Highlights

- Genome-editing is revolutionising all aspects of cellular biology, including pharmacological studies, including the ability to define ‘on-target’ effects of pharmaceutical ligands.
- The ability to eliminate expression from single or multiple genes provides ‘knock-out’ rather than ‘knock-down’ cell lines derived from widely used model cell systems, as well as the capacity to explore the specific functions of closely related proteins.
- Such approaches have provided new insights into the biological roles of a number of poorly characterised GPCRs.
- Elimination of expression of entire subfamilies of G protein α subunits and of both the widely expressed arrestin proteins has provided new understanding of the roles of G protein-mediated versus arrestin-mediated signalling from GPCRs.
- Genome-editing will be expanded rapidly to target other cell signalling pathway components.

Genome-Editing Provides New Insights into Receptor Controlled Signalling Pathways

Graeme Milligan[1] and Asuka Inoue[2]

[1] Centre for Translational Pharmacology, Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, U.K.

[2] Graduate School of Pharmaceutical Sciences Tohoku University 6-3, Aoba, Aramaki, Aoba-ku, Sendai, Miyagi, 980-8578 Japan

Short title: genome editing and GPCR signalling

Correspondence: Graeme.Milligan@glasgow.ac.uk (G.Milligan) or 井上 飛鳥 iaska@m.tohoku.ac.jp (A.Inoue).

Keywords

Genome editing, CRISPR/Cas9, G protein-coupled receptor, G protein, arrestin

Abstract

Rapid developments in genome-editing, based largely on CRISPR/Cas9 technologies, are offering unprecedented opportunities to eliminate the expression of single or multiple gene products in intact organisms and in model cell systems. Eliminating individual G protein-coupled receptors, both single and multiple G protein subunits, and arrestin adaptor proteins is providing new, and sometimes unanticipated, insights into molecular details of regulation of cell signalling pathways and the behaviour of receptor ligands. Genome-editing is certain to become a central component of therapeutic target validation and will provide pharmacologists with new understandings of the complexities of action of novel and previously studied ligands as well as the transmission of signals from individual cell surface receptors to intracellular signalling cascades.

Challenges in pharmacological research

Defining both target selectivity and 'on-target'-mediated effects of tool compounds and drug-like synthetic ligands have long been major challenges for pharmacologists. Moreover, unravelling the relative contributions of multiple intracellular signalling pathways and cascades initiated by activation of a single type of cell-surface receptor to a downstream cellular end-point or a physiological effect has also been highly reliant on the availability, effectiveness, and selectivity of various inhibitors and **antagonists** (see **Glossary**). In the absence of such tool compounds, and prior to the recent development of genome-editing approaches, particularly those based on the **CRISPR/Cas9** system (**Text Box 1**), such analyses had often been limited to the use of 'antisense' and other mRNA '**knock-down**' strategies. Using these it was anticipated that observed reduction in levels of a mRNA would be followed rapidly by substantial reduction in cellular levels and function of the corresponding protein(s). Whilst this may generally be true for proteins that turnover rapidly

it is not the case for proteins with longer cellular half-life. Knock-down strategies have produced a vast array of semi-quantitative studies that have linked the extent of assessed 'knock-down' (at both mRNA and protein level) with reduction in function. Gene '**knock-outs**' in both mice and non-mammalian model organisms have provided cells (e.g. mouse embryo fibroblasts) and tissues to explore the implications of true elimination of expression of proteins of interest. However, maintained lack of expression in intact organisms can frequently lead to substantial alterations in other cellular proteins that attempt to compensate and maintain homeostatic balance. Although this has been improved by moves from 'whole body' knock-out to more tissue-specific and conditional elimination of expression, many of the same issues remain. As discussed by Alvarez-Curto et al., [1] knock-out of expression of various genes may be better tolerated, without inducing alterations in levels of other network components, in simple model cell systems maintained in cell culture as many of these are transformed and genetically modified in their base-state (**Text Box 2**).

In this article we focus on the use of CRISPR/Cas9-based genome-editing in such model cell systems and on the lessons learned to date from such studies about signal selection and signal transduction pathways promoted by members of the **G protein-coupled receptor (GPCR)** superfamily. Many of these studies have been performed using human embryonic kidney (HEK)293 cells (**Figure 1**) and their various derivatives. This is useful because HEK293 cells are widely used by the research community to study underpinning mechanisms of cellular signalling. However, a range of other cell lines and primary cell types have also been used, and this is likely to expand rapidly as genome-editing approaches become ever more widely adopted. In the studies we will consider various G protein subunits, arrestins or other signalling pathway components have been eliminated. In a number of cases expression of GPCRs has also been targeted. This is already providing novel insights into roles of poorly characterised GPCRs. Furthermore, this approach is likely to have increasing utility in

providing **null backgrounds** for ligand screening programmes where in the cell background of choice various GPCRs are expressed at sufficiently high level to generate signals that can obfuscate efforts to define novel ligands for so called **orphan** GPCRs or complicate analysis of effects at closely-related members of the GPCR superfamily.

