Since the human rhodopsin subfamily includes 701 members, of which 241 are nonolfactory, we describe below only some of the representative members of the 13 main branches of the four groups of rhodopsin-like GPCRs. The vast majority of the members of rhodopsin subfamily have a simplified domain organization, containing only 7TM domain, and as a result, their ligands bind within a cavity between the TM regions. The only exception from this rule is some glycoprotein-binding receptors that contain a dedicated N-terminally located ligand-binding domain.

The α-group of rhodopsin receptors

Prostaglandin receptor cluster has eight prostaglandin receptors and seven orphan receptors [4]. The proteins included into this cluster are G protein coupled receptors 26, 61, 62, and 78 (GPR26, GPR61, GPR62, and GPR78 containing 337, 451, 368, and 363 residues, respectively), prostaglandin D2 receptor (PTGDR with 359 residues), prostaglandin E2 receptors subtypes EP1, EP2, EP3, and EP4 (PTGER1, PTGER2, PTGER3, and PTGER4 containing 402, 358, 390, and 488 residues, respectively), prostaglandin F2α receptor (PTGFR, 352 residues), prostacyclin receptor (PTGIR, 386 residues), super conserved receptors expressed in brain 1, 2, and 3 (SREB1, SREB2, and SREB3 containing 375, 370, and 373 residues), and thromboxane A2 receptor (TBXA2R with 343 residues). Thus, members of the prostaglandin receptor cluster range in length from 337 to 488 residues. Therefore, for more detailed characterization of these receptors, we selected three proteins with the shortest, medium, and the longest sequences.

GPR26 (UniProt ID: Q8NDV2) is a constitutively active orphan GPCR consisting of 337 residues and containing a 10-residue-long N-terminus followed by a 287-residue-long 7TM domain (residues 11–297) and a 40-residue-long cytoplasm domain (residues 298–337). GPR26 is predicted to contain 12.8% disordered residues, that include part of a short N-terminus (residues 1–5) and two relatively long IDPRs (residues 227–241 and 317–337), with the first of these longer IDPRs being in the 56-residue-long intracellular loop-3 (ICL3, residues 190–245) (Figure S16).

PTGDR (UniProt ID: Q13258) is a 359-residue-long receptor with a very short extracellular N-terminus of 21 residues, 7TM domain (residues 20–331), and a short cytoplasmic C-tail (residues 332–359). PTGDR is predicted to contain 19.5% disordered residues (see Figure S17), largest portion of which is assembled into 31-residue-long IDPR (residues 230–260), which is a part of the 51-residue-long ICL3 (residues 217–262). In addition to the canonical form, AS generates a shortened isoform that does not have C-terminal residues 297–359 and has region 283–296 changed from YRAYYGAFKDVKEK to AFVPGVPAKTPGSR. There are several phosphorylation and glycosylation sites in this receptor.

PTGER4 (UniProt ID: P35408) is the longest member of the cluster of prostaglandin receptors. This 488-residue-long GPCR serves as a receptor for prostaglandin E2 and possesses a relaxing effect on smooth muscles. The increased length of this receptor is defined by the presence of the 156-residue-long C-terminal cytoplasmic domain, whereas the remaining part of the protein is similar to other prostaglandin receptors, consisting of an extracellular N-terminus of 19 residues and a 7TM domain (residues 20–332) that contains a 56-residue-long ICL3 (residues 212–267). Almost half of PTGER4 (42.6%) is predicted to be disordered, with the longest IDPR (136 residues) being located within the cytoplasmic C-terminal domain, and with another long IDPR (residues 220-258) being a part of the ICL3. There are multiple PTM sites in this protein, which also contains five C-terminally located MoRFs (see Figure S18).

Amine receptor cluster includes adrenergic receptors (ADRA1A, ADRA1B, ADRA1D, ADRA2A, ADRA2B, ADRA2C, ADRB1, ADRB2, and ADRB3), dopamine receptors (DRD1, DRD2, DRD3, DRD4, and DRD5), histamine receptors (HRH2 and HRH4), muscarinic receptors

(CHRM1, CHRM2, CHRM3, CHRM4, and CHRM5), serotonin receptors (HTR1A, HTR1B, HTR1D, HTR1E, HTR1F HTR2A, HTR2B, HTR2C, HTR4, HTR5(HTR5A), HTR6, and HTR7), trace amine receptors (TAR1, TAR3, TAR4, TAR5), and several orphan receptors (GPR58 and GPR57) [4] activated by structurally related small amine molecules with a single aromatic ring. Peculiarities of disorder distribution and some structural information available for the ADRA1A, CHRM1, DRD1, HRH2, HTR1A, and TAR1 are presented in Supplementary Materials, and their corresponding D^2P^2 disorder profiles are shown in Supplementary Figures S19–S24.

Opsin receptors cluster includes the only known human receptors that respond to light. The members of this cluster are the cone visual pigments, such as OPN1SW, OPN1LW, and OPN1 MW, which are the light-absorbing molecules that mediate vision, the encephalopsin (OPN3), the melanopsin (OPN4), the peropsin (RRH), the retinal G protein-coupled receptor (RGR), the rod visual pigment (RHO), and two orphan receptors GPR21 and GPR52. Supplementary materials represent some structural and disorder-related information for OPN1SW, OPN1LW, OPN1 MW, RRH, RHO, GPR21, GPR52, OPN3, and OPN4, and show their D^2P^2 disorder profiles (see Figures S25–S32).

Melatonin receptor cluster includes the melatonin receptor types 1A and 1B (MTNR1A, UniProt ID: P48039, 350 residues, 13.1% PDRs and MTNR1B, UniProt ID: P49286, 362 residues, 19.1% PDRs) with core GPCR structure, as well as an orphan melatonin-related receptor GPR50 (UniProt ID: Q13585, 617 residues, 49.3% PDRs) that in addition to the 7TM domain includes a long cytoplasmic domain (residues 295-617) predicted to be excessively disordered and to contain six MoRFs (see Figures S33–S35). Some structure- and disorder-related information for the members of the MECA receptor cluster is listed in Supplementary Materials.

The β-group of rhodopsin-like receptors

All rhodopsin-like receptors from the β-group are activated by peptides. This GPCR group includes the hypocretin receptors HCRTR1 and HCRTR2; the neuropeptide FF receptors NPFF1 and NPFF2; the tachykinin receptors TACR1, TACR2, TACR3, and TAC3RL; the cholecystokinin receptors CCKAR and CCKBR; the neuropeptide Y receptors NPY1R, NPY2R, NPY5R, and PPYR1; the endothelin-related receptors EDNRA, EDNRB, ETBRLP1, and ETBRLP2; the gastrin-releasing peptide receptor GRPR; the neuromedin B receptor NMBR; the uterin bombesin receptor BRS3; the neurotensin receptors NTSR1 and NTSR2; the growth hormone secretagogues receptor GHSR; the neuromedin receptors NMU1R and NMU2R; the thyrotropin releasing hormone receptor TRHR; the arginine vasopressin receptors AVPR1A, AVPR1B, and AVPR2; the gonadotropin-releasing hormone receptors GNRHR and GNRHRII; the prolactin-releasing peptide receptor PrRP; the oxytocin receptor OXTR; the motilin receptor MLNR; and an orphan receptor GPR72 [4]. Brief description of disorder status of these receptors and several illustrative disorder profiles (see Figures S36–S46) are provided in Supplementary Materials.

The γ-group of rhodopsin receptors

SOG receptor cluster includes receptors for somatostatin (SSTR1, SSTR2, SSTR3, SSTR4, and SSTR5), opioids (OPRD1, OPRK1, OPRL1, and OPRM1), neuropeptide galanin (GALR1, GALR2, and GALR3), neuropeptide W (GPR7 and GPR8), and RF-amide (GPR54) [4]. Brief description of the disorder status of the members of this cluster is provided in Supplementary Materials, where some illustrative D^2P^2 profiles are also shown (see Figures S47–S49).

MCH receptor cluster includes just two receptors for the melanin-concentrating hormone (MCH), MCHR1 (UniProt ID: Q99705), 422 residues, 32.9% PDRs, four IDPRs (residues 1–103 (N-tail), 220–225 (part of ICL2), 307–318 (part of ICL3), and 405-422 (part of C-tail) and MCHR2 [UniProt ID: Q969V1, 340 residues, 6.8% PDRs, two IDPRs (residues 1–5 and 324–340)]. Figure S50 shows that long N-terminally located IDPR of MCHR1 contains 3 MoRFs.

Chemokine receptors cluster includes more than 40 members, that range in length from 333 to 469 residues and contain from 9.0 to 28.8% disordered residues, which are invariantly located within the N- and C-terminal tails, as well as within some of the intracellular or extracellular loops (see Supplementary Materials for more detailed description of the peculiarities of disorder distribution in these proteins). To illustrate the variability of intrinsic disorder distribution within these proteins, Figures S51 and S52 represent the D^2P^2-generated disorder profiles for the most and least disordered members of the cluster of chemokine receptors, BLTR2 ($N = 389$, PDR = 28.8%) and CCR3 ($N = 355$, PDR = 9.0%), respectively.

The δ-group of rhodopsin receptors

MAS-related receptor cluster is rather small, containing 8 members. All these receptors are rather short and are characterized by the shortened cytoplasmic loops ICL1, ICL2 and ICL3. They also show similar disorder distribution, which is illustrated by Figure S53 showing disorder profile generated by D^2P^2 for the MAS-related G protein-coupled receptor member D, MRGD, which is the most disordered member of this cluster (see Supplementary Materials for more details on the peculiarities of disorder predisposition of these proteins).

As it was already indicated, glycoprotein receptors are exceptional members of the rhodopsin family, since they

have long N-terminal extracellular domain (that ranges from 337 (LHCGR) to 349 (FSHR), or 393 (TSHR), or 409 (LGR7), or 416 (LGR8), or 520 (LGR4), or 540 (LGR5), or even 543 (LGR6) residues), which is dedicated to the ligand binding, whereas other members of the rhodopsin family bind their ligands in a cavity between their TM regions. Furthermore, glycoprotein receptors have long C-terminal cytoplasmic domains, which can be as long as 64 (LGR8), or 65 (FSHR), or 72 (LHCGR), or 82 (TSHR), or 84 (LGR5), or 85 (LGR7), or 137 (LGR6), or 157 (LGR4) residues. As a result, glycoprotein receptors are noticeably longer than other representatives of the rhodopsin family. Figures S54–S59 represent available D^2P^2 profiles for the members of this cluster, whereas some details of their disorder predisposition are listed in Supplementary Materials.

Purin receptor cluster has more than 40 members and includes a set of the formyl peptide receptors (FPRs), a set of the nucleotide receptors [such as P2Y1(P2RY1), P2Y2(P2RY2), P2Y4(P2RY4), P2Y5, P2Y6(P2RY6), P2Y9(GPR23), P2Y10, P2Y11(P2RY11), P2Y12, FKSG77(GPR86, GPR94), FKSG79, and FKSG80(GPR81)], proteinase-activated receptors (such as F2R, F2RL1, F2RL2, and F2RL3), receptors activated by cysteinyl leukotrienes (CYSLT1 and CYSLT2) or hydroxycarboxylic acid receptors C3AR(C3AR1) and HM74, as well as a large number of orphan GPCRs, such as GPR4, GPR17, GPR18, GPR35, GPR55, GPR65, GPR80(GPR99), GPR82, GPR87, GPR91, GPR92, GPR101, GPR103, GPR105, H963, OGR1(GPR68), TRHR, RE2, RGR, PTAFR, G2A, and EBI2. Since topologically members of this cluster are not too different from other representatives of the rhodopsin family (with the obvious exception to the glycoprotein receptors), we will describe below only a couple of illustrative examples.

P2Y1(P2RY1) (UniProt ID: P47900) is a 373-residue-long receptor for extracellular adenine nucleotides, such as ADP, that was shown to possess two disparate binding sites for its ligands, with the nucleotide binding site being located in a pocket created by the residues from N-terminus, ECL2 and TM helices VI and VII, whereas the second, allosteric, non-nucleotide binding site being located in a shallow pocket formed by aromatic and hydrophobic residues of TM helices I, II and III and ECL1 positioned completely outside of the helical bundle [87]. P2RY1 has a typical 7TM core structure decorated with extended N- and C- terminal regions of almost equivalent length (residues 1–51 and 326–373). This receptor is predicted to contain 16.1% disordered residues assembled in several IDPRs (residues 1–3, 23–36, 251–263, and 334–373). It contains two disulfide bonds (between residues 42 and 296, and 124 and 202) and multiple PTMs sites (see Figure S60). Figure S61 shows that a 385-residue-long receptor for activated thrombin or trypsin (F2RL3), which is involved in platelet activation, is characterized by the presence of long, mostly disordered N- and C-terminal regions and many PTM sites.

Olfactory receptor cluster, with its at least 460 members grouped into 17 main branches, clearly represents the largest group of the rhodopsin-like receptors [4]. Detailed analysis of this cluster constitutes a subject of separate study. Length of the human olfactory receptors ranges from 369 in the olfactory receptor 2T1 (OR2T1) to 298 residues in olfactory receptor 52Z1 (OR52Z1). OR2T1 (UniProt ID: O43869) is characterized by a long N-terminal domain (residues 1–76), and is predicted to have rather low intrinsic disorder level (PDR = 8.1%), containing only short IDPRs (residues 51–59, 283–287, 316–320 and 360–369) (see Figure S62). The predicted intrinsic disorder level of the shortest olfactory receptor OR52Z1 (UniProt ID: P0C646) is very low (2%).

Functional intrinsic disorder of the members of adhesion receptor family

This family of GPCRs includes 24 members, which are characterized by very peculiar topology, where in addition to the canonical 7TM domains, these proteins contain large N-terminal regions that range in length from 200 to 2800 residues, possessing various functional domains with adhesion-like motifs and often forming mucin-like stalks enriched in the glycosylation sites and proline residues [4]. Also, many members of the adhesion receptor family possess long C-terminal cytoplasmic domains.

Brain-specific angiogenesis inhibitor 1, BAI1, also known as adhesion G protein-coupled receptor B1 (ADGRB1), (UniProt ID: O14514) is a 1584 residue-long phosphatidylserine receptor with the N-terminal region of 918 residues (31–948) and a C-terminal cytoplasmic domain of 397 residues (1188–1584). For a long time, BAI1, which was discovered as a protein, whose extracellular region is able to inhibit angiogenesis in tumor models, remained an orphan receptor with elusive physiological functions. However, this protein is considered now as a phagocytic receptor recognizing apoptotic cells with exposed phosphatidylserine [88]. BAI1 can also promote cytoskeletal reorganization, which is needed to facilitate the phagocytic clearance of apoptotic cells, acting upstream of the of the ELMO/Dock180/Rac signaling module [88]. This receptor promotes myoblast fusion, enhances the engulfment of apoptotic cells [89], inhibits angiogenesis in the brain, and mediates the p53 signaling related to the glioblastoma suppression [90]. Recent work established that BAI1 and other members of this subfamily (BAI2 and BAI3) are important regulators of synaptogenesis and dendritic spine formation [91]. Proteolytic processing of BAI generates two secreted fragments, Vasculostatin-120 (residues 31–926) [92] and Vasculostatin-40 (residues 31–327) [93], that participate in the regulation of vascular homeostasis. BAI is predicted to contain 732 disordered residues (i.e.,

it has the PDR of 46.2%), with the C-terminal cytoplasmic domain being almost completely disordered and with several long IDPs being located within the long N-terminal region (e.g., residues 133–171, 200–236, 276–398, 414–503, and 539–580). Figure S63 shows that besides multiple IDPRs and PTM sites, BAI1/ADGRB1 is predicted to have multiple disorder-based interaction sites, MoRFs, many of which are located within the C-terminal cytoplasmic domain. Curiously, BAI1/ADGRB1 contains five thrombospondin (TSP) type-1 repeats in the extracellular domain (residues 261–315, 354–407, 409–462, 467–520, and 522–575) that mediate binding of this protein to phosphatidylserine and are needed for recognition and binding of bacterial outer membrane lipopolysaccharides. These TSP type-1 repeats are predicted to contain high levels of disorder, overlapping with the long N-terminal IDPRs. BAI1 also contains an integrin-binding RGD motif (residues 231–233) [94] located within an IDPR. Cytoplasmic C-terminal domain of BAI includes a proline-rich region (residues 1384–1451) containing two poly-proline motifs (residues 1411–1422 and 1425–1430) that can interact with Src homology 3 (SH3) and WW domain-containing proteins and thereby participate in the regulation of signal transduction [95]. The C-terminal most part of BAI1 is the QTEV motif that represents a binding site for PDZ domain-containing proteins, which are associated with multiple cellular functions ranging from modulation of cytoskeletal architecture to regulation of localization and trafficking of proteins [94]. Both the proline-rich region and QTEV-containing C-tail are highly disordered, suggesting importance of intrinsic disorder for the multifunctionality of the C-terminal cytoplasmic domain. Furthermore, the C-terminal BAI region (residues 1365–1584) is involved in interaction with MAGI1 (which is membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1 that plays a crucial role as scaffolding protein at cell–cell junctions) is not only predicted to be heavily disordered, but also is expected to have four MoRFs [96].

Structural, functional, and disorder-related information for BAI2, BAI3, CELSR1, CELSR2, EMR1, EMR2, EMR3, and LEC1 with corresponding D^2P^2 profiles (see Figures S64-72) is presented in Supplementary Materials.

Lectomedin-2 (LEC2, UniProt ID: O94910; PDR = 37.5%) is a 1478-long calcium-independent receptor for α-latrotoxin, which is potentially related to the regulation of exocytosis and can play a role in mediation of the heterophilic synaptic cell–cell contact and postsynaptic specialization, whereas lectomedin-3 (LEC3, UniProt ID: Q9HAR2; PDR = 31.7%) is a 1447-residue long adhesion GPCR interacting with the fibronectin-like domain-containing leucine-rich transmembrane proteins FLRT2 and FLRT3 at the surface of adjacent cells and thereby regulating cell–cell adhesion and neuron guidance [97]. N-terminal regions of these GPCRs contain 834 and 847 residues, whereas their C-terminal cytoplasmic domains are 378- and 343-residue long. Figure S73 illustrates the presence of numerous IDPRs, PTMs and MoRFs in LEC2. Unfortunately, no D^2P^2-generated profile is currently available for LEC3, although this protein is also predicted to have multiple long IDPRs (e.g., residues 393–499, 1120–1201, and 1247–1390). All members of this subfamily have multiple AS-generated isoforms.