Elimination of expression of GPCRs

mRNAs encoding a substantial number of individual GPCRs are expressed at detectable levels in virtually all tissues and cell types [2, 3, 4]. Even for widely used model cell systems, including HEK293 cells, that are often considered as a 'blank canvas' for introduction and analysis of receptors of interest, studies have shown detectable expression of a wide range of GPCR family members [5]. Unsurprisingly, pluripotent embryonic stem cells display high level expression of various members of the Frizzled subfamily of GPCRs as these play key roles in embryonic development [6]. Moreover, subsets of the members of the lysophospholipid (lysophosphatidic acid (LPA) and sphingosine 1-phosphate (S1P) activated) receptor family are widely expressed, and co-expressed. For example, three distinct S1P receptors and a further LPA receptor are expressed endogenously by CHO-K1 cells [7]. This is a model cell line widely used in ligand screening projects. This pattern of co-expression reflects the importance of these receptors in cellular growth control. This does mean, however, that cells lacking expression of genuine LPA- and S1P-activated GPCRs are uncommon. This has likely contributed to still unsubstantiated suggestions that LPA or highly related molecules are the endogenously generated agonists for a number of other poorly characterised GPCRs [8, 9] beyond the currently accepted six LPA receptors (LPA_{1-6}) and five S1P receptors ($S1P_{1-5}$).

The key signalling mechanisms of many GPCRs are well established. However, even for the β_2-adrenoceptor for which a very well described and extensive pharmacopeia is available**,**

researchers have recently employed CRISPR/Cas9-based genome-editing to define unambiguously the contribution of this receptor to a less well established signalling outcome. Galaz-Montoya et al. [10] eliminated endogenous expression of the β_2-adrenoceptor from HEK293 cells to ensure that β-adrenoceptor agonist function in these cells that resulted in elevation of intracellular Ca^{2+} levels truly reflected activation of this receptor. This was despite concurrent pharmacological studies providing all but unequivocal evidence that this must be the case. This level of analysis indicates that CRISPR/Cas9-based genome-editing is becoming available to many and is convenient to use, as long as simple and robust strategies are in place to confirm successful elimination of expression in genome-edited cell clones (**Text Box 2**). Galaz-Montoya et al. [10] then performed further pharmacological studies that appeared to rule out a role of the adenylyl cyclase stimulatory G protein G_s in generating the Ca^{2+} signal. Related studies by Stallaert et al., [11] also explored mechanisms underlying the ability of β_2-adrenoceptor agonists to increase intracellular Ca^{2+} levels in HEK293 cells. In part, these studies also used genome-edited cells, but in this case engineered to lack expression of the α subunit of G_s. Now β_2-adrenoceptor agonists were unable to generate the Ca^{2+} signal, suggesting a direct role for $G\alpha_s$. Further studies demonstrated that the mechanism reflected β_2-adrenoceptor-mediated release of ATP. This nucleotide then acted as an autocrine factor to activate an endogenously expressed P2Y purinoceptor which, via its own G_q/G_{11}-based coupling to phospholipase C, induced the Ca^{2+} elevation [11]. This highlights an example in which incomplete appreciation of the endogenous expression patterns of GPCRs can generate results and lead to conclusions that on first sight may appear unexpected or incongruous.

At the other extreme in terms of the current level of understanding of the roles of a GPCR, a genome-wide screen was conducted in the human myeloid leukemia cell line KBM7 to identify expressed proteins required for the cellular action of the *Pseudomonas aeruginosa*

produced exotoxin A. One of the identified proteins was the orphan GPCR GPR107 [12] about which virtually nothing was known. Subsequent CRISPR/Cas9-based elimination of GPR107 from HeLa cells substantially reduced the ability of exotoxin A to kill these cells. Further studies showed GPR107 to be localised to the trans-Golgi network and to play a role in intracellular vesicle trafficking [12]. In a similar vein, Bostaille et al., [13] employed CRISPR/Cas9-mediated elimination of the poorly understood adhesion family GPCR Adgra2 (also designated GPR124) from HEK293 cells to allow detailed study of trafficking of variants of this receptor in an Adgra2-'null' cell background. Potential roles of GPRC5A, another 'orphan' GPCR, but part of the class C or glutamate-like receptor subfamily, have also been studied following CPISPR/Cas9-mediated gene targeting in a number of cell lines [14]. These studies indicated an important role for GPRC5A in cell adhesion, and linked this GPCR to the function of integrins via activation of the RhoA and Rac1 small GTPases. Regulation of these pathways has also been posited to suggest a role for GPRC5A in cancer progression [14]. Also in a therapeutic context, although the outcomes were predictable because it is well established to act as a co-receptor that allows infection by strains of the human immunodeficiency virus type 1, CPISPR/Cas9-mediated elimination of the chemokine receptor CXCR4 from cells, including human primary $CD4^+$ T cells, was shown to result in resistance to viral entry [15].

Elimination of expression of subunits of heterotrimeric G proteins

Although many GPCRs are often presented as interacting predominantly with a single G protein, or with members of only one of the four subfamilies of heterotrimeric G proteins (**Figure 1**), this is frequently an over-simplification. Indeed, at least in transfected cell systems, many GPCRs can interact with multiple G protein subclasses, although there may be a rank-order or hierarchy to this. A substantial number of studies have been published in the

last two years that have used HEK293 cell lines isolated following genome-editing designed to eliminate expression of various sets of G protein α subunits.

Elimination of $G\alpha_q/G\alpha_{11}$

The first published of these studies employed a clone of HEK293 cells that had been genome-edited to eliminate expression of both $G\alpha_q$ and $G\alpha_{11}$. These two G proteins are closely related, are generally co-expressed, and are the principal means to link cell surface GPCRs to activation of phospholipase Cβ1. This results in the generation of inositol 1,4,5 trisphosphate and elevation of cytosolic Ca^{2+} levels via release from the endoplasmic reticulum. The two other members of the human $G\alpha_q$ family, $G\alpha_{14}$ and $G\alpha_{16}$, are not expressed by HEK293 cells [5] (**Figure 1**). In recent years, a number of chemical inhibitors of members of the $G\alpha_q/G\alpha_{11}$ family G proteins have become available [16-17] (**Text Box 3**). As part of a far-reaching study that defined that the depsipeptide **FR900359** was a potent and selective inhibitor of G proteins of this sub-family Schrage et al., [18] showed that whilst ATP-mediated activation of an endogenously expressed P2Y purinoceptor is observed in parental HEK293 cells, this was lacking in the $G\alpha_q/G\alpha_{11}$**-**null genome-edited cells. The effect of ATP was restored by re-introduction of $G\alpha_q$ into the genome-edited cells and in both parental HEK293 cells and $G\alpha_q$ re-transfected G_q/G_{11}**-**null cells the Ca^{2+} signal induced by ATP was fully blocked by addition of FR900359 [18]. A key component of this study was that re-introduction into the G_q/G_{11}**-**null cells of one of the proteins eliminated by the genome-editing process was both able and sufficient to fully restore function and phenotype [18]. This confirms that the deficit observed in the genome-edited cell line was directly due to the lack of this protein. Similar re-capitulation of function or phenotype studies should be an integral part of the experimental design of any project that eliminates expression of a protein of interest. Not least, this reflects that genome editing can result in 'off-target' effects that

result in the elimination of expression of gene products other than those the approach was designed to modify (**Text Box 4**). Equivalent results to those with ATP were also obtained when using the acetylcholine mimetic carbachol to activate the muscarinic M_3 receptor [18], another GPCR that is endogenously expressed to significant levels by HEK293 cells. Interestingly, lack of expression of $G\alpha_q$ and $G\alpha_{11}$ did not affect cell viability or proliferation [18]. However, compared to parental HEK293 cells migration towards foetal bovine serum of the cells lacking both $G\alpha_q$ and $G\alpha_{11}$ is markedly impaired [19]. Re-introduction of $G\alpha_q$ into the G_q/G_{11}-knock out cells enhanced the migration response to transforming growth factor-β (TGF-β), while both cell lines showed a similar signalling response (SMAD phosphorylation) to TGF-β, thus indicating that $G\alpha_q$ and/or $G\alpha_{11}$ is important for integrating migratory signals [19].

Subsequently these cells have been used in a number of other studies. Alvarez-Curto et al. [1] transfected such $G\alpha_q/G\alpha_{11}$-null cells and isolated clones stably expressing either wild type mouse free fatty acid receptor 4 (mFFA4) or a form of this receptor in which C-terminal amino acids that become phosphorylated in response to agonist activation in the wild type were altered from Ser/Thr to Ala [20]. The wild type receptor was activated in both parental and $G\alpha_q/G\alpha_{11}$-deficient cells as defined by agonist-induced phosphorylation of residues Thr^{347} and Ser^{350} within the receptor C-terminal tail. However, agonism could not generate elevated levels of either inositol phosphates or Ca^{2+} in the $G\alpha_q/G\alpha_{11}$-deficient cells [1]. More interestingly, although many GPCRs, including mFFA4, are able to enhance phosphorylation and activation of the ERK1/2 MAP kinases when expressed in parental HEK293 cells, this was completely lacking in the cells genome-edited to lack expression of $G\alpha_q$ and $G\alpha_{11}$ [1]. There is substantial debate as to the contribution of G protein- and non-canonical, potentially arrestin-mediated, signalling to this end-point following activation of

GPCRs. However, these results imply that, at least for mFFA4, signals that regulate ERK1/2 phosphorylation are dependent entirely on these G proteins. In support of this, in parental HEK293 cells mFFA4-mediated ERK1/2 MAP kinase phosphorylation was completely eliminated in the presence of an earlier described $G\alpha_q/G\alpha_{11}$ inhibitor **YM-254890** [21] (**Text Box 3**) that is closely related to FR900359. Further studies employing HEK293 cells genome-edited to lack expression of a broader range of G protein α subunits ($G\alpha_s$, $G\alpha_{olf}$, $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$ and $G\alpha_{13}$) have essentially confirmed, at least in this cell type, an absence of detectable GPCR-mediated arrestin signalling in the absence of receptor interaction with G protein α subunits [22]. The $G\alpha_q$ plus $G\alpha_{11}$ genome-edited cells have also been used to provide information to support the idea that a modified form of a muscarinic M_3 receptor, designed to retain the ability to signal via $G\alpha_q$ but to lose interactions with arrestin adaptor proteins, does indeed retain competence to activate this G protein [23]. Moreover, within a larger study designed to understand the interactions of the hallucinogen lysergic acid diethylamide (LSD) with the serotonin 5-HT_{2B} receptor Wacker et al. [24] expressed this receptor in both parental and $G\alpha_q/G\alpha_{11}$-null HEK293 cells to illustrate a minimal role for these G proteins on the rate of dissociation of LSD from the receptor. Although focussed largely on the molecular basis for why human GPRC6A is predominantly retained intracellularly when used in expression studies, Jørgensen et al. [25] also used the $G\alpha_q/G\alpha_{11}$-null genome-edited HEK293 cells to allow introduction of a previously characterised receptor-promiscuous mutant of $G\alpha_q$ ($Gly^{66}AspG\alpha_q$) [26] to define the relative signalling effectiveness of variants and mutants of GPRC6A.