Leukocyte antigen CD97 (UniProt ID: P48960) is an 835-residue-long calcium-dependent receptor containing several EGF-like domains and regulating adhesion and signaling processes after leukocyte activation. N- and C-terminal regions of CD97 are of 532 and 46 residues, respectively. Most of the 308 disordered residues of this protein are concentrated within the N-terminally located long IDPR (residues 40–250) that includes most of its EGF-like domains. Curiously, although the N-terminal region containing EGF-like domains is predicted to be disordered, a crystal structure was solved for the construct containing EGF-like domains 1 and 2 (residues 25–118) directly linked to the EGF-like domain 5 (residues 212–260) of human CD97 in the presence of calcium an barium (PDB ID: 2BOU) [98], suggesting the involvement of the binding-induced folding mechanism in the formation of protein–cation structure. In addition to multiple PTMs, CD97 contains one MoRF (Figure S74) and is known to have multiple AS-generated isoforms.

GPR56 is a 693-residue-long adhesion GPCR G1, also known as protein TM7XN1 (UniProt ID: Q9Y653), with N- and C-terminal regions of 377 and 36 residues, respectively. GPR56 plays a crucial role in cortical development, with mutations in this protein causing malformations of the cerebral cortex [99]. This adhesion GPCR is characterized by rather unusual activation, where the large N-terminal (NT) region (residues 26–382) is cleaved from the rest of the GPR56 receptor during processing but remains non-covalently bound to the 7TM domain. This NT region inhibits the receptor activity, but this inhibition can be released via the NT binding to extracellular ligand and/or due to the GPR56 NT homophilic trans–trans interactions [99]. Here, ligand binding to the NT region leads to its release from the receptor, resulting in the exposure of a new N-terminal stalk required for the stimulation of the G protein-dependent signaling activity [100]. Figure S75 shows that GPR56 contains significant level of intrinsic disorder (PDR = 31.2%), possessing multiple IDPRs (including one long IDPR, residues 126–218), numerous PTMs, and one MoRF (residues 182–189).

The remaining members of the adhesion receptor family are orphan receptors with practically unknown physiological functions. Since the functional roles of their disordered regions were not established yet, Supplementary Materials provide a very superficial description of these adhesion GPCRs, which are split into two groups, long and short,

based on the length of their amino acid sequences. The only member of this group that will be mentioned below is GPR112 (UniProt ID: Q8IZF6; $N = 3080$), which, with its 64.4% disordered residues, is one of the longest and most disordered GPCRs (see Figure S76). N-terminal region of this protein contains 2714 residues, whereas its C-tail is 94-residue long. The longest IDPR (449 residues) of GPR12 is longer than many full-length GPCRs from other families. Numerous MoRFs and PTM sites are spread over the disordered regions of this protein.

Intrinsic disorder in functionality of the frizzled/taste2 receptor family members

GPCRs included into the frizzled/taste2 GPCR family form two distinct groups, frizzled (that in addition to 10 frizzled receptors (FZDs) contains the smoothened receptor (SMO)) and taste receptors that includes 13 GPCRs [4].

Frizzled-1 (FZD1, UniProt ID: Q9UP38; PDR = 31.4) is a 647-residue-long receptor for several WNT proteins that can be activated by WNT1, WNT3, WNT3A, and to a lesser extent WNT2, and that is involved in the canonical WNT/β-catenin signaling pathway [101], which plays a number of crucial roles in the organism development, regulates maintenance, self-renewal, and differentiation of adult mammalian tissue stem cells [102], and the malfunctions of which are associated with various diseases including cancer, cardiovascular diseases, and Parkinson's disease [103–107]. Overall, Wnt signaling is involved in multiple physiological and pathological processes, plays an important role in angiogenesis [108, 109], and is well studied in embryogenesis and carcinogenesis [110]. Human FZD1 contains a 250-residue-long N-terminal extracellular domain (residues 73–322) containing FZ domain (residues 111–230) responsible for the WNT binding and short C-terminal cytoplasmic tail (residues 623–647). Significant part of the N-terminal region is intrinsically disordered. In fact, Figure S77 shows that two long IDPRs (residues 1–115 and 230–275) that surround the FZ domain of human FZD1 are located here. Furthermore, part of the C-tail (residues 639–647) is disordered, too. Although the WNT-binding FZ domain is predicted to be mostly structured, which is not a big surprise, since this domain includes five disulfide bridges, the presence of PTM sites in IDPRs adjacent to this domain suggests that function of FZ domain can be modulated by PTMs. Since N-tail of FZD1 is predicted to have MoRF region, and since there are multiple PTM sites in this protein, it is clear that at least some functionality of FZD1 is intrinsic disorder dependent. Data collected in Supplementary Materials show that human FZD2 (Figure S78), FZD4, and FZD7 have structural organization and disorder profiles similar to those of FZD1, whereas FZD3, FZD5, FZD6, FZD9, and FZD10 are characterized by the presence of long C-terminal domains, which are comparable in length with their N-terminal extracellular regions (see Figure S79), and FZD8 is characterized by a unique disorder profile containing three long IDPRs (Figure S80).

The last member of the frizzled group is the Smoothened homolog (SMOH, UniProt ID: Q99835), which is a 787-residue-long GPCR that serves as a key receptor in regulation of the Hedgehog (HH) signaling pathway [111]. Curiously, activation of SMOH is associated with the interaction of the secreted HH protein with the 12TM Patched-1 (PTCH) protein, which is the membrane-embedded receptor that, in its HH-unbound form, interacts with SMOH thereby inhibiting its activity [112]. Since HH pathway serves as one of the major regulators of development, cell proliferation, and maintenance of the stem cells [113–115], its action is strictly controlled and tightly regulated, e.g., via regulation and control of the SMOH receptor. In fact, insufficient activity of HH signaling is associated with various developmental defects (e.g., cyclopia and holoprosencephaly), whereas excessive activity of this pathway is related to the cancer pathogenesis [111, 116–118]. This is the reason why SMOH is commonly considered as an oncoprotein [119]. Although no structural information is currently available for the full-length SMOH protein, crystal structures were solved for some of its parts, such as 7TM domain [e.g., PDB ID: 4QIN [120]) and a multidomain construct containing FZ domain (also known as cysteine-rich domain (CRD)] and 7TM domain (e.g., PDB ID: 5L7D [121]). Topologically, SMOH is characterized by the presence of long N- and C-terminal regions, which are of 206 and 242 residues long, respectively. Figure S81 shows that although the 7TM domain of SMOH (residues 234-545) is clearly defined as mostly ordered (it possesses only a few short IDPRs), the FZ/CRD of this protein is predicted to be rather disordered, which makes it quite different from FZDs, where this domain is predicted to be mostly different. Overall, high level of predicted disorder in SMOH (PDR = 38.6%) and the presence of long N- and C-terminally located IDPs (residues 1–75 and 644–787) combined with the existence of numerous MoRFs (located in both tails) and multiple PTMs clearly indicates that intrinsic disorder is needed for SMOH function, and that this protein contains all types of the proteoforms, conformational/basic, inducible/modified, and functioning.

Finally, a group of taste receptors includes a set of short GPCRs with almost absent N-terminal elements and very short C-terminal cytoplasmic regions. As a matter of fact, the length of the extracellular N-terminal regions in these proteins ranges from 1 to 14, whereas their cytoplasmic C-tails are 13- to 35-residue long. As a result, these GPCRs do not contain long IDPRs and their overall level of disorder ranges from 3.4% to 15.5%.

Functional intrinsic disorder of the members of secretin receptor family

The last family of human GPCRs is represented by 15 receptors that bind rather large peptides. Among the members of this family are the calcitonin receptors CALCR and CALCRL, the corticotropin-releasing hormone receptors CRHR1 and CRHR2, the gastric inhibitory polypeptide receptor GIPR, the glucagon receptor GCGR, the growth hormone-releasing hormone receptor GHRHR, the glucagon-like peptide receptors GLP1R and GLP2R, the pituitary adenylyl cyclase-activating protein PACAP, the parathyroid hormone receptors PTHR1 and PTHR2, the secretin receptor SCTR, and the vasoactive intestinal peptide receptors VIPR1 and VIPR2 [4].

CALCR (UniProt ID: P30988; $N=508$) and CALCRL (UniProt ID: Q16602; $N=461$) are the receptors for calcitonin and calcitonin-gene-related peptide (CGRP), respectively. The 7TM domains of CALCR and CALCRL (residues 172–429 and 147–388, respectively) are inserted between the relatively long N- and C-terminal regions consisting of 129 and 79 residues in CALCR and 124 and 73 residues in CALCRL. Using the Volta phase plate single-particle cryo-electron microscopy, the low-resolution structure of the full-length CALCR complexed with the peptide ligand and heterotrimeric $G\alpha_s\beta\gamma$ protein has been determined (PDB ID: 5UZ7 [122]). Although the full-length mature CALCR was used in this study, no structure was determined for residues 43–153, 224–230, 350–354, 379–383, and 437–508 that, therefore, constitute regions of missing electron density; i.e., regions with high structural dynamics. In other words, based on this structural analysis, about 200 residues of human CALCR (i.e., 39.4%) are expected to be rather mobile. As follows from evaluation of the disorder predisposition of this protein, the indicated segments either corresponded to the flexible regions or IDPRs (see Fig. 4a). In fact, CALCR is predicted to have 89 disordered residues (PDR = 17.5%), more than half of which being located within the long C-terminally located IDPR (residues 453–508). Similarly, crystal structure was solved for the N-terminal ligand recognizing domain of CALCRL (e.g., PDB ID: 3AQF, [123]), whereas utilization of the cryo-electron microscopy generated low-resolution structure of the mature CALCRL (residues 22-461) complexed with the calcitonin-gene-related

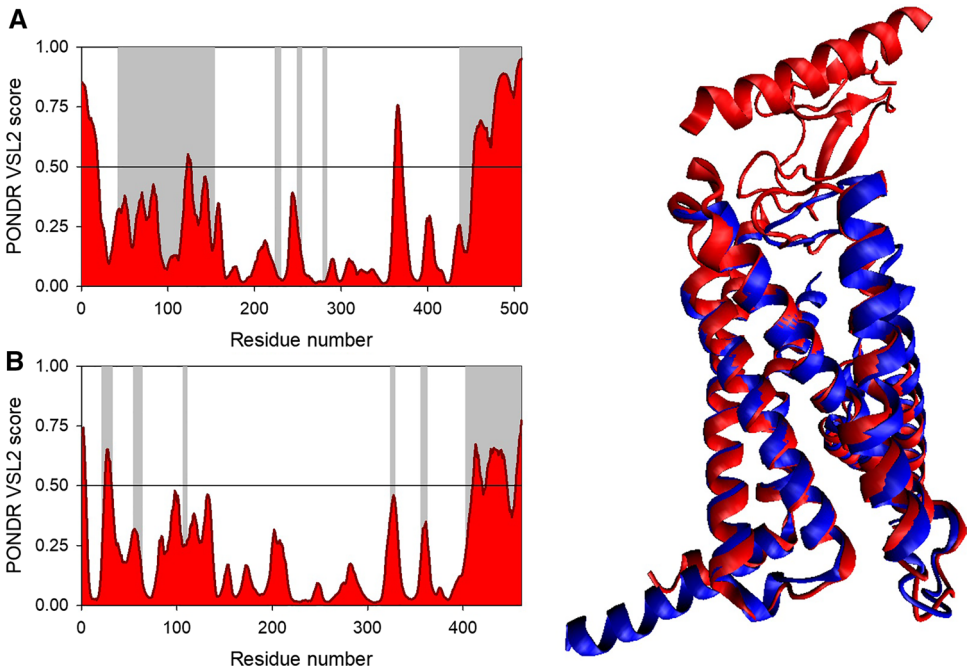


Fig. 4 Analysis of intrinsic disorder predisposition and structure of human CALCR (UniProt ID: P30988; $N=508$) and CALCRL (UniProt ID: Q16602; $N=461$). **a**, **b** show intrinsic disorder predispositions of these proteins in a form of predicted disorder profiles generated by a set of commonly used disorder predictors, such as. **c** Represents structural alignment of bound form of CALCR complexed with the peptide ligand and heterotrimeric $G\alpha\beta\gamma$ protein has been determined (PDB ID: 5UZ7 [122], blue structure) and bound form of the mature CALCRL (residues 22–461) complexed with the calcitonin-gene-related peptide (CGRP), receptor activity-modifying protein 1 (RAMP1), and the $G\alpha_s$-protein heterotrimer (PDB ID: 6E3Y [124], red structure). In both cases, structures were determined using the Volta phase plate single-particle cryo-electron microscopy. Note that although the reported 5UZ7 and 6E3Y structures included complexes of the human CALCR or CALCRL with the corresponding $G\alpha_s$-protein heterotrimers and other binding partners, only structures of human CALCR or CALCRL are shown here for simplicity. **a**, **b** Positions of regions with missing electron density are shown as gray vertical bars

peptide (CGRP), receptor activity-modifying protein 1 (RAMP1), and the $G\alpha_s$-protein heterotrimer (PDB ID: 6E3Y [124]). Analysis of this structure revealed that residues 22–32, 55–63, 107–109, 324–328, 356–362, and 403–461 of CALCRL constitute regions of missing electron density, indicating high conformational flexibility of these 94 residues (20.4%). This is in good agreement with the results of the analysis of disorder predisposition of this protein, where the regions with missing electron density were either predicted to be disordered or flexible (i.e., were characterized by the disorder scores exceeding 0.5 and 0.25, respectively, see Fig. 4b). Figure 4b shows that the overall level of predicted intrinsic disorder in CALCRL is not very high (PDR = 11.1%), and this protein does not possess long IDPRs. Therefore, CALCR and CALCRL are among a very few human GPCRs containing long N-terminal extracellular and C-terminal cytoplasmic regions, for which structure is known for almost entire protein. Figure 4c represents aligned structures of these proteins and shows close structural similarity of resolved regions (particularly 7TM domains and other regions in their close proximity). In fact, aligned structures covered 240 residues in each protein (of 266 and 346 residues of CALCR and CALCRL used in this analysis) and RMSD of heavy atoms in this alignment was impressively low (1.23 Å), which is an impressive accuracy taking into account that the corresponding structures were resolved to 3.3 Å.

The corticotropin-releasing hormone receptors CRHR1 (UniProt ID: P34998; $N = 444$; $N_{NT} = 88$; $N_{CT} = 47$; PDR = 12.2%) and CRHR2 (UniProt ID: Q13324; $N = 411$; $N_{NT} = 108$; $N_{CT} = 47$; PDR = 10.7%) are activated by highly specific interaction with corticotropin-releasing factor (CRF) and urocortin (UCN) accompanied by the conformation changes in these GPCRs that trigger down-stream signaling [125]. Figure 5 represents an NMR solution structure of a complex between the N-terminal extracellular domain (ECD) of human CRHR1 (residues 25–108) and a high affinity peptide agonist, α-helical cyclic CRF (PDB ID: 2L27; [126]). Importantly for subsequent functionality of CRHR1, formation of this ECD–CRF complex promotes induction of the α-helical conformation in the N-terminus of ECD leading to the activation of CRHR1 [127]. It was pointed out that the overall fold of the ECD represents "the short consensus repeat (SCR) motif with the three disulfide bonds between cysteine residues, Cys_{30}–Cys_{54}, Cys_{44}–Cys_{87}, and Cys_{68}–Cys_{102}. The SCR includes a short N-terminal α-helix (Asp_{27}–Glu_{31}) and two anti-parallel β-sheet regions around residues Ser_{47}–Val_{48} (β_1 strand), Cys_{54}–Trp_{55} (β_2 strand), Leu_{63}–Arg_{66} (β_3 strand), and Gly_{83}–Glu_{86} (β_4 strand)" [126]. Because of the presence of three disulfide bridges, the ECD of human CRHR1 is predicted to be mostly ordered. However, since interaction with CRF induced structural changes in ECD leading to the activation of CRHR1, this system represents a clear example of functioning proteoform. Furthermore, both CRHR1 and CRHR2 have multiple AS-generated isoforms and also have multiple PTM sites, indicating that these proteins also have inducible proteoforms.

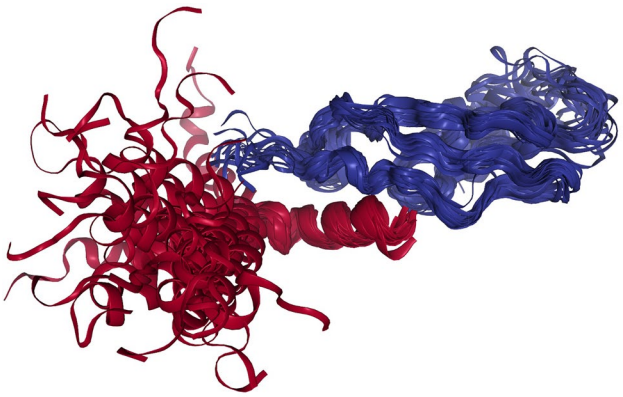


Fig. 5 NMR solution structure of a complex between the N-terminal extracellular domain (ECD) of human CRHR1 (residues 25–108, blue) and a high affinity peptide agonist, α-helical cyclic CRF (red) (PDB ID: 2L27; [126])

Disorder-related information and some D^2P^2 generated illustrative examples of functional disorder profiles for the remaining members of the secretin family (Figures S82–S84) are presented in Supplementary Materials.

Intrinsic disorder and multifunctionality of human G proteins

Activation of 800+ human GPCRs by numerous ligands (> 1000) of different nature (ranging from photons to many hormones and neurotransmitters) represents the very first step in numerous G protein signaling pathways leading to the triggering of the multiple physiological responses to the various external stimuli. At the next step, a cytoplasmic domain of an activated GPCR interacts with one of the intracellularly located guanine nucleotide-binding proteins (G proteins), which are heterotrimers composed of α, β, and γ subunits. This GPCR–G protein interaction increases the GDP to GTP exchange in the Gα subunit, causing Gα activation and dissociation from the Gβγ dimer, therefore generating two functional subunits (Gα-GTP and Gβγ) that can control different cellular pathways [128–132]. So, G proteins play crucial role in transmission of signals from a variety of stimuli outside a cell to its interior, thereby serving as important intracellular molecular switches.