Elimination of $G\alpha_s/G\alpha_{olf}$

The family of adenylyl cyclase stimulatory G proteins contains splice variants of $G\alpha_s$ that are expressed ubiquitously, as well as the highly related G protein $G\alpha_{olf}$ (olfactory) (**Figure 1**). Initially considered to be limited to the nasal neuroepithelium, $G\alpha_{olf}$ is now appreciated to have a wider expression pattern, particularly in regions of the brain. As noted earlier, Stallaert et al. [11] generated $G\alpha_s/G\alpha_{olf}$-deficient HEK293 cells and demonstrated that β_2-adrenoceptor-induced Ca^{2+} response was mediated by G_s family proteins.

As part of an extensive study into mechanisms linking the β_2-adrenoceptor to phosphorylation of ERK1/2 MAP kinases O'Hayre et al. [27] used HEK293 cells genome-edited to lack expression of both $G\alpha_s$ and $G\alpha_{olf}$. Lack of these G proteins all but eliminated the ability of the β_2-adrenoceptor agonist isoproterenol to induce ERK1/2 MAP kinase phosphorylation. Importantly, this was restored by re-introduction of $G\alpha_s$ into the cells [27]. As for FFA4 [1], expression of a form of the β_2-adrenoceptor that lacked key sites of agonist-induced phosphorylation, and that therefore interacts poorly with arrestins, resulted in greater levels of ERK1/2 phosphorylation in parental cells than was generated by the wild type β_2-adrenoceptor [27]. In concert with [1] and [22] these studies indicate that for distinct, well characterised GPCRs, even though they interact predominantly with different G protein classes, regulation of the ERK1/2 MAP kinases is largely, if not exclusively, G protein-dependent. Further studies that consider this question in cells lacking expression of arrestins will be reviewed later. HEK293 cells lacking $G\alpha_s/G\alpha_{olf}$ have also been used within a broader set of studies designed to assess the contribution of G_s-regulation of p70 S6 kinase activity within a pre-assembled β-arrestin/p70 S6 kinase/ribosomal protein S6 complex that enhances mRNA translation [28].

In efforts to assess conformational changes in activated G proteins and how this might correlate with agonist efficacy, Furness et al. [29] stably expressed the calcitonin receptor in these $G\alpha_s/G\alpha_{olf}$-null HEK293 cells. They then introduced transiently sets of G_s G protein

FRET-sensor pairings to focus on these, without receptor interactions also occurring with endogenously expressed G_s. By expressing such FRET sensors in the $G\alpha_s/G\alpha_{olf}$-null cells these studies showed the FRET sensors mediate cAMP response to various ligands in a manner similar to native G_s. When membranes from such cells were challenged with different agonists of the calcitonin receptor clear, if subtle, differences were noted in the organisation of the G protein complex when using either human or salmon calcitonin as agonist [29].

The *GNAS* gene that encodes $G\alpha_s$ can also generate a longer, N-terminally extended, variant described as $XL\alpha_s$. It was recently indicated that $XL\alpha_s$ can interact with various intracellular sorting proteins that are key contributors to clathrin-mediated endocytosis, including nexin-9 and dynamins [30]. Elimination of $XL\alpha_s$ from the osteocyte-like cell line Ocy454 using CRISPR/Cas9 increased internalisation of transferrin, whilst further studies indicated this to be a specific role for $XL\alpha_s$ rather than for conventional $G\alpha_s$ [30].

Elimination of $G\alpha_{12}/G\alpha_{13}$, or $G\alpha_{12}/G\alpha_{13}$ and $G\alpha_q/G\alpha_{11}$

As discussed earlier, key roles of $G\alpha_q$ and/ or $G\alpha_{11}$ are broadly understood. However, despite studies on knock-out mouse lines and the production and use of mouse embryo fibroblast cells derived from such animals, the roles of $G\alpha_{12}$ and /or $G\alpha_{13}$ are generally poorly understood [31, 32]. In significant part this reflects that activation of these G proteins does not directly regulate second messenger levels or other easy-to-measure and high-throughput endpoints. HEK293 cells express both $G\alpha_{12}$ and $G\alpha_{13}$ [5] (**Figure 1**) and genome-editing has been used to generate clones of HEK293 cells lacking both $G\alpha_{12}$ and $G\alpha_{13}$ [33, 34] and even lines lacking each of $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$, and $G\alpha_{13}$ [22, 34, 35, 36].