There are four major families of the Gα subunits ($G\alpha_s$, $G\alpha_i$, $G\alpha_q$, and $G\alpha_{12}$) encoded by 16 human genes [3, 5, 6]. Figure 6a represents phylogenetic relationship of human Gα subunits and their expression. $G\alpha_s$ subfamily includes two members, $G\alpha_s$ and $G\alpha_{olf}$, each serving as a subunit of the

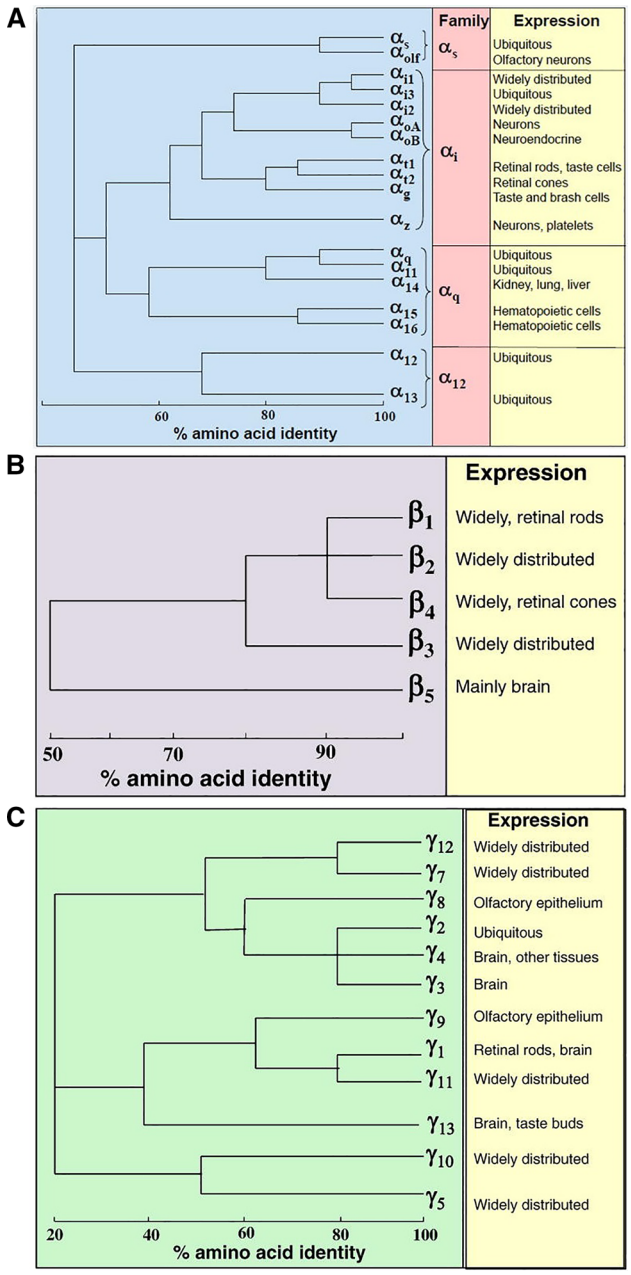


Fig. 6 Phylogenetic relationship of human Gα (**a**), Gβ (**b**), and Gγ subunits and their expression. Reproduced with permission from [131]

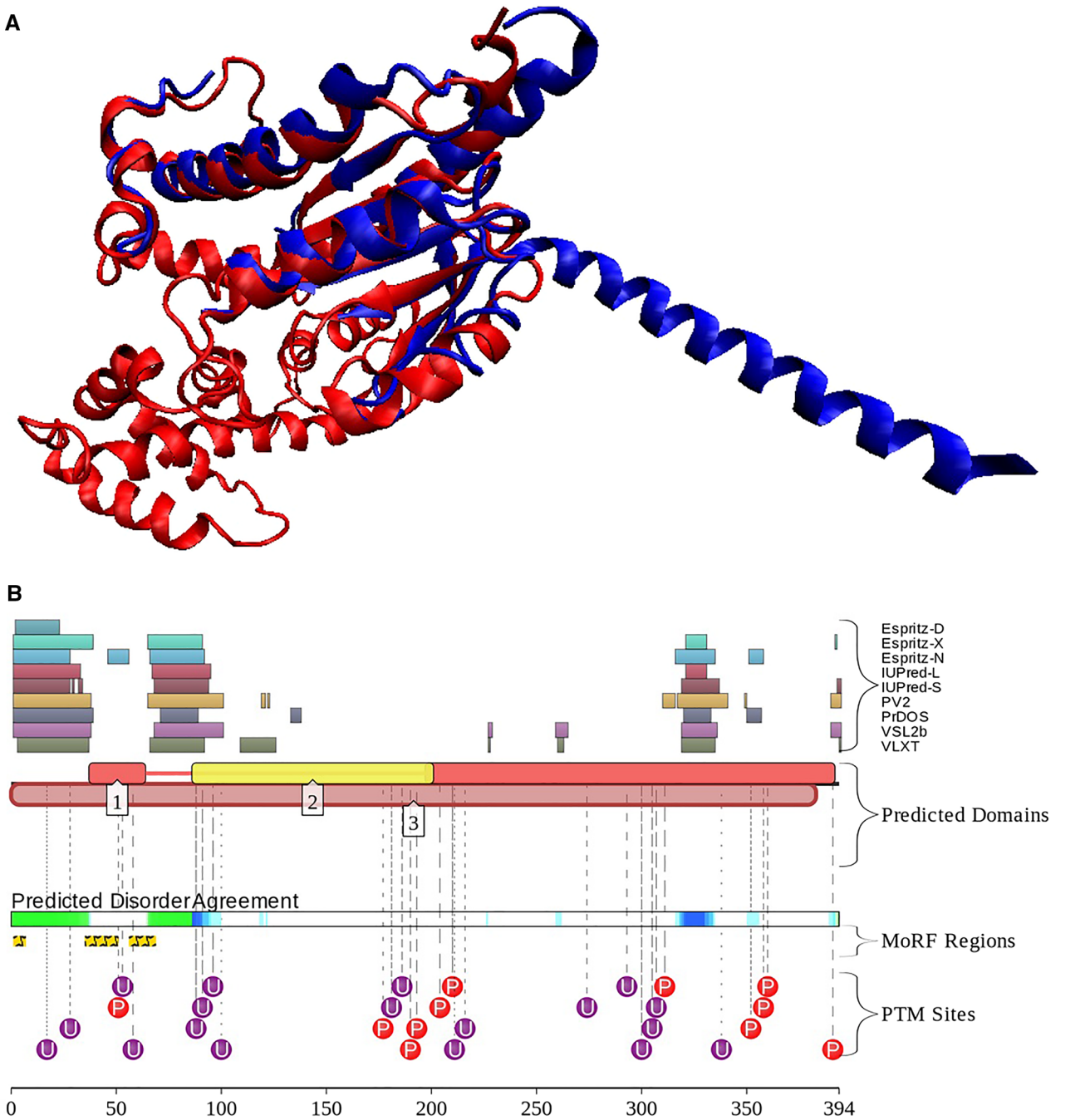
heterotrimeric G protein responsible for the activation of adenylyl cyclases controlling the cAMP-dependent pathway, where the stimulatory $G\alpha_s$ subunit is ubiquitously expressed in most cell types, whereas $G\alpha_{olf}$ is specifically expressed in the olfactory sensory neurons [131]. In the $G\alpha_i$ family, there are seven members. These include: three inhibitory $G\alpha_i$ proteins ($G\alpha_{i1}$, $G\alpha_{i2}$, and $G\alpha_{i3}$ that are expressed in most cell types, where they lead to reduced production of intracellular cAMP); the major Gα subunit of the mammalian brain, $G\alpha_o$, which has two alternatively spliced variants, $G\alpha_{oA}$ and $G\alpha_{oB}$; the $G\alpha_t$ subunit of the transducin that has two isoforms, $G\alpha_{t1}$ and $G\alpha_{t2}$, which are expressed in the rod and cone photoreceptor cells of the eye, respectively; the $G\alpha_g$ subunit of gustducin expressed in the taste receptor cells; and the $G\alpha_z$ subunit of the inhibitory G protein found in neuronal tissues and in platelets. There are four members of the human $G\alpha_q$ family: ubiquitously expressed $G\alpha_q$ and $G\alpha_{11}$; preferentially found in kidney, liver, and lung $G\alpha_{14}$; and specifically expressed in hematopoietic cells $G\alpha_{16}$ (which is equivalent to $G\alpha_{15}$ in mouse). Finally, the $G\alpha_{12}$ family includes ubiquitously expressed subunits $G\alpha_{12}$ and $G\alpha_{13}$ [131].

Although Gα subunits are GTPases that are important for recognition of the activated GPCRs, the ability of G proteins to directly relay the signals from GPCRs to various downstream pathways and function as molecular binary switches with biological activities determined by the bound nucleotide depends on the GPCR-induced dissociation of the heterotrimeric Gαβγ protein into two functional subunits Gα and Gβγ, both transmitting signals to various cellular pathways [128–132]. So, the tightly associated Gβγ heterodimer represents an important regulatory entity. Similar to the Gα subunits encoded by multiple genes, human genome contains 5 Gβ and 12 Gγ genes. Figures 6b, c show phylogenetic relationship of human Gβ and Gγ subunits, respectively, and peculiarities of their expression in different tissues and organs.

Intrinsic disorder in human Gα subunits

Human $G\alpha_s$ family

Human $G\alpha_s$ subunits act downstream of several GPCRs and activate adenylyl cyclases, resulting in increased levels of the second messenger cAMP [133, 134]. They are encoded by the bicistronic gene *GNAS* that in addition to several alternatively spliced isoforms of the $G\alpha_s$ subunit also encodes, from an overlapping reading frame, the ALEX protein (alternative gene product encoded by XL-exon, $N=626$), which inhibits the adenylyl cyclase-stimulating activity of $G\alpha_s$ subunit produced from the same locus in a different open reading frame. Human $G\alpha_s$ is characterized by the presence of several isoforms generated by alternative splining. Although GNAS-2 (UniProt ID: P63092-2, $N=380$, also known as short $G\alpha_s$ subunit) and isoforms 3 (GNAS-3, UniProt ID: P63092-3; $N=379$) and 4 (GNAS-4, UniProt ID: P63092-4; $N=395$) are not too different from the canonical isoform GNAS-1 (UniProt ID: P63092-1; $N=394$, also known as long $G\alpha_s$ subunit), possessing some variability in the 71-86 region, sequences of the extra-large isoforms XLas-1 (UniProt ID: Q5JWF2-1; $N=1037$), XLas-2 (UniProt ID: Q5JWF2-2; $N=1023$), and XLas-3 (UniProt ID: Q5JWF2-3; $N=752$) are significantly divergent from other isoforms, and, as a


A
B
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
1
2
3
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
0
50
100
150
200
250
300
350
394
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] P-loop containing nucleoside triphosphate hydrolases
[2] Transducin (alpha subunit), insertion domain
Pfams:
[3] G-protein alpha subunit

◂**Fig. 7** Analysis of structure and intrinsic disorder predisposition of human $G\alpha_s$ subunit. **a** Multiple structure alignment of free $G\alpha_s$ subunit (red; PDB ID: 6AU6 [136]) and $G\alpha_s$ subunit complexed with the calcitonin receptor, nanobody-35, and the Gβγ dimer (blue; PDB ID: 5UZ7 [122], note that only structure of the $G\alpha_s$ subunit is presented here). Multiple structural alignment was conducted using MultiProt algorithm [225]. **b** Intrinsic disorder propensity and some important disorder-related functional information generated for human $G\alpha_s$ subunit by the D^2P^2 database (http://d2p2.pro/), that uses outputs of several disorder predictors (see differently colored bars at the top of the plot), such as ESpritz_DisProt, ESpritz_X-ray, and ESpritz_NMR (shown as ESpritz-D, ESpritz-X, and ESpritz-N, respectively), IUPred_long and IUPred_short (shown as IUPred-l and IUPred-s, respectively), PV2, PrDOS, PONDR® VSL2B, and PONDR® VLXT. This is complemented by the information on the location of domains predicted by Superfamily and Pfam platforms (http://supfam.org/SUPERFAMILY/ and https://pfam.xfam.org/, respectively). The level of agreement between all of the disorder predictors is shown in the middle of the plot as color intensity in an aligned gradient. The green segments represent disorder that is not found within a predicted domain, whereas the blue segments are where the disorder predictions intersect the domain prediction. Positions of disorder-based interactions sites (MoRFs) and sites of curated posttranslational modifications (phosphorylation and ubiquitination) are shown by yellow blocks with zigzag infill and by red and purple circles, respectively

result, they are annotated in different UniProt entry. Furthermore, the isoform NESP55 (UniProt ID: O95467; $N=245$), which is known as the neuroendocrine secretory protein 55, shares no sequence similarity with other isoforms, being initiated from a promoter ~ 45 kb upstream of *GNAS*, and containing two *NESP*-specific exons that are spliced onto exons 2–12 of *GNAS*, thereby forming an alternatively spliced transcript of *GNAS*, although the NESP55 open reading frame is confined to one upstream exon [135]. The *GNAS* locus is characterized by a very complex imprinting, where different proteins are expressed biallelically, maternally, or paternally. In particular, the GNAS isoforms are expressed biallelically, the NESP55 isoform is maternally derived, and the extra-large XLas isoforms are paternally derived.

Structures of short isoform of human $G\alpha_s$ subunit (GNAS-2, UniProt ID: P63092-2, $N=380$) and several its fragments alone or in complex with GPCR were determined by X-ray crystallography and Volta phase plate electron cryo-microscopy (cryo-EM). Comparison of the structures of free (PDB ID: 6AU6 [136]) and complexed (PDB ID: 5UZ7 [122]) forms of $G\alpha_s$ subunit revealed that interaction of this protein with the calcitonin receptor, nanobody-35, and the Gβγ dimer causes noticeable structural changes in the $G\alpha_s$ subunit, with the most noticeable being the binding-induced formation of the long α-helix in the N-terminus, which was disordered in the unbound state (see Fig. 7a). In fact, although residues 4-35 in the unbound $G\alpha_s$ subunit constitute a region of missing electron density, a large part of this N-terminus possesses an α-helical structure in bound state of this protein. Furthermore, the authors of the cryo-EM study indicated that the α-helical domain (AHD) of the complexed $G\alpha_s$ (residues 62–206) possessed large variability that prevented reconstruction of the density of this region [122]. This form also contained several short regions of missing electron density (residues 251-263, 293-308, 323-330, and 366-369) suggesting the presence of noticeable conformational flexibility in the $G\alpha_s$ subunit, which is preserved even in its bound form. Figure 7b represents a D^2P^2 generated disorder profile of the long isoform of human $G\alpha_s$ subunit and shows that this protein is predicted to have several IDPRs together with three disorder-based binding sites, MoRFs, and is heavily decorated by multiple PTMs. To conclude discussion of the prevalence of intrinsic disorder in proteins encoded by the human *GNAS* gene, Fig. 8 represents disorder profiles of all its proteinaceous products, GNAS-1 (PDR = 25.4%), GNAS-2 (PDR = 20.0%), GNAS-3 (PDR = 19.0%), GNAS-4 (PDR = 25.6%), XLas-1 (PDR = 71.6%), XLas-2 (PDR = 70.2%), XLas-3 (PDR = 94.9% %), NESP55 (PDR = 86.5%), and ALEX (PDR = 99.0%) and clearly shows that all GNAS proteins are expected to contain significant levels of intrinsic disorder; whereas XLas proteins, NESP55, and ALEX are predicted to be mostly disordered. Furthermore, Fig. 8a provides a strong support to the idea that intrinsic disorder and conformational flexibility could be of great importance for the functionality of GNAS-1. In fact, it is clearly seen that the sites known to be involved in the nucleotide binding (residues 47–55, 197–204, 223–227, and 292–295) either coincide or are located in close proximity to the IDPRs or regions with increased flexibility (i.e., regions for which disorder scores exceed 0.25).

The $G\alpha_{olf}$ subunit (GNAL, UniProt ID: P38405; $N=381$) is characterized by a moderate level of predicted disorder (PDR = 19.4%) and contains a highly disordered N-terminus (residues 1–50) that includes two MoRFs (see Figure S85). This subunit has several PTM sites and multiple AS-generated isoforms that preferentially affect the highly disordered N-terminal region.