Ligand screening strategies for novel activators of GPCRs have often centred to trying to develop assays that can be utilised broadly, no matter the G protein-coupling preference and profile of the receptor. Approaches to this have included transfection alongside a receptor of

interest of the 'universal' G protein $G\alpha_{16}$ [37] and the co-transfection of a cocktail of **chimeric G protein α subunits** [38]. Recently a transforming growth factor α (TGFα) **shedding assay** was introduced [39] and shown to be able to report activation of a broad range of heterologously expressed GPCRs. Sergeev et al. [34] demonstrated, by using free fatty acid receptor 2 (FFA2), that the TGFα shedding assay is suitable if the receptor couples effectively to any combination of $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$, and/or $G\alpha_{13}$. This conclusion was based on the lack of agonist-induced shedding of TGFα when FFA2 was introduced into HEK293 cells that had been genome-edited to eliminate all four of these G protein α subunits, and the restoration of such activity with re-introduction of any one of these four G proteins into the genome-edited cells [34] (**Figure 2**). As FFA2 is also able to activate various pertussis toxin-sensitive G_i-family G proteins [40, 41] and a number of these are also expressed endogenously by HEK293 cells [5] (**Figure 1**), these studies also illustrated that activation of members of this G protein group is not able to induce TGFα shedding. The FFA2 allosteric agonist **AZ1729**, although able to produce a conformation of the receptor able to engage with G_i-family G proteins, is not able to do so for $G\alpha_q$/$G\alpha_{11}$ G proteins [41]. The extreme G_i-directed 'bias' of this ligand was now also shown to extend to a lack of capacity to activate $G\alpha_{12}$ and/or $G\alpha_{13}$ via FFA2 [34]. By contrast orthosteric agonists, including the short chain fatty acids that are the endogenous agonists for this receptor, allowed engagement with all of these G proteins subtypes [34]. Devost et al. [36] used these same quadruple G protein α subunit-deleted cells to show loss of coupling of various GPCRs known to engage with these G proteins to stimulate a serum response factor-based gene reporter construct. Once again, individual re-introduction of each of $G\alpha_q$, $G\alpha_{11}$, $G\alpha_{12}$, or $G\alpha_{13}$ restored responsiveness to intramolecular BRET-based biosensors of the angiotensin AT_1 receptor and in a manner dependent on the specific construction of the sensor [36].

Elimination of arrestins

Four members of the arrestin family play key roles in desensitisation of responses of GPCRs following initial agonist-induced receptor activation [42]. Although arrestin-1 and arrestin-4 have very limited expression profiles, both arrestin-2 and arrestin-3 (also still widely designated β-arrestin-1 and β-arrestin-2 respectively) are expressed ubiquitously, including by HEK293 cells [5] (**Figure 1**). When bound to an agonist-activated GPCR arrestins block access of G proteins to the receptor. In recent years much broader roles of arrestins have been uncovered in which they may initiate signal transduction events distinct from, and apparently not dependent upon, G protein activation [42, 43, 44]. Understanding the contributions and roles of arrestin-2 and arrestin-3 as initiators of signal transduction via distinct pathways is at the heart of the push to identify 'bias' ligands that may lead to the development of improved therapeutics [41, 42, 43]. Indeed, results from mouse knock-out models clearly implicate arrestins in roles beyond, and distinct from, their capacity to limit and desensitise G protein-mediated signalling [44, 45]. Although often assumed to be functionally redundant, arrestin-2 and arrestin-3 clearly fulfil distinct roles [45]. Genome editing has been used to eliminate the expression of both arrestin-2 and arrestin-3 in HEK293 cells [1, 27], and such arrestin-null cells have been central to a number of recent studies. As highlighted earlier, alongside cells genome-edited to lack both $G\alpha_q$ and $G\alpha_{11}$, [1] expressed both wild type and phosphorylation-deficient mouse FFA4 in HEK293 cells genome-edited to lack both arrestin-2 and arrestin-3. Although wild type FFA4 internalises extensively and rapidly from the surface of parental HEK293 cells upon addition of an agonist [46], occupation of either form of this receptor now failed to cause substantial internalisation in the arrestin-2 plus arrestin-3 null cells [1]. FFA4 agonist-mediated phosphorylation of the ERK1/2 MAP kinases was also preserved in the absence of arrestins [1]. Moreover, in both wild type and arrestin-null HEK293 cells much greater elevation of ERK1/2 MAP kinase

317 phosphorylation was produced by the phosphorylation-deficient mFFA4 receptor. This
318 indicates that this effect cannot simply reflect lack of arrestin-mediated desensitisation of the
319 receptor but must be linked intrinsically to its phosphorylation state [1]. As well as using the
320 same arrestin2/3-null HEK293 cells [27] recently additionally employed selective small
321 interfering RNA knockdown of each of arrestin-2 and arrestin-3 to indicate a key role for
322 specifically arrestin-3 in agonist-induced internalisation of the β_2-adrenoceptor. These
323 workers then extended this to studies that used **TALEN** (transcription activator–like effector
324 nuclease)-induced elimination of arrestin-2. Once again β_2-adrenoceptor-mediated ERK1/2
325 phosphorylation was enhanced, rather than decreased, in the complete absence of arrestins
326 [27]. By contrast, the ERK1/2 phosphorylation response was completely silenced in
327 $G\alpha_s/G\alpha_{olf}$-null cells [27] as mentioned previously. It is important to note, however, that
328 whilst [47] also noted enhanced ERK1/2 phosphorylation in arrestin2/3-null HEK293 cells
329 upon agonist activation of the A variant of the chemokine CXCR3 receptor, they rather
330 reported reduced activity via the B variant of this receptor. This suggests that there may be
331 substantial variation in mechanisms of regulation of ERK1/2 phosphorylation by different
332 GPCRs when expressed in the same cell background. The HEK293 cells genome-edited to
333 lack expression of both arrestin-2 and arrestin-3 have also been used to complement other
334 approaches. In studies of mutated variants of arrestin-3 that appeared to interact
335 spontaneously with the M_2 muscarinic acetylcholine receptor it was concluded that Lys^{139}Ile-
336 arrestin-3 was indeed able to interact with this receptor in an agonist-independent fashion
337 [48]. In a similar vein, [49] used HEK293 cells lacking both arrestin-2 and arrestin-3, but also
338 engineered additionally to lack expression of the endogenously expressed β_2-adrenoceptor, to
339 study functional outcomes of two conformations (tail and core) of GPCR-arrestin complexes.
340 By expressing a mutant arrestin that only forms a tail conformation alongside either a β_2-
341 adrenoceptor or a chimeric β_2-adrenoceptor/V_2 vasopressin receptor in these triple knock-out