Human $G\alpha_i$ family

Similar to the other members of the class of Gα proteins, $G\alpha_{i1}$ subunit has two structural domains, a nucleotide-binding domain (the Ras-like domain) with GTPase activity and an α-helical domain (the α-H domain). Since in the nucleotide-bound form, the α-H domain partially occludes the bound nucleotide, at least partial opening of the two domains is required for the release of nucleotides [137]. Another important feature of the $G\alpha_{i1}$ subunit is the presence of the N-terminally located myristoylation site that is required for the association of Ras-like domain with the membrane surface. Activation of Gα proteins is accompanied by the dramatic structural rearrangements of three segments, known as switches I–III (residues 177–187, 199–216, and

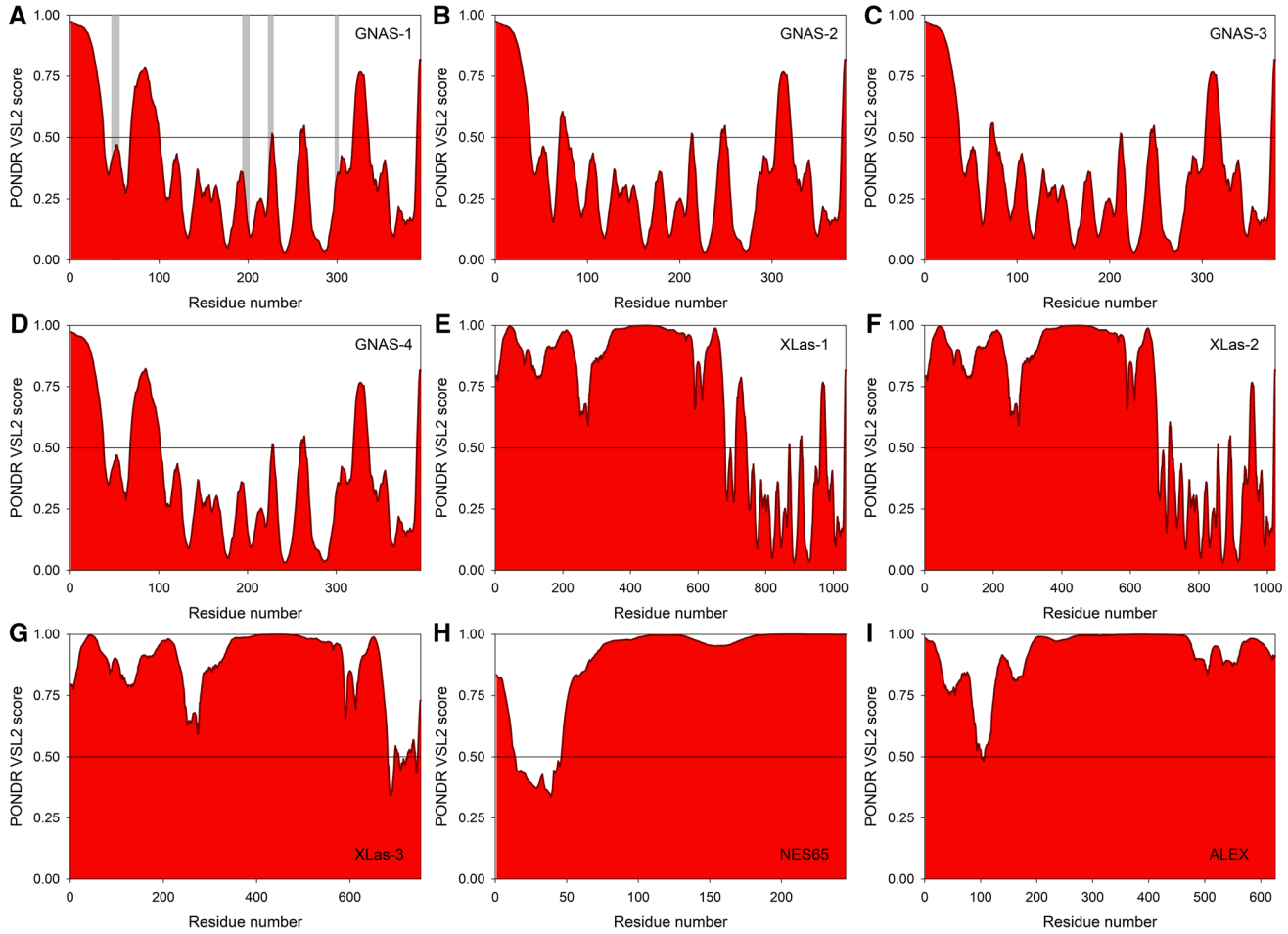


Fig. 8 Analysis of the intrinsic disorder status of various proteins encoded by the human *GNAS* gene. Disorder profiles were generated by PONDR® VSL2 for: **a** GNAS-1 (UniProt ID: P63092-1; $N=394$); **b** GNAS-2 (UniProt ID: P63092-2, $N=380$); **c** GNAS-3 (UniProt ID: P63092-3; $N=379$); **d** GNAS-4 (UniProt ID: P63092-4; $N=395$); **e** XLas-1 (UniProt ID: Q5JWF2-1; $N=1037$); **f** XLas-2 (UniProt ID: Q5JWF2-2; $N=1023$); **g** XLas-3 (UniProt ID: Q5JWF2-3; $N=752$); **h** NESP55 (UniProt ID: O95467; $N=245$); and **i** ALEX (UniProt ID: P84996; $N=626$). In these analyses, disorder scores above the 0.5 threshold correspond to predicted intrinsically disordered residues and regions. In plot **a**, locations of known nucleotide binding sites are shown by gray vertical bars

231–242, respectively), which are loops between α-helices whose conformations are sensitive to guanine nucleotides [137, 138]. It was pointed out that switch I undergoes the GTPγS-dependent conformational change. On the other hand, switches II and III are disordered in the GDP-bound form of the protein [139], but undergo disorder–order transition upon GTPγS binding [140]. NMR analysis of the solution structure of this protein in its apo- and nucleotide-bound forms revealed that the apo- and GDP-bound forms of the $G\alpha_{i1}$ subunit are characterized by rather open and dynamic structure, whereas the $G\alpha_{i1}$ subunit with bound GTP analogs possesses a rigid and closed arrangement of the Gα subdomain [137]. Human inhibitory $G\alpha_{i1}$ subunit (GNAI1, UniProt ID: P63096; $N=354$; PDR = 24.6%) has a long N-terminally located IDPR and contains numerous PTM sites (see Figure S86). Furthermore, the switches I–III with the nucleotide-dependent structures were predicted to be either flexible (switch I) or intrinsically disordered (switches II and III). Finally, human $G\alpha_{i1}$ subunit has an alternatively spliced isoform-2 with missing residues 1-52 that correspond to the long N-terminal IDPR of the canonical form of this protein.

Disorder profiles of human $G\alpha_{i2}$ (GNAI2, UniProt ID: P04899; $N=355$; PDR = 24.5%) and $G\alpha_{i3}$ proteins (GNAI3, UniProt ID: P08754; N = 354; PDR = 24.9%) are remarkably similar to that of the $G\alpha_{i1}$ subunit, with each of these two proteins possessing a long N-terminal IDPRs, multiple PTMs sites (see Figures S87 and S88), and a multitude of AS isoforms with missing or changed N-terminal regions or with modified C-tails. Same is also applicable to the $G\alpha_{o}$ protein (GNAO, UniProt ID: P09471-1; $N=354$; PDR = 24.9%), which is the major Gα subunit found in the human brain, and its alternatively spliced isoform GNAO Alpha-2 (UniProt ID: P09471-2; $N=354$; PDR = 24.9%) (see Figures S89 and S90) that is different from the canonical

form of this protein by changed C-terminal region (residues 249-354). Similarly, $G\alpha_{t1}$ (GNAT-1, UniProt ID: P11488; $N = 350$; PDR = 23.7%) and $G\alpha_{t2}$ (GNAT-2, UniProt ID: P19087; $N = 354$; PDR = 22.9%) subunits of transducins expressed in the rod and cone cells of the eye, respectively, as well as the $G\alpha_g$ subunit of gustducin (GNAT3, UniProt ID: A8MTJ3; $N = 354$; PDR = 21.2%) and the $G\alpha_z$ subunit (GNAZ; UniProt ID: P19086; $N = 355$; PDR = 23.7%) also have disorder profiles relatively close to those of other members of this family (see Figures S91–S94). This similarity of disorder profiles between the members of human $G\alpha_i$ family is not surprising, since these proteins show relatively high sequence identities that range from 54.29 to 94.35% [131].

Human $G\alpha_q$ family

In terms of the distribution of intrinsic disorder propensity within the amino acid sequences of the members of human $G\alpha_q$ family, $G\alpha_q$ (GNAQ, UniProt ID: P50148; $N = 359$; PDR = 25.6%), $G\alpha_{11}$ (GNA11, UniProt ID: P29992; $N = 359$; PDR = 24.2%), $G\alpha_{14}$ (GNA14, UniProt ID: O95837; $N = 355$; PDR = 23.1%), and $G\alpha_{16}$ (GNA16, UniProt ID: P30679; $N = 374$; PDR = 23.3%), are similar to each other (see Figures S95–98) and are not very different from the disorder profiles of the members of $G\alpha_s$ and $G\alpha_i$ families. In fact, all these Gα proteins have a long N-terminally located IDPR, several short IDPRs, and numerous sites of various PTMs spread over their sequences.

Human $G\alpha_{12}$ family

The members of $G\alpha_{12}$ family $G\alpha_{12}$ (GNA12, UniProt ID: Q03113; $N = 381$; PDR = 22.8%) and $G\alpha_{13}$ (GNA13, UniProt ID: Q14344; $N = 377$; PDR = 15.9%) are different from most other human Gα proteins (with the exception for $G\alpha_{16}$ and several AS isoforms of $G\alpha_{s1}$) by possessing longer amino acid sequences. As a result, in addition to the long N-terminally located IDPRs, these Gα subunits are characterized by the presence of rather disordered C-tails (see Figures S99 and S100).

Functionality and intrinsic disorder of human Gβ subunits

The Gβγ functional subunit represents a tightly associated heterodimer, where the structure of the Gβ represents a β-propeller with 7 blades, each consisting of four-stranded β-sheets, and where the N-terminally located α-helical segment is engaged in the formation of the coiled-coil with the Gγ subunit (see Fig. 9a). The ability of the Gβ subunits to form a seven-bladed β-propeller structure is determined by the presence in these proteins of 7 WD motifs (also known as β-transducin repeats), which are short (~40-residue-long) structural motifs often containing a tryptophan–aspartic acid (W-D) dipeptide at their C-termini, and which, being present as tandem copies (typically between 5 and 7) can fold into a circular or closed solenoid structure [141].

Figure 6b shows that sequences of $G\beta_1$ (GNB1, UniProt ID: P62873; $N = 340$; PDR = 10.3%), $G\beta_2$ (GNB2, UniProt ID: P62879; $N = 340$; PDR = 11.8%), $G\beta_3$ (GNB3, UniProt ID: P16520; $N = 340$; PDR = 10.3%), and $G\beta_4$ (GNB4, UniProt ID: Q9HAV0; $N = 340$; PDR = 11.8%) subunits are characterized by high sequence identity, which ranges from 80 to 90%, whereas Gβ5 (GNB5, UniProt ID: O14775; N = 395; PDR = 16.7%) is different from other Gβ subunits, sharing with them ~ 50% identical residues. In agreement with this high sequence similarity, all Gβ subunits are characterized by rather similar disorder profiles (see Fig. 9b). Importantly, the N-terminal regions of all Gβ subunits involved in the formation of α-helical coiled-coil structures (which are structural elements built by two or more α-helices that wind around each other to form a supercoil) with the Gγ subunits are predicted to be mostly disordered, indicating that these regions undergo disorder-to-order transitions at the formation of the tightly bound Gβγ subunits. This observation is in line with the important notion that many proteins involved in the formation of actin filament, cell junction, ciliary rootlet, lamellipodium, nucleoli, and various macromolecular complexes frequently contain disordered regions that can form coiled-coils [19, 63, 142, 143].

Intrinsic disorder and functionality of human Gγ subunits

As it follows from Fig. 9a, in its Gβ-bound form, human Gγ subunit possesses a highly extended α-helical structure, where the entire sequence is engaged in the coiled-coil formation at interaction with the N-terminal tail of the Gβ subunit. If one looks at the structure of the bound form of Gγ subunit computationally taken out of the complex context, it become evident that such a structure cannot exist in the unbound form, since there are no interhelical contents in this structure (see Fig. 9c), clearly indicating that such extended α-helical structure is supported by the set of specific interaction with the corresponding helical segments of the Gβ subunit. This is in line with the results of a comprehensive analyses of ordered and disordered protein complexes, which are either formed via interaction of ordered proteins (ordered oligomers) or represent a result of the concomitant folding and binding events (disordered oligomers), which revealed that the per-residue surface and interface areas are significantly greater in disordered oligomers than in ordered oligomers [144, 145].

Figure 6c illustrates that there is noticeable sequence variability among human Gγ subunits that share 20% to 80% identical residues. In line with these considerations

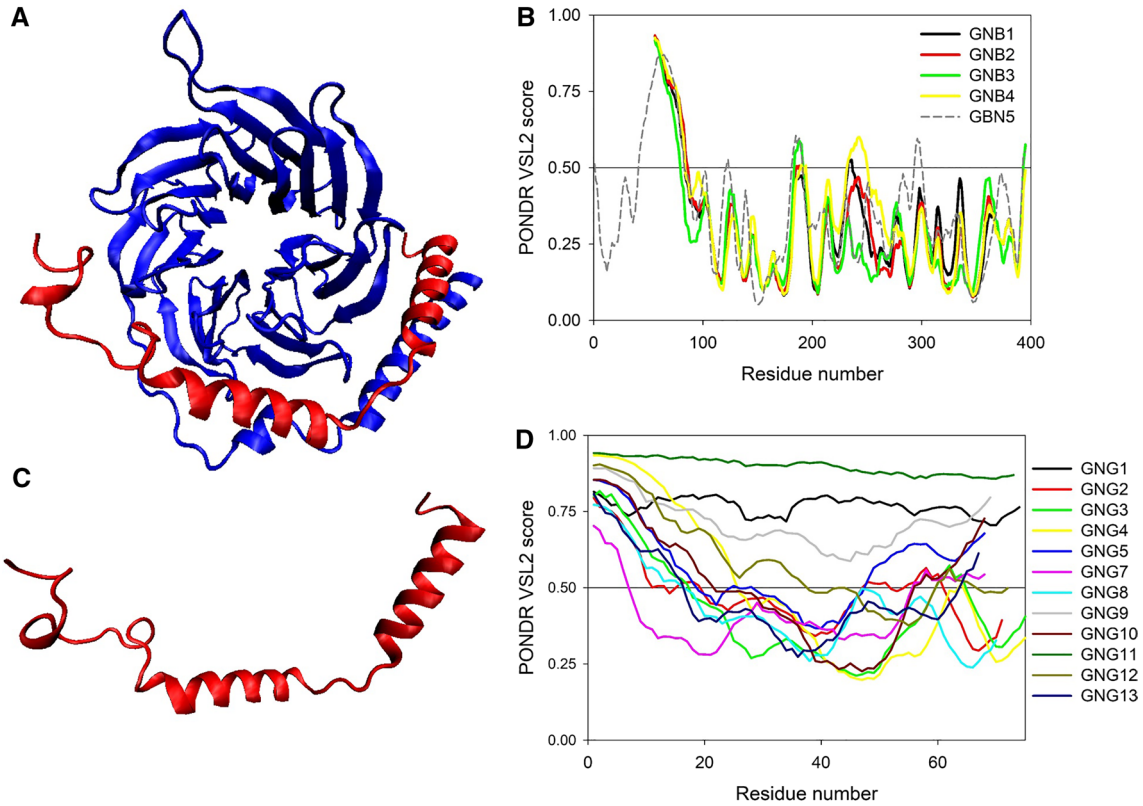


Fig. 9 **a** Crystal structure of the tightly associated Gβγ heterodimer (PDB ID: 5HE0; [226]). Here, structures of the Gβ and Gγ subunits are shown by blue and red color. Note that although the reported 5HE0 structure included a complex between human $G\beta_1\gamma_2$ heterodimer and bovine β-adrenergic receptor kinase 1, only structure of the $G\beta_1\gamma_2$ heterodimer is shown here for simplicity. **b** Comparison of the intrinsic disorder propensities of human Gβ subunits: $G\beta_1$ (GNB1, UniProt ID: P62873; $N=340$; PDR = 10.3%), $G\beta_2$ (GNB2, UniProt ID: P62879; $N=340$; PDR = 11.8%), $G\beta_3$ (GNB3, UniProt ID: P16520; $N=340$; PDR = 10.3%), $G\beta_4$ (GNB4, UniProt ID: Q9HAV0; $N=340$; PDR = 11.8%), and Gβ5 (GNB5, UniProt ID: O14775; $N=395$; PDR = 16.7%). **c** Crystal structure of the Gβ-bound form of the Gγ subunit computationally taken out of the complex context (PDB ID: 5HE0; [226]). **d** Comparison of the intrinsic disorder propensities of human Gγ subunits: $G\gamma_1$ (GNG1/GNGT1, UniProt ID: P63211; $N=74$; PDR = 100.0%), $G\gamma_2$ (GNG2, UniProt ID: P59768; $N=71$; PDR = 38.3%), $G\gamma_3$ (GNG3, UniProt ID: P63215; $N=75$; PDR = 29.3%), $G\gamma_4$ (GNG4, UniProt ID: P50150; $N=75$; PDR = 34.7%), $G\gamma_5$ (GNG5, UniProt ID: P63218; $N=68$; PDR = 60.3%), $G\gamma_7$ (GNG7, UniProt ID: O60262; $N=68$; PDR = 27.9%), $G\gamma_8$ (GNG8, UniProt ID: Q9UK08; $N=70$; PDR = 22.9%), $G\gamma_9$ (GNG9/GNGT2, UniProt ID: O14610; $N=69$; PDR = 100.0%), $G\gamma_{10}$ (GNG10, UniProt ID: P50151; $N=68$; PDR = 50.0%), $G\gamma_{11}$ (GNG11, UniProt ID: P61952; $N=73$; PDR = 100.0%), $G\gamma_{12}$ (GNG12, UniProt ID: Q9UBI6; $N=72$; PDR = 61.1%), and $G\gamma_{13}$ (GNG13, UniProt ID: Q9UBI6; $N=67$; PDR = 29.9%)

related to the correlation between the intrinsic disorder status of a protein or protein region and its ability to form coiled-coil, all human Gγ subunits, such as $G\gamma_1$ (GNG1/GNGT1, UniProt ID: P63211; $N=74$; PDR = 100.0%), $G\gamma_2$ (GNG2, UniProt ID: P59768; $N=71$; PDR = 38.3%), $G\gamma_3$ (GNG3, UniProt ID: P63215; $N=75$; PDR = 29.3%), $G\gamma_4$ (GNG4, UniProt ID: P50150; $N=75$; PDR = 34.7%), $G\gamma_5$ (GNG5, UniProt ID: P63218; $N=68$; PDR = 60.3%), $G\gamma_7$ (GNG7, UniProt ID: O60262; $N=68$; PDR = 27.9%), $G\gamma_8$ (GNG8, UniProt ID: Q9UK08; $N=70$; PDR = 22.9%), $G\gamma_9$ (GNG9/GNGT2, UniProt ID: O14610; $N=69$; PDR = 100.0%), $G\gamma_{10}$ (GNG10, UniProt ID: P50151; $N=68$; PDR = 50.0%), $G\gamma_{11}$ (GNG11, UniProt ID: P61952; $N=73$; PDR = 100.0%), $G\gamma_{12}$ (GNG12, UniProt ID: Q9UBI6; $N=72$; PDR = 61.1%), and $G\gamma_{13}$ (GNG13, UniProt ID: Q9UBI6; N = 67; PDR = 29.9%), are expected to be highly disordered (see Fig. 9d).

Experimental evidence of the structural and functional flexibility of GPCRs and G proteins

Multifunctionality and structural polymorphism of GPCRs

Results of our computational analyses of the intrinsic disorder predisposition of GPCRs and G proteins are supported by several experimental studies, where the presence of noticeable structural flexibility was reported for several

components of this important machinery. For example, it was postulated that the ability of GPCRs to interact with several intracellular proteins, such as arrestins, G proteins, and kinases, and to be activated by signals of very different nature requires high structural plasticity, suggesting that structure of these proteins represents dynamic conformational ensembles, with multiple functional conformations providing means for the conformation-dependent interactions of GPCRs with different partners [146–149].

In particular, based on the structural analysis of the light-activated rhodopsin in a lipid environment (using phospholipid nanodiscs, in which the receptor has properties similar to those in native membranes), it was concluded that this receptor existed in the manifold of conformations, and distribution of different forms within this conformational ensemble showed strong pH dependence, suggesting that these conformations were in equilibrium [147]. In this study, the site-directed spin labeling (SDSL) and double electron–electron resonance (DEER) approach were utilized to investigate the positions of TM5, TM6, and TM7 helices of the rhodopsin by distance mapping of pairs of spin labels located on the outer surface of the target helices. This analysis revealed that in a native-like phospholipid environment, activated rhodopsin existed as manifold of conformations, where there is an equilibrium between comparable populations of multiple conformations defined by the positions of TM6 and TM7 at the cytoplasmic surface [147]. Furthermore, this conformational ensemble of the nanodisk-embedded activated rhodopsin was able to spontaneously relax to the conformation typical for the inactivated receptor [147]. Importantly, although G protein apparently was able to predominantly select one of the activated rhodopsin conformations, the helices remained in exchange between more than one states, indicating the presence of significant conformational dynamics in the rhodopsin, even in its G protein-bound state [147]. Similarly, utilization of the NMR spectroscopy of $^{13}CH_3$-ε-methionines combined with DEER spectroscopy revealed the presence of significant conformational heterogeneity in the transmembrane core of the β_2-adrenergic receptor (ADRB2, residues 1–365), and these conformational flexibility was at least partially preserved when this receptor was bound to agonist or inverse agonist [148–150].

It is important to emphasize here that the aforementioned manifolds of rhodopsin and ADRB2 conformations reflect the structural flexibility of the transmembrane 7TM domain of GPCR. Currently available information on the flexibility and conformational dynamics of the transmembrane domains of GPCRs is described in great detail in a recent review [9]. Particularly, it was pointed out that due to their noticeable structural dynamics, structure of GPCRs can be described as a conformational ensemble containing almost infinite number of conformations that tend to group into clusters known as conformational states, with rapid structural exchange among conformations within a conformational state and slower exchange between the conformational states [9]. It was also indicated that although the intracellular side of the GPCR undergoes the largest conformational changes, the extracellular part of the receptor also possesses noticeable conformational dynamics and this conformational flexibility contributes to the efficiency and specificity of ligand binding [9]. The presence of multiple conformational states in GPCRs is related to their biased signaling, also known as functional selectivity, where the selection of which of many possible signaling pathways will be stimulated by a given GPCR is determined by a ligand it binds [9, 151, 152]. This concept of biased signaling suggests that GPCRs do not act as simple on–off switches controlled by a ligand (i.e., do not have just two "inactive" and "active" forms with transition between these two forms controlled by the extracellular ligand binding). Instead, ligands can select among numerous GPCR conformational states possessing abilities to interact with different intracellular binding partners thereby stimulating different signaling pathways [153–165]. Furthermore, in addition to the 7TM domain invariantly present in all GPCRs, many classes of these receptors contain various functional domains that can be rather long and that are often expected to contain high levels of intrinsic disorder.

It is also important to remember that GPCRs initiate and coordinate multiple highly conserved signaling and regulatory pathways by interactions with heterotrimeric G proteins that stimulate generation of second messengers, G protein-coupled receptor kinases (GRKs) phosphorylating the activated receptors, and arrestins that bind to the phosphorylated receptors and mitigate signaling [166]. However, in addition to this G protein/GRK/arrestin triad GPCRs can interact with multiple other proteins. Some of these interactions lead to the formation of stable multiprotein complexes with the receptor [167, 168]. In fact, in several cases, it was shown that the GPCR activity requires interaction with specific non-triad proteins. For example, functioning of the calcitonin receptor-like receptor (CRLR) and calcitonin receptor relies on the GPCR interaction with the receptor activity-modifying proteins (RAMPs) and the receptor component protein (RCP) [169, 170]. In fact, the CRLR–RAMP1 complex acts as a receptor for a pleiotropic family of the calcitonin-related peptides; i.e., neuropeptides with homology to calcitonin, amylin, and adrenomedullin. On the other hand, the CRLR–RAMP2–RCP complex functions as an adrenomedullin receptor, and calcitonin receptor acts as an amylin receptor when naturally occurring splice variant of the calcitonin receptor forms complexes with RAMP1 or RAMP3 [168–170].

Intracellular domains of several GPCRs were shown to interact with some non-G protein effectors, thereby opening a possibility for the non-G protein-mediated GPCR

signaling [171–174]. Among such non-G protein effectors that might function as alternative GPCR signal transducers are guanine nucleotide exchange factors (GEFs), non-receptor tyrosine kinases, and several proteins acting as adaptors or scaffolds [168, 175]. For example, interaction of the PDZ domain (postsynaptic density protein of 95 kDa/disc-large/zona occludens-1) of the cAMP-regulated Ras GEF with a specific motif located within the C-terminus of the β_1-adrenergic receptor defines the ability of this GPCR to stimulate guanine nucleotide exchange and activation of the protein Ras [176]. C-termini of the β_2-adrenergic and parathyroid hormone/parathyroid hormone-related protein receptors contain specific motifs which are recognized by the PDZ domain-containing protein, the Na^+/H^+ exchanger regulatory factor/ezrin binding protein 50 (NHERF-1), thereby regulating activity of the Na^+/H^+ exchanger 3 [177, 178]. Other PDZ domain-containing proteins can play a role in enhancing various signaling pathways via interaction with specific GPCRs. The examples of such PDZ scaffold regulators include the multi-PDZ protein Mupp1 that noticeably boosts the $G\alpha_i$-mediated signaling by interaction with the stimulated $GABA_B$ receptors [179] and MT_1 melatonin receptors [180], the PDZ scaffold MAGI-3 that augments the receptor-mediated activation of ERK/mitogen-activated protein (MAP) kinase pathways by interaction with Frizzled-4 [181] and LPA_2 receptor [182], and the PDZ scaffolds PDZ-RhoGEF and LARG that played a role in stimulation of Rho signaling resulted in the modification of the cytoskeleton dynamics via interaction with the LPA_2 receptor [183]. In addition to PDZ domain-containing proteins, GPCRs can interact with several proteins containing Src homology 2 and 3 (SH2− and SH3−) domains [184, 185].

Also, a number of GPCRs (such as AT1 angiotensin receptor [186, 187] and the platelet activating factor receptor [188, 189]) can initiate cellular signaling through agonist-promoted interactions with non-receptor protein tyrosine kinases [e.g., members of the Janus kinase (Jak) family and tyrosine kinase 2 (Tyk2)] [175]. The efficiency of G protein-mediated signaling can be also modulated by other GPCR-associated proteins, such as members of the A-kinase anchoring protein (AKAP) family and Homer proteins, interacting, respectively, with the β-adrenergic receptors [190–194] and the metabotropic glutamate receptors mGluR1 and mGluR5 [195–198]. Because of their capability to interact with the proline-rich motifs of mGluRs and also bind to intracellular IP_3 receptors, Homer proteins increase the efficiency of the mGluR-stimulated calcium signaling by linking mGluRs, intracellular IP_3 receptors, and other cellular components [197, 199, 200]. Similarly, being able to interact with GPCRs and protein kinase A (PKA), AKAPs (such as AKAP79 and AKAP250) tether PKA in the close proximity of the β-adrenergic receptors thereby modulating the efficacy of the PKA-mediated phosphorylation of different downstream substrates [190–193]. Curiously, AKAP250 binding to β_2-adrenergic receptor involves cytoplasmic Arg-329 to Leu-413 domain of this GPCR that contains a multitude of phosphorylation sites and is predicted to be highly disordered.

Furthermore, GPCR can also interact with several integral transmembrane proteins, such as other GPCRs, ion channels, transporters, and receptors from other families [201–203]. For example, melatonin receptor MT1 was shown to physically interact with the voltage-gated calcium channel Cav 2.2 and this interaction resulted in inhibition of the Cav 2.2-promoted Ca^{2+} entry in an agonist-independent manner [202]. More generally, MT1 receptor was shown to serve as an important integral part of several presynaptic protein complexes in neurons [202]. Finally, the functionality of GPCRs can be regulated not only by intracellular or transmembrane proteins, but also by some extracellular proteins. An illustrative example of such allosteric regulators of the GPCR activity is given by the extracellular leucine-rich repeat fibronectin type III domain containing 1 (ELFN1) protein that was shown to bind exclusively to the members of the group III mGluRs (mGluR4, mGluR6, mGluR7, and mGluR8) and allosterically modulate the ability of these receptors to suppress the cAMP accumulation by altering the ability of mGluRs to activate G proteins [204].

Obviously, considered in the previous paragraphs, illustrative examples of the non-triad proteins involved in interaction and regulation of GPCRs represent a very top of an iceberg, since many other intracellular, transmembrane, and extracellular proteins might have a possibility to be engaged in binding to GPCRs. Recently, to get a better picture of the breadth of the GPCR interactome reflecting the involvement of various proteins in GPCR-mediated signaling, a modified membrane yeast two-hybrid (MYTH) approach (which is specifically designed for the identification of proteins interacting with the full-length integral membrane proteins [205–210]) was used to identify interacting partners for a set of 48 selected human inactivated full-length GPCRs [201]. Analyzed GPCRs included 44 rhodopsin-like receptors, 2 secretin-like receptors (vasoactive intestinal peptide receptor 2 and retinoic acid-induced gene 2 protein), and 2 Frizzled/taste2 receptors (smoothened and Frizzled-7). This analysis revealed that the GPCR interactome of the 48 selected full-length human GPCRs included 686 proteins (including 299 membrane proteins) connected by 987 unique interactions [201].

Multifunctionality and structural polymorphism of G proteins

The ability to interact with multiple partners (not only numerous GPCRs and somewhat less numerous effectors) was also described for G proteins. Since in humans there

are 23 Gα, 6 Gβ, and 12 Gγ subunits that can be assembled into numerous different heterotrimers [211], since there are 850+ different GPCRs [1–4] and ~ 40 of G protein effectors, modulators, and scaffold proteins collectively known as regulator of G signaling (RGS) proteins that are able to interact with various Gαβγ heterotrimers or their dissociated subunits [212], and since a greatly simplified G protein signaling pathway is present in plants [213], researchers typically utilize *Arabidopsis thaliana* as a model organism to analyze functionality of G proteins [213]. In fact, the heterotrimeric G protein complex in *Arabidopsis* includes one canonical Gα subunit (AtGPA1), one Gβ subunit (AGB1), one of three Gγ subunits (AGG1, 2, and 3), at least one subunit of Regulator of G Signaling protein (AtRGS1), and one of three atypical Extra-Large G proteins (XLG1, 2, and 3) [214–220]. Utilization of the comprehensive high-throughput yeast two-hybrid (Y2H) screening of several prey libraries made from diverse *Arabidopsis* tissues using seven primary baits, ranging from the Gα (GPA1), Gβ/Gγ1 (AGB1/AGG1), and Gβ/Gγ2 (AGB1/AGG2) subunits of the heterotrimeric G proteins, the N-myc downregulated-like1 (NDL1), Pirin (PRN), receptor for activated C kinase1A (RACK1A), and AtRGS1 revealed the existence of a well-developed and highly connected G protein interactome [221]. This interactome included 434 proteins involved in 544 interactions, with 68 proteins being engaged in 167 interactions, thereby forming a highly interconnected core of this G protein interactome with an average node degree of 4.1, and with many proteins identified as interaction partners of two or more of the bait proteins [221]. Based on the functional analysis of a set of core G protein interactome proteins, a novel role for the G proteins in regulating cell wall modification was proposed [221]. Recently, the analysis of the *Arabidopsis* G protein interactome was extended by assembling a set of XLG-interacting proteins using the full-length XLG proteins as baits in the yeast two-hybrid complementation-based screen of glucose-treated seedlings, roots, and *Arabidopsis* cells in culture [220]. This analysis showed that *Arabidopsis* contains at least 72 XLG-interacting proteins with various biological functions, including developmental processes, cell organization and biogenesis, as well as control of responses to stress and abiotic stimuli, transcriptional regulation, and guidance of cell trafficking, including partner-dependent subcellular localization of the XLG proteins in the nucleus, endosome, and plasma membrane [220].

Structure–function continuum of human GPCRs and G proteins in light of the intrinsic disorder-based multifunctionality and proteoforms

Although the data assembled in this article seem to be excessive, they generate a very clear picture on the importance of intrinsic disorder for the functionality of human GPCRs and G proteins. We show here that the vast majority of members of these two important groups of proteins possess noticeable amounts of intrinsic disorder, which typically ranges from 15 to 30%, but can reach 100% for some Gγ subunits. In GPCRs, intrinsically disordered regions are most commonly found in the extracellular N-terminal and intracellular C-terminal regions encompassing 7TM domain. Several GPCR families use long N-terminal regions that frequently contain high levels of intrinsic disorder for recognition of various ligands. Intrinsically disordered intracellular C-tails are commonly utilized for interaction with G proteins and arrestins. Ligand binding and interaction with G proteins and other effectors are frequently associated with structural rearrangements of the corresponding regions and domains of GPCRs. All GPCRs have numerous post-translational modifications utilized for the regulation of their activity. Furthermore, due to the extensive alternative splicing, vast majority of GPCRs are synthesized as multiple isoforms. Similar observations are also applicable to human G proteins, which contain functional IDPRs, multiple PTMs, and multitude of AS-generated isoforms. An example of the crucial importance of alternative splicing for structure and functionality of G proteins is given by $G\alpha_s$ subunit. Figure 8 shows that human bicistronic *GNAS* gene encodes multiple proteins generated by AS, or that are characterized by very different levels of intrinsic disorder (GNAS-1 (PDR = 25.4%), GNAS-2 (PDR = 20.0%), GNAS-3 (PDR = 19.0%), GNAS-4 (PDR = 25.6%), XLas-1 (PDR = 71.6%), XLas-2 (PDR = 70.2%), XLas-3 (PDR = 94.9% %), NESP55 (PDR = 86.5%), and ALEX (PDR = 99.0%). Furthermore, although GNAS and XLas isoforms all act as activators of adenylyl cyclases, NESP55 protein is a chromogranin-like polypeptide associated with the constitutive secretory pathway [222, 223], and ALEX, which is a completely disordered protein produced from the same *GNAS* locus as GNAS and XLas isoforms, but in a different open reading frame, acts as an inhibitor of the adenylyl cyclase-stimulating activity of the $G\alpha_s$ subunit [224].

Finally, one should keep in mind that 800+ human GPCRs that are capable of recognition of thousands of various external signals and that transmit the related information to trigger distinct signaling cascades inside the

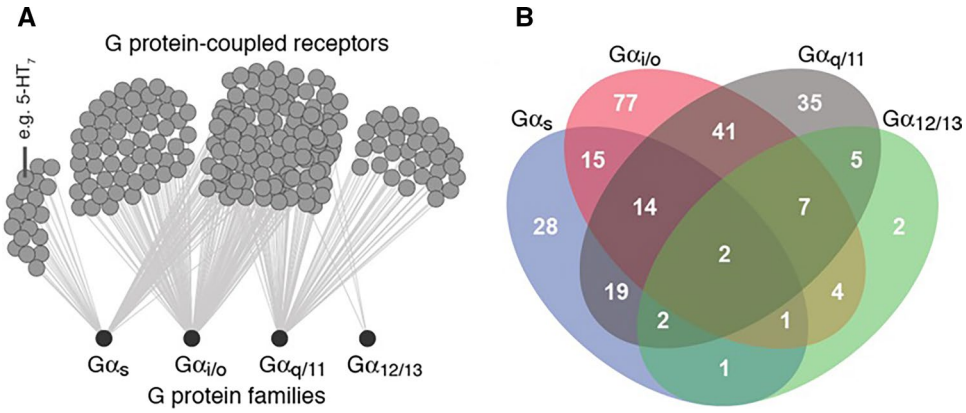


Fig. 10 Complex interactivity of human GPCRs and G proteins. **a** Network representation of the currently available G protein coupling data. **b** Numbers of receptors coupling to different G proteins or their sets. Reproduced with permission from [5]

cells, do so via the precise and specific coupling to cytosolic adaptor proteins, G proteins. G proteins themselves have heterotrimeric Gαβγ structure, with Gα, Gβ, and Gγ subunits being encoded by 16 (assembled in four families), 5, and 12 genes, respectively, and with different combinations of these Gα, Gβ, and Gγ subunits being assembled into the heterotrimers. The general mechanism of the activation of G protein-associated pathways is described as a set of consecutive events, where receptor is first activated by binding of the external ligand. This results in structural changes in the receptor that lead to the recruitment of a heterotrimeric Gαβγ, followed by the nucleotide exchange in Gα, and finally resulting in the dissociation of the Gαβγ protein into Gα and G βγ subunits. However, combinatorics of this process is very large, since one receptor can interact with multiple Gα proteins and multiple receptors can couple to the same Gα protein [5]. This point is illustrated by Fig. 10, where a network of the currently available G protein coupling data is represented together with the diagram showing the numbers of GPCRs which are known to couple to different G proteins or their sets [5].

Based on all these observations, one can conclude that human GPCRs and G proteins represent dynamic conformational ensembles containing multiple conformational or basic proteoforms originated due to the presence of multiple IDPRs. Because GPCRs and G proteins contain multiple PTMs, and since a multitude of their isoforms is generated via the alternative splicing or due to the use the alternative promoters, these proteins also have inducible or modified proteoforms. Finally, because of the conformational ensembles of GPCRs and G proteins might undergo dramatic changes due to their interaction with specific binding partners (as a matter of fact, the entire mode of action of these proteins is based on the recognition of a signal followed by conformational change needed for recognition of another partner that is crucial for the downstream transmission of the signal) these proteins also contain functioning proteoforms. Therefore, multifunctionality of GPCRs and G proteins, which is needed for them to recognize a wide variety of extracellular signals and to transmit this information for initiation of a multitude of cellular pathways, is determined by the presence of all the proteoforms. Thus, this intrinsic disorder and proteoform-based multifunctionality represents an important illustration of the structure–function continuum concept applied to cellular signaling.

Acknowledgements This work was supported in part by the President of the Russian Federation Scholarship SP-3665.2018.4 (A.V.F.), grants from Russian Science Foundation RSCF 18-75-10115 (A.V.F.) and RSCF 19-15-00107 (K.K.T.).