cells, Cahill 3rd et al. [49] showed that this conformation suffices to recruit an arrestin effector molecule and to internalise ligand-activated receptors, whilst desensitisation of G protein signalling was not induced.

Concluding Remarks

Applications of genome-editing approaches, particularly in well studied and widely used model cell systems, have already provided important new insights into the relative roles of G protein- and non G protein-mediated signalling from variety of cell surface GPCRs. As with earlier methods able to 'knock-down' expression of gene products, the earliest described CRISPR/Cas9-based methods were recognised to have potential 'off-target' effects (**Text Box 4**) [50]. However, as methods have improved and developed this is now less of a concern (**Text Box 4**). As with many other revolutions in molecular biology, pharmacologists need to rapidly adopt and appreciate the power and potential of genome-editing. Ways in which it will influence understanding of the modes of action of drugs and tool compounds and of the mechanisms of signal transduction are already manifest (**Outstanding questions**), and its applications and potential seem limited only by the imagination of experimental scientists.

Acknowledgements

Work described in this article was supported by Biotechnology and Biosciences Research Council U.K. [grants BB/L027887/1 and BB/K019864/1)] (GM) and Japan Science and Technology Agency (JST) Grant Number JPMJPR1331, Japan Agency for Medical Research and Development (AMED) Grant Number JP17gm5910013 and Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 17K08264 (AI).

We thank Kouki Kawakami (Tohoku University, Japan) and Yu Hisano (Harvard Medical School, USA) for advice on designing of sgRNA sequences.

Authors Contributions

Both authors wrote the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interests

References

1. Alvarez-Curto, E. *et al.* (2016) Targeted Elimination of G Proteins and Arrestins Defines Their Specific Contributions to Both Intensity and Duration of G Protein-coupled Receptor Signaling. *J. Biol. Chem.* 291, 27147-27159
2. Regard, J.B., Sato, I.T., and Coughlin, S.R. (2008) Anatomical profiling of G protein-coupled receptor expression. *Cell* 135, 561-571
3. Amisten, S., Salehi, A., Rorsman, P., Jones, P.M. and Persaud, S.J. (2013) An atlas and functional analysis of G-protein coupled receptors in human islets of Langerhans. *Pharmacol. Ther.* 139, 359-391
4. Insel, P.A. *et al.* (2015) G Protein-Coupled Receptor (GPCR) Expression in Native Cells: "Novel" endoGPCRs as Physiologic Regulators and Therapeutic Targets. *Mol. Pharmacol.* 88, 181-187
5. Atwood, B.K., Lopez, J., Wager-Miller, J., Mackie, K., and Straiker, A. (2011) Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. *BMC Genomics* 7, 12:14
6. Layden, B.T., Newman, M., Chen, F., Fisher, A., and Lowe, W.L. Jr. (2010) G protein coupled receptors in embryonic stem cells: a role for Gs-alpha signaling. *PLoS One.* 5, e9105

7. Holdsworth, G., Slocombe, P., Hutchinson, G., and Milligan, G. (2005) Analysis of endogenous S1P and LPA receptor expression in CHO-K1 cells. *Gene.* 350, 59-63
8. Yin. H., *et al.* (2009) Lipid G protein-coupled receptor ligand identification using beta-arrestin PathHunter assay. *J. Biol. Chem.* 284, 12328-12338
9. Southern, C., *et al.* (2013) Screening β-arrestin recruitment for the identification of natural ligands for orphan G-protein-coupled receptors. *J Biomol Screen*. 18, 599-609
10. Galaz-Montoya, M., Wright, S.J., Rodriguez, G.J., Lichtarge, O., and Wensel, T.G. (2017) β_2-Adrenergic receptor activation mobilizes intracellular calcium via a non-canonical cAMP-independent signaling pathway. *J. Biol. Chem.* 292, 9967-9974
11. Stallaert, W. *et al.* (2017) Purinergic Receptor Transactivation by the β_2-Adrenergic Receptor Increases Intracellular Ca^{2+} in Nonexcitable Cells. *Mol. Pharmacol.* 91, 533-544
12. Tafesse, F.G., *et al.* (2014) GPR107, a G-protein-coupled receptor essential for intoxication by Pseudomonas aeruginosa exotoxin A, localizes to the Golgi and is cleaved by furin. *J. Biol. Chem.* 289, 24005-24018
13. Bostaille, N., Gauquier, A., Twyffels, L., and Vanhollebeke, B. (2016) Molecular insights into Adgra2/Gpr124 and Reck intracellular trafficking. *Biol. Open.* 5, 1874-1881
14. Bulanova, D.R. *et al.* (2016) Orphan G protein-coupled receptor GPRC5A modulates integrin β1-mediated epithelial cell adhesion. *Cell Adh. Migr.* 7, 1-13
15. Hou, P., *et al.* (2015) Genome editing of CXCR4 by CRISPR/cas9 confers cells resistant to HIV-1 infection. *Sci. Rep.* 5, 15577
16. Xiong, X.F., *et al.* (2016) Total synthesis and structure-activity relationship studies of a series of selective G protein inhibitors. *Nat. Chem.* 8, 1035-1041
17. Matthey, M., *et al.* (2017) Targeted inhibition of G_q signaling induces airway relaxation in mouse models of asthma. *Sci. Transl. Med.* 13, 9(407)