References

1. Bjarnadottir TK, Gloriam DE, Hellstrand SH, Kristiansson H, Fredriksson R, Schioth HB (2006) Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. Genomics 88(3):263–273. https://doi.org/10.1016/j.ygeno.2006.04.001
2. Anantharaman V, Abhiman S, de Souza RF, Aravind L (2011) Comparative genomics uncovers novel structural and functional features of the heterotrimeric GTPase signaling system. Gene 475(2):63–78. https://doi.org/10.1016/j.gene.2010.12.001
3. Southan C, Sharman JL, Benson HE, Faccenda E, Pawson AJ, Alexander SP, Buneman OP, Davenport AP, McGrath JC, Peters JA, Spedding M, Catterall WA, Fabbro D, Davies JA, Nc I (2016) The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. Nucleic Acids Res 44(D1):D1054–D1068. https://doi.org/10.1093/nar/gkv1037
4. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 63(6):1256–1272. https://doi.org/10.1124/mol.63.6.1256
5. Flock T, Hauser AS, Lund N, Gloriam DE, Balaji S, Babu MM (2017) Selectivity determinants of GPCR–G-protein binding. Nature 545(7654):317–322. https://doi.org/10.1038/nature22070
6. Isberg V, de Graaf C, Bortolato A, Cherezov V, Katritch V, Marshall FH, Mordalski S, Pin JP, Stevens RC, Vriend G, Gloriam DE (2015) Generic GPCR residue numbers—aligning topology maps while minding the gaps. Trends Pharmacol Sci 36(1):22–31. https://doi.org/10.1016/j.tips.2014.11.001
7. Neves SR, Ram PT, Iyengar R (2002) G protein pathways. Science 296(5573):1636–1639. https://doi.org/10.1126/science.1071550

8. Marinissen MJ, Gutkind JS (2001) G-protein-coupled receptors and signaling networks: emerging paradigms. Trends Pharmacol Sci 22(7):368–376. https://doi.org/10.1016/S0165-6147(00)01678-3
9. Latorraca NR, Venkatakrishnan AJ, Dror RO (2017) GPCR dynamics: structures in motion. Chem Rev 117(1):139–155. https://doi.org/10.1021/acs.chemrev.6b00177
10. Flower DR (1999) Modelling G-protein-coupled receptors for drug design. Biochim Biophys Acta 1422(3):207–234. https://doi.org/10.1016/S0304-4157(99)00006-4
11. Fischer E (1894) Einfluss der configuration auf die wirkung der enzyme. Ber Dt Chem Ges 27:2985–2993
12. Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. Protein Sci 11(4):739–756. https://doi.org/10.1110/ps.4210102
13. Uversky VN (2002) What does it mean to be natively unfolded? Eur J Biochem 269(1):2–12. https://doi.org/10.1046/j.0014-2956.2001.02649.x
14. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C-H, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. J Mol Graph Model 19(1):26–59. https://doi.org/10.1016/S1093-3263(00)00138-8
15. Tompa P (2003) The functional benefits of protein disorder. J Mol Struct Theochem 666:361–371. https://doi.org/10.1016/j.theochem.2003.08.047
16. Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 41(3):415–427. https://doi.org/10.1002/1097-0134(20001115)41:3%3c415:AID-PROT130%3e3.0.CO;2-7
17. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. J Mol Biol 293(2):321–331. https://doi.org/10.1006/jmbi.1999.3110
18. Xue B, Williams RW, Oldfield CJ, Dunker AK, Uversky VN (2010) Archaic chaos: intrinsically disordered proteins in Archaea. BMC Syst Biol 4(Suppl 1):S1. https://doi.org/10.1186/1752-0509-4-S1-S1
19. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, Babu MM (2014) Classification of intrinsically disordered regions and proteins. Chem Rev 114(13):6589–6631. https://doi.org/10.1021/cr400525m
20. Habchi J, Tompa P, Longhi S, Uversky VN (2014) Introducing protein intrinsic disorder. Chem Rev 114(13):6561–6588. https://doi.org/10.1021/cr400514h
21. Dunker AK, Garner E, Guilliot S, Romero P, Albrecht K, Hart J, Obradovic Z, Kissinger C, Villafranca JE (1998) Protein disorder and the evolution of molecular recognition: theory, predictions and observations. Pac Symp Biocomput 473–484
22. Tompa P (2002) Intrinsically unstructured proteins. Trends Biochem Sci 27(10):527–533. https://doi.org/10.1016/S0968-0004(02)02169-2
23. Daughdrill GW, Pielak GJ, Uversky VN, Cortese MS, Dunker AK (2005) Natively disordered proteins. In: Buchner J, Kiefhaber T (eds) Handbook of protein folding. Wiley, Weinheim, pp 271–353
24. Uversky VN (2013) Unusual biophysics of intrinsically disordered proteins. Biochim Biophys Acta 1834(5):932–951. https://doi.org/10.1016/j.bbapap.2012.12.008
25. Uversky VN, Dunker AK (2013) The case for intrinsically disordered proteins playing contributory roles in molecular recognition without a stable 3D structure. F1000 Biol Rep 5:1. https://doi.org/10.3410/b5-1
26. Peng Z, Yan J, Fan X, Mizianty MJ, Xue B, Wang K, Hu G, Uversky VN, Kurgan L (2015) Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. Cell Mol Life Sci 72(1):137–151. https://doi.org/10.1007/s00018-014-1661-9
27. Na I, Redmon D, Kopa M, Qin Y, Xue B, Uversky VN (2013) Ordered disorder of the astrocytic dystrophin-associated protein complex in the norm and pathology. PLoS One 8(8):e73476. https://doi.org/10.1371/journal.pone.0073476
28. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol 337(3):635–645. https://doi.org/10.1016/j.jmb.2004.02.002
29. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ (2000) Intrinsic protein disorder in complete genomes. Genome Inform Ser Workshop Genome Inform 11:161–171
30. Hu G, Wang K, Song J, Uversky VN, Kurgan L (2018) Taxonomic landscape of the dark proteomes: whole-proteome scale interplay between structural darkness, intrinsic disorder, and crystallization propensity. Proteomics 18(21–22):e1800243. https://doi.org/10.1002/pmic.201800243
31. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) Intrinsic disorder and protein function. Biochemistry 41(21):6573–6582. https://doi.org/10.1021/bi012159+
32. Dunker AK, Brown CJ, Obradovic Z (2002) Identification and functions of usefully disordered proteins. Adv Protein Chem 62:25–49. https://doi.org/10.1016/S0065-3233(02)62004-2
33. Uversky VN (2003) Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? Cell Mol Life Sci 60(9):1852–1871. https://doi.org/10.1007/s00018-003-3096-6
34. Uversky VN (2013) A decade and a half of protein intrinsic disorder: biology still waits for physics. Protein Sci 22(6):693–724. https://doi.org/10.1002/pro.2261
35. Dyson HJ, Wright PE (2002) Coupling of folding and binding for unstructured proteins. Curr Opin Struct Biol 12(1):54–60. https://doi.org/10.1016/S0959-440X(02)00289-0
36. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6(3):197–208. https://doi.org/10.1038/nrm1589
37. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z (2007) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res 6(5):1882–1898. https://doi.org/10.1021/pr060392u
38. Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. Curr Opin Struct Biol 18(6):756–764. https://doi.org/10.1016/j.sbi.2008.10.002
39. Dunker AK, Uversky VN (2008) Signal transduction via unstructured protein conduits. Nat Chem Biol 4(4):229–230. https://doi.org/10.1038/nchembio0408-229
40. Oldfield CJ, Meng J, Yang JY, Yang MQ, Uversky VN, Dunker AK (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genom 9(Suppl 1):S1. https://doi.org/10.1186/1471-2164-9-S1-S1
41. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J 272(20):5129–5148. https://doi.org/10.1111/j.1742-4658.2005.04948.x
42. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol 323(3):573–584. https://doi.org/10.1016/S0022-2836(02)00969-5
43. Uversky VN, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept.

Annu Rev Biophys 37:215–246. https://doi.org/10.1146/annurev.biophys.37.032807.125924
44. Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC (2014) Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem Rev 114(13):6844–6879. https://doi.org/10.1021/cr400713r
45. Dunker AK, Obradovic Z (2001) The protein trinity–linking function and disorder. Nat Biotechnol 19(9):805–806. https://doi.org/10.1038/nbt0901-805
46. Romero PR, Zaidi S, Fang YY, Uversky VN, Radivojac P, Oldfield CJ, Cortese MS, Sickmeier M, LeGall T, Obradovic Z, Dunker AK (2006) Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. Proc Natl Acad Sci USA 103(22):8390–8395. https://doi.org/10.1073/pnas.0507916103
47. Jakob U, Kriwacki R, Uversky VN (2014) Conditionally and transiently disordered proteins: awakening cryptic disorder to regulate protein function. Chem Rev 114(13):6779–6805. https://doi.org/10.1021/cr400459c
48. Uversky VN (2016) p53 proteoforms and intrinsic disorder: an illustration of the protein structure-function continuum concept. Int J Mol Sci 17(11):1874. https://doi.org/10.3390/ijms17111874
49. Schluter H, Apweiler R, Holzhutter HG, Jungblut PR (2009) Finding one's way in proteomics: a protein species nomenclature. Chem Cent J 3:11. https://doi.org/10.1186/1752-153X-3-11
50. Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Linskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skollermo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan JH, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Ponten F (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 4(12):1920–1932. https://doi.org/10.1074/mcp.M500279-MCP200
51. Farrah T, Deutsch EW, Omenn GS, Sun Z, Watts JD, Yamamoto T, Shteynberg D, Harris MM, Moritz RL (2014) State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven human proteome project. J Proteome Res 13(1):60–75. https://doi.org/10.1021/Pr4010037
52. Farrah T, Deutsch EW, Hoopmann MR, Hallows JL, Sun Z, Huang CY, Moritz RL (2013) The State of the human proteome in 2012 as viewed through PeptideAtlas. J Proteome Res 12(1):162–171. https://doi.org/10.1021/Pr301012j
53. Reddy PJ, Ray S, Srivastava S (2015) The quest of the human proteome and the missing proteins: digging deeper. Omics-a J Integr Biol 19(5):276–282. https://doi.org/10.1089/omi.2015.0035
54. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabuddhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LDN, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang TC, Zhong J, Wu XY, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TSK, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H, Pandey A (2014) A draft map of the human proteome. Nature 509(7502):575. https://doi.org/10.1038/nature13302
55. Consortium TEP (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489(7414):57–74. https://doi.org/10.1038/nature11247
56. Ponomarenko EA, Poverennaya EV, Ilgisonis EV, Pyatnitskiy MA, Kopylov AT, Zgoda VG, Lisitsa AV, Archakov AI (2016) The size of the human proteome: the width and depth. Int J Anal Chem 2016:7436849. https://doi.org/10.1155/2016/7436849
57. Smith LM, Kelleher NL, Consortium for Top Down P (2013) Proteoform: a single term describing protein complexity. Nat Methods 10(3):186–187. https://doi.org/10.1038/nmeth.2369
58. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 32(3):1037–1049. https://doi.org/10.1093/nar/gkh253
59. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P (2014) The structural and functional signatures of proteins that undergo multiple events of post-translational modification. Protein Sci 23(8):1077–1093. https://doi.org/10.1002/pro.2494
60. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK (2005) Coupled folding and binding with alpha-helix-forming molecular recognition elements. Biochemistry 44(37):12454–12470. https://doi.org/10.1021/bi050736e
61. Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK (2007) Intrinsic disorder and functional proteomics. Biophys J 92(5):1439–1456. https://doi.org/10.1529/biophysj.106.094045
62. Uversky VN, Dunker AK (2010) Understanding protein non-folding. Biochim Biophys Acta 1804(6):1231–1264. https://doi.org/10.1016/j.bbapap.2010.01.017
63. Uversky VN (2011) Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. Chem Soc Rev 40(3):1623–1634. https://doi.org/10.1039/c0cs00057d
64. Uversky VN (2012) Disordered competitive recruiter: fast and foldable. J Mol Biol 418(5):267–268. https://doi.org/10.1016/j.jmb.2012.02.034
65. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK, Uversky VN (2006) Analysis of molecular recognition features (MoRFs). J Mol Biol 362(5):1043–1059. https://doi.org/10.1016/j.jmb.2006.07.087
66. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK (2007) Characterization of molecular recognition features, MoRFs, and their binding partners. J Proteome Res 6(6):2351–2366. https://doi.org/10.1021/pr0701411
67. Beadle GW, Tatum EL (1941) Genetic control of biochemical reactions in neurospora. Proc Natl Acad Sci USA 27(11):499–506
68. Uversky VN (2015) Functional roles of transiently and intrinsically disordered regions within proteins. FEBS J 282(7):1182–1189. https://doi.org/10.1111/febs.13202
69. Uversky VN (2018) Functions of short lifetime biological structures at large: the case of intrinsically disordered proteins. Brief Funct Genomics. https://doi.org/10.1093/bfgp/ely023
70. Venkatakrishnan AJ, Flock T, Prado DE, Oates ME, Gough J, Madan Babu M (2014) Structured and disordered facets of the GPCR fold. Curr Opin Struct Biol 27:129–137. https://doi.org/10.1016/j.sbi.2014.08.002

71. Attwood TK, Findlay JB (1994) Fingerprinting G-protein-coupled receptors. Protein Eng 7(2):195–203. https://doi.org/10.1093/protein/7.2.195
72. Kolakowski LF Jr (1994) GCRDb: a G-protein-coupled receptor database. Recept Channels 2(1):1–7
73. Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, Conn PJ, Stevens RC (2014) Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. Science 344(6179):58–64. https://doi.org/10.1126/science.1249489
74. Nomura R, Suzuki Y, Kakizuka A, Jingami H (2008) Direct detection of the interaction between recombinant soluble extracellular regions in the heterodimeric metabotropic gamma-aminobutyric acid receptor. J Biol Chem 283(8):4665–4673. https://doi.org/10.1074/jbc.M705202200
75. Geng Y, Bush M, Mosyak L, Wang F, Fan QR (2013) Structural mechanism of ligand activation in human GABA(B) receptor. Nature 504(7479):254–259. https://doi.org/10.1038/nature12725
76. Burmakina S, Geng Y, Chen Y, Fan QR (2014) Heterodimeric coiled-coil interactions of human GABAB receptor. Proc Natl Acad Sci USA 111(19):6958–6963. https://doi.org/10.1073/pnas.1400081111
77. Margeta-Mitrovic M, Jan YN, Jan LY (2000) A trafficking checkpoint controls GABA(B) receptor heterodimerization. Neuron 27(1):97–106. https://doi.org/10.1016/S0896-6273(00)00012-X
78. Pagano A, Rovelli G, Mosbacher J, Lohmann T, Duthey B, Stauffer D, Ristig D, Schuler V, Meigel I, Lampert C, Stein T, Prezeau L, Blahos J, Pin J, Froestl W, Kuhn R, Heid J, Kaupmann K, Bettler B (2001) C-terminal interaction is essential for surface trafficking but not for heteromeric assembly of GABA(b) receptors. J Neurosci 21(4):1189–1202. https://doi.org/10.1523/JNEUROSCI.21-04-01189.2001
79. Kaupmann K, Malitschek B, Schuler V, Heid J, Froestl W, Beck P, Mosbacher J, Bischoff S, Kulik A, Shigemoto R, Karschin A, Bettler B (1998) GABA(B)-receptor subtypes assemble into functional heteromeric complexes. Nature 396(6712):683–687. https://doi.org/10.1038/25360
80. White JH, Wise A, Main MJ, Green A, Fraser NJ, Disney GH, Barnes AA, Emson P, Foord SM, Marshall FH (1998) Heterodimerization is required for the formation of a functional GABA(B) receptor. Nature 396(6712):679–682. https://doi.org/10.1038/25354
81. Galvez T, Duthey B, Kniazeff J, Blahos J, Rovelli G, Bettler B, Prezeau L, Pin JP (2001) Allosteric interactions between GB1 and GB2 subunits are required for optimal GABA(B) receptor function. EMBO J 20(9):2152–2159. https://doi.org/10.1093/emboj/20.9.2152
82. Geng Y, Mosyak L, Kurinov I, Zuo H, Sturchler E, Cheng TC, Subramanyam P, Brown AP, Brennan SC, Mun HC, Bush M, Chen Y, Nguyen TX, Cao B, Chang DD, Quick M, Conigrave AD, Colecraft HM, McDonald P, Fan QR (2016) Structural mechanism of ligand activation in human calcium-sensing receptor. Elife. https://doi.org/10.7554/elife.13662
83. Riccardi D, Brennan SC, Chang W (2013) The extracellular calcium-sensing receptor, CaSR, in fetal development. Best Pract Res Clin Endocrinol Metab 27(3):443–453. https://doi.org/10.1016/j.beem.2013.02.010
84. Ruat M, Traiffort E (2013) Roles of the calcium sensing receptor in the central nervous system. Best Pract Res Clin Endocrinol Metab 27(3):429–442. https://doi.org/10.1016/j.beem.2013.03.001
85. Bandyopadhyay S, Tfelt-Hansen J, Chattopadhyay N (2010) Diverse roles of extracellular calcium-sensing receptor in the central nervous system. J Neurosci Res 88(10):2073–2082. https://doi.org/10.1002/jnr.22391
86. Conigrave AD, Hampson DR (2006) Broad-spectrum L-amino acid sensing by class 3 G-protein-coupled receptors. Trends Endocrinol Metab 17(10):398–407. https://doi.org/10.1016/j.tem.2006.10.012
87. Zhang D, Gao ZG, Zhang K, Kiselev E, Crane S, Wang J, Paoletta S, Yi C, Ma L, Zhang W, Han GW, Liu H, Cherezov V, Katritch V, Jiang H, Stevens RC, Jacobson KA, Zhao Q, Wu B (2015) Two disparate ligand-binding sites in the human P2Y1 receptor. Nature 520(7547):317–321. https://doi.org/10.1038/nature14287
88. Park D, Ravichandran KS (2010) Emerging roles of brain-specific angiogenesis inhibitor 1. Adv Exp Med Biol 706:167–178. https://doi.org/10.1007/978-1-4419-7913-1_15
89. Das S, Sarkar A, Ryan KA, Fox S, Berger AH, Juncadella IJ, Bimczok D, Smythies LE, Harris PR, Ravichandran KS, Crowe SE, Smith PD, Ernst PB (2014) Brain angiogenesis inhibitor 1 is expressed by gastric phagocytes during infection with Helicobacter pylori and mediates the recognition and engulfment of human apoptotic gastric epithelial cells. FASEB J 28(5):2214–2224. https://doi.org/10.1096/fj.13-243238
90. Duda DG, Sunamura M, Lozonschi L, Yokoyama T, Yatsuoka T, Motoi F, Horii A, Tani K, Asano S, Nakamura Y, Matsuno S (2002) Overexpression of the p53-inducible brain-specific angiogenesis inhibitor 1 suppresses efficiently tumour angiogenesis. Br J Cancer 86(3):490–496. https://doi.org/10.1038/sj.bjc.6600067
91. Stephenson JR, Purcell RH, Hall RA (2014) The BAI subfamily of adhesion GPCRs: synaptic regulation and beyond. Trends Pharmacol Sci 35(4):208–215. https://doi.org/10.1016/j.tips.2014.02.002
92. Kaur B, Brat DJ, Devi NS, Van Meir EG (2005) Vasculostatin, a proteolytic fragment of brain angiogenesis inhibitor 1, is an antiangiogenic and antitumorigenic factor. Oncogene 24(22):3632–3642. https://doi.org/10.1038/sj.onc.1208317
93. Cork SM, Kaur B, Devi NS, Cooper L, Saltz JH, Sandberg EM, Kaluz S, Van Meir EG (2012) A proprotein convertase/MMP-14 proteolytic cascade releases a novel 40 kDa vasculostatin from tumor suppressor BAI1. Oncogene 31(50):5144–5152. https://doi.org/10.1038/onc.2012.1
94. Cork SM, Van Meir EG (2011) Emerging roles for the BAI1 protein family in the regulation of phagocytosis, synaptogenesis, neurovasculature, and tumor development. J Mol Med (Berl) 89(8):743–752. https://doi.org/10.1007/s00109-011-0759-x
95. Oda K, Shiratsuchi T, Nishimori H, Inazawa J, Yoshikawa H, Taketani Y, Nakamura Y, Tokino T (1999) Identification of BAIAP2 (BAI-associated protein 2), a novel human homologue of hamster IRSp53, whose SH3 domain interacts with the cytoplasmic domain of BAI1. Cytogenet Cell Genet 84(1–2):75–82. https://doi.org/10.1159/000015219
96. Shiratsuchi T, Futamura M, Oda K, Nishimori H, Nakamura Y, Tokino T (1998) Cloning and characterization of BAI-associated protein 1: a PDZ domain-containing protein that interacts with BAI1. Biochem Biophys Res Commun 247(3):597–604. https://doi.org/10.1006/bbrc.1998.8603
97. Lu YC, Nazarko OV, Sando R 3rd, Salzman GS, Li NS, Sudhof TC, Arac D (2015) Structural basis of latrophilin-FLRT-UNC5 interaction in cell adhesion. Structure 23(9):1678–1691. https://doi.org/10.1016/j.str.2015.06.024
98. Abbott RJ, Spendlove I, Roversi P, Fitzgibbon H, Knott V, Teriete P, McDonnell JM, Handford PA, Lea SM (2007) Structural and functional characterization of a novel T cell receptor co-regulatory protein complex, CD97-CD55. J Biol Chem 282(30):22023–22032. https://doi.org/10.1074/jbc.M702588200
99. Paavola KJ, Stephenson JR, Ritter SL, Alter SP, Hall RA (2011) The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. J Biol Chem 286(33):28914–28921. https://doi.org/10.1074/jbc.M111.247973

100. Stoveken HM, Hajduczok AG, Xu L, Tall GG (2015) Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. Proc Natl Acad Sci USA 112(19):6194–6199. https://doi.org/10.1073/pnas.1421785112
101. Gazit A, Yaniv A, Bafico A, Pramila T, Igarashi M, Kitajewski J, Aaronson SA (1999) Human frizzled 1 interacts with transforming Wnts to transduce a TCF dependent transcriptional response. Oncogene 18(44):5959–5966. https://doi.org/10.1038/sj.onc.1202985
102. Van Camp JK, Beckers S, Zegers D, Van Hul W (2014) Wnt signaling and the control of human stem cell fate. Stem Cell Rev 10(2):207–229. https://doi.org/10.1007/s12015-013-9486-8
103. Marchetti B, L'Episcopo F, Morale MC, Tirolo C, Testa N, Caniglia S, Serapide MF, Pluchino S (2013) Uncovering novel actors in astrocyte-neuron crosstalk in Parkinson's disease: the Wnt/beta-catenin signaling cascade as the common final pathway for neuroprotection and self-repair. Eur J Neurosci 37(10):1550–1563. https://doi.org/10.1111/ejn.12166
104. Planutis K, Planutiene M, Nguyen AV, Moyer MP, Holcombe RF (2013) Invasive colon cancer, but not non-invasive adenomas induce a gradient effect of Wnt pathway receptor frizzled 1 (Fz1) expression in the tumor microenvironment. J Transl Med 11:50. https://doi.org/10.1186/1479-5876-11-50
105. Clevers H, Nusse R (2012) Wnt/beta-catenin signaling and disease. Cell 149(6):1192–1205. https://doi.org/10.1016/j.cell.2012.05.012
106. Gay A, Towler DA (2017) Wnt signaling in cardiovascular disease: opportunities and challenges. Curr Opin Lipidol 28(5):387–396. https://doi.org/10.1097/MOL.0000000000000445
107. Nusse R, Clevers H (2017) Wnt/beta-catenin signaling, disease, and emerging therapeutic modalities. Cell 169(6):985–999. https://doi.org/10.1016/j.cell.2017.05.016
108. Masckauchan TN, Kitajewski J (2006) Wnt/Frizzled signaling in the vasculature: new angiogenic factors in sight. Physiology (Bethesda) 21:181–188. https://doi.org/10.1152/physiol.00058.2005
109. Zhang B, Ma JX (2010) Wnt pathway antagonists and angiogenesis. Protein Cell 1(10):898–906. https://doi.org/10.1007/s13238-010-0112-0
110. Clevers H (2006) Wnt/beta-catenin signaling in development and disease. Cell 127(3):469–480. https://doi.org/10.1016/j.cell.2006.10.018
111. Gorojankina T (2016) Hedgehog signaling pathway: a novel model and molecular mechanisms of signal transduction. Cell Mol Life Sci 73(7):1317–1332. https://doi.org/10.1007/s00018-015-2127-4
112. Gong X, Qian H, Cao P, Zhao X, Zhou Q, Lei J, Yan N (2018) Structural basis for the recognition of Sonic Hedgehog by human Patched1. Science. https://doi.org/10.1126/science.aas8935
113. Ingham PW, Nakano Y, Seger C (2011) Mechanisms and functions of Hedgehog signalling across the metazoa. Nat Rev Genet 12(6):393–406. https://doi.org/10.1038/nrg2984
114. Briscoe J, Therond PP (2013) The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol 14(7):416–429. https://doi.org/10.1038/nrm3598
115. McMahon AP, Ingham PW, Tabin CJ (2003) Developmental roles and clinical significance of hedgehog signaling. Curr Top Dev Biol 53:1–114. https://doi.org/10.1016/S0070-2153(03)53002-2
116. Taipale J, Beachy PA (2001) The Hedgehog and Wnt signalling pathways in cancer. Nature 411(6835):349–354. https://doi.org/10.1038/35077219
117. Zardawi SJ, O'Toole SA, Sutherland RL, Musgrove EA (2009) Dysregulation of Hedgehog, Wnt and Notch signalling pathways in breast cancer. Histol Histopathol 24(3):385–398. https://doi.org/10.14670/HH-24.385
118. Toftgard R (2000) Hedgehog signalling in cancer. Cell Mol Life Sci 57(12):1720–1731. https://doi.org/10.1007/PL00000654
119. Sharpe HJ, Wang W, Hannoush RN, de Sauvage FJ (2015) Regulation of the oncoprotein Smoothened by small molecules. Nat Chem Biol 11(4):246–255. https://doi.org/10.1038/nchembio.1776
120. Wang C, Wu H, Evron T, Vardy E, Han GW, Huang XP, Hufeisen SJ, Mangano TJ, Urban DJ, Katritch V, Cherezov V, Caron MG, Roth BL, Stevens RC (2014) Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. Nat Commun 5:4355. https://doi.org/10.1038/ncomms5355
121. Byrne EFX, Sircar R, Miller PS, Hedger G, Luchetti G, Nachtergaele S, Tully MD, Mydock-McGrane L, Covey DF, Rambo RP, Sansom MSP, Newstead S, Rohatgi R, Siebold C (2016) Structural basis of Smoothened regulation by its extracellular domains. Nature 535(7613):517–522. https://doi.org/10.1038/nature18934
122. Liang YL, Khoshouei M, Radjainia M, Zhang Y, Glukhova A, Tarrasch J, Thal DM, Furness SGB, Christopoulos G, Coudrat T, Danev R, Baumeister W, Miller LJ, Christopoulos A, Kobilka BK, Wootten D, Skiniotis G, Sexton PM (2017) Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. Nature 546(7656):118–123. https://doi.org/10.1038/nature22327
123. Kusano S, Kukimoto-Niino M, Hino N, Ohsawa N, Okuda K, Sakamoto K, Shirouzu M, Shindo T, Yokoyama S (2012) Structural basis for extracellular interactions between calcitonin receptor-like receptor and receptor activity-modifying protein 2 for adrenomedullin-specific binding. Protein Sci 21(2):199–210. https://doi.org/10.1002/pro.2003
124. Liang YL, Khoshouei M, Deganutti G, Glukhova A, Koole C, Peat TS, Radjainia M, Plitzko JM, Baumeister W, Miller LJ, Hay DL, Christopoulos A, Reynolds CA, Wootten D, Sexton PM (2018) Cryo-EM structure of the active, Gs-protein complexed, human CGRP receptor. Nature 561(7724):492–497. https://doi.org/10.1038/s41586-018-0535-y
125. Tao J, Hildebrand ME, Liao P, Liang MC, Tan G, Li S, Snutch TP, Soong TW (2008) Activation of corticotropin-releasing factor receptor 1 selectively inhibits CaV3.2 T-type calcium channels. Mol Pharmacol 73(6):1596–1609. https://doi.org/10.1124/mol.107.043612
126. Grace CR, Perrin MH, Gulyas J, Rivier JE, Vale WW, Riek R (2010) NMR structure of the first extracellular domain of corticotropin-releasing factor receptor 1 (ECD1-CRF-R1) complexed with a high affinity agonist. J Biol Chem 285(49):38580–38589. https://doi.org/10.1074/jbc.M110.121897
127. Gulyas J, Rivier C, Perrin M, Koerber SC, Sutton S, Corrigan A, Lahrichi SL, Craig AG, Vale W, Rivier J (1995) Potent, structurally constrained agonists and competitive antagonists of corticotropin-releasing factor. Proc Natl Acad Sci USA 92(23):10575–10579
128. Gilman AG (1987) G proteins: transducers of receptor-generated signals. Annu Rev Biochem 56:615–649. https://doi.org/10.1146/annurev.bi.56.070187.003151
129. Engelhardt S, Rochais F (2007) G proteins: more than transducers of receptor-generated signals? Circ Res 100(8):1109–1111. https://doi.org/10.1161/01.RES.0000266971.15127.e8
130. Simon MI, Strathmann MP, Gautam N (1991) Diversity of G proteins in signal transduction. Science 252(5007):802–808. https://doi.org/10.1126/science.1902986
131. Syrovatkina V, Alegre KO, Dey R, Huang XY (2016) Regulation, signaling, and physiological functions of G-proteins. J Mol Biol 428(19):3850–3868. https://doi.org/10.1016/j.jmb.2016.08.002
132. Bourne HR, Sanders DA, McCormick F (1990) The GTPase superfamily: a conserved switch for diverse cell functions. Nature 348(6297):125–132. https://doi.org/10.1038/348125a0
133. Brand CS, Sadana R, Malik S, Smrcka AV, Dessauer CW (2015) Adenylyl cyclase 5 regulation by gbetagamma involves

isoform-specific use of multiple interaction sites. Mol Pharmacol 88(4):758–767. https://doi.org/10.1124/mol.115.099556
134. Farfel Z, Iiri T, Shapira H, Roitman A, Mouallem M, Bourne HR (1996) Pseudohypoparathyroidism, a novel mutation in the betagamma-contact region of Gsalpha impairs receptor stimulation. J Biol Chem 271(33):19653–19655. https://doi.org/10.1074/jbc.271.33.19653
135. Plagge A, Isles AR, Gordon E, Humby T, Dean W, Gritsch S, Fischer-Colbrie R, Wilkinson LS, Kelsey G (2005) Imprinted Nesp55 influences behavioral reactivity to novel environments. Mol Cell Biol 25(8):3019–3026. https://doi.org/10.1128/MCB.25.8.3019-3026.2005
136. Hu Q, Shokat KM (2018) Disease-causing mutations in the G protein Galphas subvert the roles of GDP and GTP. Cell 173(5):1254-1264 e1211. https://doi.org/10.1016/j.cell.2018.03.018
137. Goricanec D, Stehle R, Egloff P, Grigoriu S, Pluckthun A, Wagner G, Hagn F (2016) Conformational dynamics of a G-protein alpha subunit is tightly regulated by nucleotide binding. Proc Natl Acad Sci USA 113(26):E3629–E3638. https://doi.org/10.1073/pnas.1604125113
138. Oldham WM, Van Eps N, Preininger AM, Hubbell WL, Hamm HE (2007) Mapping allosteric connections from the receptor to the nucleotide-binding pocket of heterotrimeric G proteins. Proc Natl Acad Sci USA 104(19):7927–7932. https://doi.org/10.1073/pnas.0702623104
139. Mixon MB, Lee E, Coleman DE, Berghuis AM, Gilman AG, Sprang SR (1995) Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis. Science 270(5238):954–960. https://doi.org/10.1126/science.270.5238.954
140. Coleman DE, Berghuis AM, Lee E, Linder ME, Gilman AG, Sprang SR (1994) Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. Science 265(5177):1405–1412. https://doi.org/10.1126/science.8073283
141. Chen CK, Chan NL, Wang AH (2011) The many blades of the beta-propeller proteins: conserved but versatile. Trends Biochem Sci 36(10):553–561. https://doi.org/10.1016/j.tibs.2011.07.004
142. Anurag M, Singh GP, Dash D (2012) Location of disorder in coiled coil proteins is influenced by its biological role and subcellular localization: a GO-based study on human proteome. Mol BioSyst 8(1):346–352. https://doi.org/10.1039/c1mb05210a
143. Gazi AD, Bastaki M, Charova SN, Gkougkoulia EA, Kapellios EA, Panopoulos NJ, Kokkinidis M (2008) Evidence for a coiled-coil interaction mode of disordered proteins from bacterial type III secretion systems. J Biol Chem 283(49):34062–34068. https://doi.org/10.1074/jbc.M803408200
144. Gunasekaran K, Tsai CJ, Nussinov R (2004) Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. J Mol Biol 341(5):1327–1341. https://doi.org/10.1016/j.jmb.2004.07.002
145. Wu Z, Hu G, Yang J, Peng Z, Uversky VN, Kurgan L (2015) In various protein complexes, disordered protomers have large per-residue surface areas and area of protein-, DNA- and RNA-binding interfaces. FEBS Lett 589(19 Pt A):2561–2569. https://doi.org/10.1016/j.febslet.2015.08.014
146. Bock A, Kostenis E, Trankle C, Lohse MJ, Mohr K (2014) Pilot the pulse: controlling the multiplicity of receptor dynamics. Trends Pharmacol Sci 35(12):630–638. https://doi.org/10.1016/j.tips.2014.10.002
147. Van Eps N, Caro LN, Morizumi T, Kusnetzow AK, Szczepek M, Hofmann KP, Bayburt TH, Sligar SG, Ernst OP, Hubbell WL (2017) Conformational equilibria of light-activated rhodopsin in nanodiscs. Proc Natl Acad Sci USA 114(16):E3268–E3275. https://doi.org/10.1073/pnas.1620405114
148. Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, Lerch MT, Kobilka TS, Thian FS, Hubbell WL, Prosser RS, Kobilka BK (2015) Structural insights into the dynamic process of beta2-adrenergic receptor signaling. Cell 161(5):1101–1111. https://doi.org/10.1016/j.cell.2015.04.043
149. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK (2013) The dynamic process of beta(2)-adrenergic receptor activation. Cell 152(3):532–542. https://doi.org/10.1016/j.cell.2013.01.008
150. Kofuku Y, Ueda T, Okude J, Shiraishi Y, Kondo K, Maeda M, Tsujishita H, Shimada I (2012) Efficacy of the beta(2)-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. Nat Commun 3:1045. https://doi.org/10.1038/ncomms2046
151. Violin JD, Crombie AL, Soergel DG, Lark MW (2014) Biased ligands at G-protein-coupled receptors: promise and progress. Trends Pharmacol Sci 35(7):308–316. https://doi.org/10.1016/j.tips.2014.04.007
152. Rajagopal S, Rajagopal K, Lefkowitz RJ (2010) Teaching old receptors new tricks: biasing seven-transmembrane receptors. Nat Rev Drug Discov 9(5):373–386. https://doi.org/10.1038/nrd3024
153. Strachan RT, Sun JP, Rominger DH, Violin JD, Ahn S, Rojas Bie Thomsen A, Zhu X, Kleist A, Costa T, Lefkowitz RJ (2014) Divergent transducer-specific molecular efficacies generate biased agonism at a G protein-coupled receptor (GPCR). J Biol Chem 289(20):14211–14224. https://doi.org/10.1074/jbc.M114.548131
154. Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, Liu W, Xu HE, Cherezov V, Roth BL, Stevens RC (2013) Structural features for functional selectivity at serotonin receptors. Science 340(6132):615–619. https://doi.org/10.1126/science.1232808
155. Wootten D, Simms J, Miller LJ, Christopoulos A, Sexton PM (2013) Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. Proc Natl Acad Sci USA 110(13):5211–5216. https://doi.org/10.1073/pnas.1221585110
156. Katritch V, Cherezov V, Stevens RC (2013) Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol 53:531–556. https://doi.org/10.1146/annurev-pharmtox-032112-135923
157. Mary S, Damian M, Louet M, Floquet N, Fehrentz JA, Marie J, Martinez J, Baneres JL (2012) Ligands and signaling proteins govern the conformational landscape explored by a G protein-coupled receptor. Proc Natl Acad Sci USA 109(21):8304–8309. https://doi.org/10.1073/pnas.1119881109
158. Rahmeh R, Damian M, Cottet M, Orcel H, Mendre C, Durroux T, Sharma KS, Durand G, Pucci B, Trinquet E, Zwier JM, Deupi X, Bron P, Baneres JL, Mouillac B, Granier S (2012) Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. Proc Natl Acad Sci USA 109(17):6733–6738. https://doi.org/10.1073/pnas.1201093109
159. Liu JJ, Horst R, Katritch V, Stevens RC, Wuthrich K (2012) Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR. Science 335(6072):1106–1110. https://doi.org/10.1126/science.1215802
160. Reiter E, Ahn S, Shukla AK, Lefkowitz RJ (2012) Molecular mechanism of beta-arrestin-biased agonism at seven-transmembrane receptors. Annu Rev Pharmacol Toxicol 52:179–197. https://doi.org/10.1146/annurev.pharmtox.010909.105800
161. Kahsai AW, Xiao K, Rajagopal S, Ahn S, Shukla AK, Sun J, Oas TG, Lefkowitz RJ (2011) Multiple ligand-specific conformations of the beta2-adrenergic receptor. Nat Chem Biol 7(10):692–700. https://doi.org/10.1038/nchembio.634
162. Rajagopal S, Ahn S, Rominger DH, Gowen-MacDonald W, Lam CM, Dewire SM, Violin JD, Lefkowitz RJ (2011) Quantifying