18. Schrage, R., *et al.* (2015) The experimental power of FR900359 to study Gq-regulated biological processes. *Nat. Commun.* 6:10156

19. Doçi, C.L., *et al.* (2017) Epidermal loss of Gαq confers a migratory and differentiation defect in keratinocytes. *PLoS One.* 12, e0173692

20. Prihandoko, R., *et al.* (2016) Distinct Phosphorylation Clusters Determine the Signaling Outcome of Free Fatty Acid Receptor 4/G Protein-Coupled Receptor 120. *Mol. Pharmacol.* 89, 505-520

21. Takasaki, J., *et al.* (2004) A novel Galphaq/11-selective inhibitor. *J. Biol. Chem.* 279, 47438-47445

22. Grundmann, M., *et al*. (2018) Lack of beta-arrestin signalling in the absence of active G proteins *Nat Commun*. 9:341

23. Hu, J., *et al.* (2016) A G Protein-biased Designer G Protein-coupled Receptor Useful for Studying the Physiological Relevance of Gq/11-dependent Signaling Pathways. *J. Biol. Chem.* 291, 7809-7820

24. Wacker, D., *et al.* (2017) Crystal Structure of an LSD-Bound Human Serotonin Receptor. *Cell.* 168, 377-389

25. Jørgensen, S., *et al.* (2017) Genetic Variations in the Human G Protein-coupled Receptor Class C, Group 6, Member A (GPRC6A) Control Cell Surface Expression and Function. *J. Biol. Chem*. 292, 1524-1534

26. Heydorn, A., *et al.* (2004) Identification of a novel site within G protein alpha subunits important for specificity of receptor-G protein interaction. *Mol. Pharmacol.* 66, 250-259

27. O'Hayre, M., *et al.* (2017) Genetic evidence that β-arrestins are dispensable for the initiation of β_2-adrenergic receptor signaling to ERK. *Sci. Signal.* 10 (484)

28. Tréfier, A., *et al.* (2017) G protein-dependent signaling triggers a β-arrestin-scaffolded p70S6K/ rpS6 module that controls 5'TOP mRNA translation. *FASEB J.* 2017 Oct 30. pii: fj.201700763R. doi: 10.1096/fj.201700763R

29. Furness, S.G., *et al.* (2017) Ligand-Dependent Modulation of G Protein Conformation Alters Drug Efficacy. *Cell.* 167, 739-749

30. He, Q., *et al.* (2017) Large G protein α-subunit XLαs limits clathrin-mediated endocytosis and regulates tissue iron levels in vivo. *Proc. Natl. Acad. Sci. U S A.* 114, E9559-E9568

31. Worzfeld, T., Wettschureck, N., and Offermanns, S. (2008) G(12)/G(13)-mediated signalling in mammalian physiology and disease. *Trends Pharmacol. Sci.* 29, 582-589

32. Kozasa, T., Hajicek, N., Chow, C.R., and Suzuki, N. (2011) Signalling mechanisms of RhoGTPase regulation by the heterotrimeric G proteins G12 and G13. *J. Biochem.* 150, 357-369

33. O'Hayre, M., *et al.* (2016) Inactivating mutations in GNA13 and RHOA in Burkitt's lymphoma and diffuse large B-cell lymphoma: a tumor suppressor function for the Gα13/RhoA axis in B cells. *Oncogene*. 35, 3771-3780

34. Sergeev, E., *et al.* (2017) A single extracellular amino acid in Free Fatty Acid Receptor 2 defines antagonist species selectivity and G protein selection bias. *Sci. Rep.* 7:13741

35. Sleno, R., *et al.* (2017) Conformational biosensors reveal allosteric interactions between heterodimeric AT1 angiotensin and prostaglandin F2α receptors. *J. Biol. Chem.* 292, 12139-12152

36. Devost, D., *et al.* (2017) Conformational Profiling of the AT1 Angiotensin II Receptor Reflects Biased Agonism, G Protein Coupling, and Cellular Context. *J. Biol.Chem.* 292, 5443-5456

37. Milligan, G., Marshall, F., and Rees, S. (1996) G16 as a universal G protein adapter: implications for agonist screening strategies. *Trends Pharmacol. Sci.* 17, 235-237