ligand bias at seven-transmembrane receptors. Mol Pharmacol 80(3):367–377. https://doi.org/10.1124/mol.111.072801
163. Kenakin T (2011) Functional selectivity and biased receptor signaling. J Pharmacol Exp Ther 336(2):296–302. https://doi.org/10.1124/jpet.110.173948
164. Kobilka BK, Deupi X (2007) Conformational complexity of G-protein-coupled receptors. Trends Pharmacol Sci 28(8):397–406. https://doi.org/10.1016/j.tips.2007.06.003
165. Yao X, Parnot C, Deupi X, Ratnala VR, Swaminath G, Farrens D, Kobilka B (2006) Coupling ligand structure to specific conformational switches in the beta2-adrenoceptor. Nat Chem Biol 2(8):417–422. https://doi.org/10.1038/nchembio801
166. Pierce KL, Premont RT, Lefkowitz RJ (2002) Seven-transmembrane receptors. Nat Rev Mol Cell Biol 3(9):639–650. https://doi.org/10.1038/nrm908
167. Luttrell LM, Maudsley S, Gesty-Palmer D (2018) Translating in vitro ligand bias into in vivo efficacy. Cell Signal 41:46–55. https://doi.org/10.1016/j.cellsig.2017.05.002
168. Maudsley S, Martin B, Luttrell LM (2005) The origins of diversity and specificity in g protein-coupled receptor signaling. J Pharmacol Exp Ther 314(2):485–494. https://doi.org/10.1124/jpet.105.083121
169. McLatchie LM, Fraser NJ, Main MJ, Wise A, Brown J, Thompson N, Solari R, Lee MG, Foord SM (1998) RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature 393(6683):333–339. https://doi.org/10.1038/30666
170. Evans BN, Rosenblatt MI, Mnayer LO, Oliver KR, Dickerson IM (2000) CGRP-RCP, a novel protein required for signal transduction at calcitonin gene-related peptide and adrenomedullin receptors. J Biol Chem 275(40):31438–31443. https://doi.org/10.1074/jbc.M005604200
171. Luttrell LM, Ferguson SS, Daaka Y, Miller WE, Maudsley S, Della Rocca GJ, Lin F, Kawakatsu H, Owada K, Luttrell DK, Caron MG, Lefkowitz RJ (1999) Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes. Science 283(5402):655–661
172. Maudsley S, Martin B, Janssens J, Etienne H, Jushaj A, van Gastel J, Willemsen A, Chen H, Gesty-Palmer D, Luttrell LM (2016) Informatic deconvolution of biased GPCR signaling mechanisms from in vivo pharmacological experimentation. Methods 92:51–63. https://doi.org/10.1016/j.ymeth.2015.05.013
173. Maudsley S, Martin B, Gesty-Palmer D, Cheung H, Johnson C, Patel S, Becker KG, Wood WH 3rd, Zhang Y, Lehrmann E, Luttrell LM (2015) Delineation of a conserved arrestin-biased signaling repertoire in vivo. Mol Pharmacol 87(4):706–717. https://doi.org/10.1124/mol.114.095224
174. Leysen H, van Gastel J, Hendrickx JO, Santos-Otte P, Martin B, Maudsley S (2018) G protein-coupled receptor systems as crucial regulators of DNA damage response processes. Int J Mol Sci. https://doi.org/10.3390/ijms19102919
175. Ritter SL, Hall RA (2009) Fine-tuning of GPCR activity by receptor-interacting proteins. Nat Rev Mol Cell Biol 10(12):819–830. https://doi.org/10.1038/nrm2803
176. Pak Y, Pham N, Rotin D (2002) Direct binding of the beta1 adrenergic receptor to the cyclic AMP-dependent guanine nucleotide exchange factor CNrasGEF leads to Ras activation. Mol Cell Biol 22(22):7942–7952
177. Hall RA, Premont RT, Chow CW, Blitzer JT, Pitcher JA, Claing A, Stoffel RH, Barak LS, Shenolikar S, Weinman EJ, Grinstein S, Lefkowitz RJ (1998) The beta2-adrenergic receptor interacts with the Na+/H+-exchanger regulatory factor to control Na+/H+ exchange. Nature 392(6676):626–630. https://doi.org/10.1038/33458
178. Mahon MJ, Donowitz M, Yun CC, Segre GV (2002) Na(+)/H(+) exchanger regulatory factor 2 directs parathyroid hormone 1 receptor signalling. Nature 417(6891):858–861. https://doi.org/10.1038/nature00816
179. Balasubramanian S, Fam SR, Hall RA (2007) GABAB receptor association with the PDZ scaffold Mupp1 alters receptor stability and function. J Biol Chem 282(6):4162–4171. https://doi.org/10.1074/jbc.M607695200
180. Guillaume JL, Daulat AM, Maurice P, Levoye A, Migaud M, Brydon L, Malpaux B, Borg-Capra C, Jockers R (2008) The PDZ protein mupp1 promotes Gi coupling and signaling of the Mt1 melatonin receptor. J Biol Chem 283(24):16762–16771. https://doi.org/10.1074/jbc.M802069200
181. Yao R, Natsume Y, Noda T (2004) MAGI-3 is involved in the regulation of the JNK signaling pathway as a scaffold protein for frizzled and Ltap. Oncogene 23(36):6023–6030. https://doi.org/10.1038/sj.onc.1207817
182. Zhang H, Wang D, Sun H, Hall RA, Yun CC (2007) MAGI-3 regulates LPA-induced activation of Erk and RhoA. Cell Signal 19(2):261–268. https://doi.org/10.1016/j.cellsig.2006.06.008
183. Yamada T, Ohoka Y, Kogo M, Inagaki S (2005) Physical and functional interactions of the lysophosphatidic acid receptors with PDZ domain-containing Rho guanine nucleotide exchange factors (RhoGEFs). J Biol Chem 280(19):19358–19363. https://doi.org/10.1074/jbc.M414561200
184. Yun HM, Kim S, Kim HJ, Kostenis E, Kim JI, Seong JY, Baik JH, Rhim H (2007) The novel cellular mechanism of human 5-HT6 receptor through an interaction with Fyn. J Biol Chem 282(8):5496–5505. https://doi.org/10.1074/jbc.M606215200
185. Hall RA, Lefkowitz RJ (2002) Regulation of G protein-coupled receptor signaling by scaffold proteins. Circ Res 91(8):672–680
186. Marrero MB, Venema VJ, Ju H, Eaton DC, Venema RC (1998) Regulation of angiotensin II-induced JAK2 tyrosine phosphorylation: roles of SHP-1 and SHP-2. Am J Physiol 275(5 Pt 1):C1216–C1223
187. Godeny MD, Sayyah J, VonDerLinden D, Johns M, Ostrov DA, Caldwell-Busby J, Sayeski PP (2007) The N-terminal SH2 domain of the tyrosine phosphatase, SHP-2, is essential for Jak2-dependent signaling via the angiotensin II type AT1 receptor. Cell Signal 19(3):600–609. https://doi.org/10.1016/j.cellsig.2006.08.010
188. Lukashova V, Asselin C, Krolewski JJ, Rola-Pleszczynski M, Stankova J (2001) G-protein-independent activation of Tyk2 by the platelet-activating factor receptor. J Biol Chem 276(26):24113–24121. https://doi.org/10.1074/jbc.M100720200
189. Lukashova V, Chen Z, Duhe RJ, Rola-Pleszczynski M, Stankova J (2003) Janus kinase 2 activation by the platelet-activating factor receptor (PAFR): roles of Tyk2 and PAFR C terminus. J Immunol 171(7):3794–3800
190. Shih M, Lin F, Scott JD, Wang HY, Malbon CC (1999) Dynamic complexes of beta2-adrenergic receptors with protein kinases and phosphatases and the role of gravin. J Biol Chem 274(3):1588–1595
191. Fraser ID, Cong M, Kim J, Rollins EN, Daaka Y, Lefkowitz RJ, Scott JD (2000) Assembly of an A kinase-anchoring protein-beta(2)-adrenergic receptor complex facilitates receptor phosphorylation and signaling. Curr Biol 10(7):409–412
192. Fan G, Shumay E, Wang H, Malbon CC (2001) The scaffold protein gravin (cAMP-dependent protein kinase-anchoring protein 250) binds the beta 2-adrenergic receptor via the receptor cytoplasmic Arg-329 to Leu-413 domain and provides a mobile scaffold during desensitization. J Biol Chem 276(26):24005–24014. https://doi.org/10.1074/jbc.M011199200
193. Tao J, Wang HY, Malbon CC (2003) Protein kinase A regulates AKAP250 (gravin) scaffold binding to the beta2-adrenergic receptor. EMBO J 22(24):6419–6429. https://doi.org/10.1093/emboj/cdg628

194. Gardner LA, Tavalin SJ, Goehring AS, Scott JD, Bahouth SW (2006) AKAP79-mediated targeting of the cyclic AMP-dependent protein kinase to the beta1-adrenergic receptor promotes recycling and functional resensitization of the receptor. J Biol Chem 281(44):33537–33553. https://doi.org/10.1074/jbc.M601809200
195. Brakeman PR, Lanahan AA, O'Brien R, Roche K, Barnes CA, Huganir RL, Worley PF (1997) Homer: a protein that selectively binds metabotropic glutamate receptors. Nature 386(6622):284–288. https://doi.org/10.1038/386284a0
196. Xiao B, Tu JC, Petralia RS, Yuan JP, Doan A, Breder CD, Ruggiero A, Lanahan AA, Wenthold RJ, Worley PF (1998) Homer regulates the association of group 1 metabotropic glutamate receptors with multivalent complexes of homer-related, synaptic proteins. Neuron 21(4):707–716
197. Tu JC, Xiao B, Yuan JP, Lanahan AA, Leoffert K, Li M, Linden DJ, Worley PF (1998) Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. Neuron 21(4):717–726
198. Kato A, Ozawa F, Saitoh Y, Fukazawa Y, Sugiyama H, Inokuchi K (1998) Novel members of the Vesl/Homer family of PDZ proteins that bind metabotropic glutamate receptors. J Biol Chem 273(37):23969–23975
199. Kammermeier PJ, Worley PF (2007) Homer 1a uncouples metabotropic glutamate receptor 5 from postsynaptic effectors. Proc Natl Acad Sci USA 104(14):6055–6060. https://doi.org/10.1073/pnas.0608991104
200. Kammermeier PJ (2008) Endogenous homer proteins regulate metabotropic glutamate receptor signaling in neurons. J Neurosci 28(34):8560–8567. https://doi.org/10.1523/JNEUROSCI.1830-08.2008
201. Sokolina K, Kittanakom S, Snider J, Kotlyar M, Maurice P, Gandia J, Benleulmi-Chaachoua A, Tadagaki K, Oishi A, Wong V, Malty RH, Deineko V, Aoki H, Amin S, Yao Z, Morato X, Otasek D, Kobayashi H, Menendez J, Auerbach D, Angers S, Przulj N, Bouvier M, Babu M, Ciruela F, Jockers R, Jurisica I, Stagljar I (2017) Systematic protein-protein interaction mapping for clinically relevant human GPCRs. Mol Syst Biol 13(3):918. https://doi.org/10.15252/msb.20167430
202. Benleulmi-Chaachoua A, Chen L, Sokolina K, Wong V, Jurisica I, Emerit MB, Darmon M, Espin A, Stagljar I, Tafelmeyer P, Zamponi GW, Delagrange P, Maurice P, Jockers R (2016) Protein interactome mining defines melatonin MT1 receptors as integral component of presynaptic protein complexes of neurons. J Pineal Res 60(1):95–108. https://doi.org/10.1111/jpi.12294
203. Gutierrez-Rodriguez M, Herranz R (2015) From multiple PAR1 receptor/protein interactions to their multiple therapeutic implications. Curr Top Med Chem 15(20):2080–2114
204. Dunn HA, Patil DN, Cao Y, Orlandi C, Martemyanov KA (2018) Synaptic adhesion protein ELFN1 is a selective allosteric modulator of group III metabotropic glutamate receptors in trans. Proc Natl Acad Sci USA 115(19):5022–5027. https://doi.org/10.1073/pnas.1722498115
205. Iyer K, Burkle L, Auerbach D, Thaminy S, Dinkel M, Engels K (2005) Stagljar I (2005) Utilizing the split-ubiquitin membrane yeast two-hybrid system to identify protein–protein interactions of integral membrane proteins. Sci STKE 275:pl3. https://doi.org/10.1126/stke.2752005pl3
206. Snider J, Kittanakom S, Damjanovic D, Curak J, Wong V, Stagljar I (2010) Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast. Nat Protoc 5(7):1281–1293. https://doi.org/10.1038/nprot.2010.83
207. Asseck LY, Grefen C (2018) Detecting interactions of membrane proteins: the split-ubiquitin system. Methods Mol Biol 1794:49–60. https://doi.org/10.1007/978-1-4939-7871-7_4
208. Petschnigg J, Wong V, Snider J, Stagljar I (2012) Investigation of membrane protein interactions using the split-ubiquitin membrane yeast two-hybrid system. Methods Mol Biol 812:225–244. https://doi.org/10.1007/978-1-61779-455-1_13
209. Thaminy S, Miller J, Stagljar I (2004) The split-ubiquitin membrane-based yeast two-hybrid system. Methods Mol Biol 261:297–312. https://doi.org/10.1385/1-59259-762-9:297
210. Stagljar I, Korostensky C, Johnsson N, te Heesen S (1998) A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. Proc Natl Acad Sci USA 95(9):5187–5192
211. Downes GB, Gautam N (1999) The G protein subunit gene families. Genomics 62(3):544–552. https://doi.org/10.1006/geno.1999.5992
212. Temple BR, Jones CD, Jones AM (2010) Evolution of a signaling nexus constrained by protein interfaces and conformational states. PLoS Comput Biol 6(10):e1000962. https://doi.org/10.1371/journal.pcbi.1000962
213. Jones AM, Assmann SM (2004) Plants: the latest model system for G-protein research. EMBO Rep 5(6):572–578. https://doi.org/10.1038/sj.embor.7400174
214. Pandey S, Chen JG, Jones AM, Assmann SM (2006) G-protein complex mutants are hypersensitive to abscisic acid regulation of germination and postgermination development. Plant Physiol 141(1):243–256. https://doi.org/10.1104/pp.106.079038
215. Ding L, Pandey S, Assmann SM (2008) *Arabidopsis* extra-large G proteins (XLGs) regulate root morphogenesis. Plant J 53(2):248–263. https://doi.org/10.1111/j.1365-313X.2007.03335.x
216. Chakravorty D, Trusov Y, Zhang W, Acharya BR, Sheahan MB, McCurdy DW, Assmann SM, Botella JR (2011) An atypical heterotrimeric G-protein gamma-subunit is involved in guard cell K(+)-channel regulation and morphological development in *Arabidopsis thaliana*. Plant J 67(5):840–851. https://doi.org/10.1111/j.1365-313X.2011.04638.x
217. Urano D, Chen JG, Botella JR, Jones AM (2013) Heterotrimeric G protein signalling in the plant kingdom. Open Biol 3(3):120186. https://doi.org/10.1098/rsob.120186
218. Wolfenstetter S, Chakravorty D, Kula R, Urano D, Trusov Y, Sheahan MB, McCurdy DW, Assmann SM, Jones AM, Botella JR (2015) Evidence for an unusual transmembrane configuration of AGG3, a class C Ggamma subunit of *Arabidopsis*. Plant J 81(3):388–398. https://doi.org/10.1111/tpj.12732
219. Urano D, Maruta N, Trusov Y, Stoian R, Wu Q, Liang Y, Jaiswal DK, Thung L, Jackson D, Botella JR, Jones AM (2016) Saltational evolution of the heterotrimeric G protein signaling mechanisms in the plant kingdom. Sci Signal 9(446):ra93. https://doi.org/10.1126/scisignal.aaf9558
220. Liang Y, Gao Y, Jones AM (2017) Extra large G-protein interactome reveals multiple stress response function and partner-dependent XLG subcellular localization. Front Plant Sci 8:1015. https://doi.org/10.3389/fpls.2017.01015
221. Klopffleisch K, Phan N, Augustin K, Bayne RS, Booker KS, Botella JR, Carpita NC, Carr T, Chen JG, Cooke TR, Frick-Cheng A, Friedman EJ, Fulk B, Hahn MG, Jiang K, Jorda L, Kruppe L, Liu C, Lorek J, McCann MC, Molina A, Moriyama EN, Mukhtar MS, Mudgil Y, Pattathil S, Schwarz J, Seta S, Tan M, Temp U, Trusov Y, Urano D, Welter B, Yang J, Panstruga R, Uhrig JF, Jones AM (2011) *Arabidopsis* G-protein interactome reveals connections to cell wall carbohydrates and morphogenesis. Mol Syst Biol 7:532. https://doi.org/10.1038/msb.2011.66
222. Fischer-Colbrie R, Eder S, Lovisetti-Scamihorn P, Becker A, Laslop A (2002) Neuroendocrine secretory protein 55: a novel marker for the constitutive secretory pathway. Ann N Y Acad Sci 971:317–322. https://doi.org/10.1111/j.1749-6632.2002.tb04486.x

223. Bastepe M (2007) The GNAS locus: quintessential complex gene encoding Gsalpha, XLalphas, and other imprinted transcripts. Curr Genomics 8(6):398–414. https://doi.org/10.2174/138920207783406488
224. Freson K, Jaeken J, Van Helvoirt M, de Zegher F, Wittevrongel C, Thys C, Hoylaerts MF, Vermylen J, Van Geet C (2003) Functional polymorphisms in the paternally expressed XLalphas and its cofactor ALEX decrease their mutual interaction and enhance receptor-mediated cAMP formation. Hum Mol Genet 12(10):1121–1130. https://doi.org/10.1093/hmg/ddg130
225. Shatsky M, Nussinov R, Wolfson HJ (2004) A method for simultaneous alignment of multiple protein structures. Proteins 56(1):143–156. https://doi.org/10.1002/prot.10628
226. Waldschmidt HV, Homan KT, Cruz-Rodriguez O, Cato MC, Waninger-Saroni J, Larimore KM, Cannavo A, Song J, Cheung JY, Kirchhoff PD, Koch WJ, Tesmer JJ, Larsen SD (2016) Structure-based design, synthesis, and biological evaluation of highly selective and potent G protein-coupled receptor kinase 2 inhibitors. J Med Chem 59(8):3793–3807. https://doi.org/10.1021/acs.jmedchem.5b02000