38. Milligan, G., and Kostenis, E. (2006) Heterotrimeric G-proteins: a short history. *Br. J. Pharmacol.* 147 Suppl 1:S46-55

39. Inoue, A. *et al.* (2012) TGFα shedding assay: an accurate and versatile method for detecting GPCR activation. *Nat. Methods.* 9, 1021-1029

40. Brown, A.J. *et al.* (2003) The Orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *J. Biol. Chem.* 278, 11312-11319

41. Bolognini, D., *et al.* (2016) A Novel Allosteric Activator of Free Fatty Acid 2 Receptor Displays Unique Gi-functional Bias. *J. Biol. Chem.* 291, 18915-18931

42. Smith, J.S., and Rajagopal, S. (2016) The β-Arrestins: Multifunctional Regulators of G Protein-coupled Receptors. *J. Biol. Chem.* 291, 8969-8977

43. Peterson, Y.K., and Luttrell, L.M. (2017) The Diverse Roles of Arrestin Scaffolds in G Protein-Coupled Receptor Signaling. *Pharmacol. Rev.* 69, 256-297

44. Jean-Charles, P.Y., Kaur, S., and Shenoy, S.K. (2017) G Protein-Coupled Receptor Signaling Through β-Arrestin-Dependent Mechanisms. *J. Cardiovasc. Pharmacol*. 70, 142-158

45. Schmid, C.L., and Bohn, L.M. (2009) Physiological and pharmacological implications of beta-arrestin regulation. *Pharmacol. Ther.* 121, 285-293

46. Hudson, B.D., *et al.* (2013) The pharmacology of TUG-891, a potent and selective agonist of the free fatty acid receptor 4 (FFA4/GPR120), demonstrates both potential opportunity and possible challenges to therapeutic agonism. *Mol. Pharmacol.* 84, 710-725

47. Smith, J.S., *et al.* (2017) C-X-C Motif Chemokine Receptor 3 Splice Variants Differentially Activate Beta-Arrestins to Regulate Downstream Signaling Pathways. *Mol. Pharmacol*. 92, 136-150

48. Prokop, S., *et al.* (2017) Differential manipulation of arrestin-3 binding to basal and agonist-activated G protein-coupled receptors. *Cell. Signal.* 36, 98-107

49. Cahill, T.J. 3rd, *et al.* (2017) Distinct conformations of GPCR-β-arrestin complexes mediate desensitization, signaling, and endocytosis. *Proc. Natl. Acad. Sci. U S A.* 114, 2562-2567

50. Ran, F.A., *et al.* (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281-2308

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

52. Maquat, L.E. *et al.* (2004) Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol*. 5, 89–99

53. Shukla, A.K. *et al.* (2013) Structure of active β-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. *Nature.* 497, 137-141

54. Stepanenko, A.A., *et al.* (2015) HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene*. 569, 182-190

55. Ran, R.A., *et al.* (2013 Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 154, 1380-1389

56. Kleinstiver, B.P., *et al.*, (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 529, 490-495

57. Slaymaker, I.M., *et al*., (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science.* 351, 84-88

58. Chen, J.S., *et al.*, (2017) Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature.* 550, 407-410

59. Kim, J., *et al.*, (2005) Functional antagonism of different G protein-coupled receptor kinases for beta-arrestin-mediated angiotensin II receptor signalling. *Proc Natl Acad Sci U S A.* 102, 1442-1447

Text Boxes

Abbreviations used in Text Boxes

DSB, double-strand break

FACS, fluorescence activated cell sorting

NHEJ, non-homologous end joining

PAM, protospacer-associated motif

RE, restriction enzyme

sgRNA, single guide RNA

Text Box 1

A CRISPR/Cas9 system utilises a RNA-guided endonuclease derived from bacteria or archaea [50]. The RNA loaded in a Cas9 endonuclease (CRISPR RNA, crRNA) guides the Cas9 complex to complementary sequence of target DNA and allow the Cas9 endonuclease to induce a double-strand break (DSB) in this DNA. Cleavage of target DNA by Cas9-crRNA requires a short nucleotide sequence called a protospacer-associated motif (PAM), flanking the 3´ site of the crRNA complementary sequence. *Streptococcus pyogenes* Cas9, the most routinely used Cas9, recognises 5'-NGG-3' PAM. When applied to mammalian cells, expression of two components is sufficient to induce a DSB: Cas9 that contains a nuclear localisation signal and a single guide RNA (sgRNA) that mimics function of crRNA. After induction of a DSB, host cells repair the cleavage and random insertion/deletion is introduced by a non-homologous end joining (NHEJ) pathway. This results in disruption (insertion, deletion or truncation) of an encoded protein.

Effective targeting of a gene locus depends on functional domains/motifs of its encoded protein. For G protein α subunits amino acid residues critical for effective interactions with GPCRs are located within the extreme C-terminal region [51]. Thus, a frame-shift mutation at any position in the protein open-reading frame is expected to nullify G protein function due to a lack of GPCR-induced activation. A frame-shift mutation usually leads to the emergence of a premature stop codon. When this type of stop codon appears 5' upstream of the last exon the mRNA undergoes nonsense-mediated mRNA decay [52]. This results in smaller amounts of a truncated protein than of the full length native protein. As a NHEJ repair can introduce a random, in-frame mutation, it is preferable to design a sgRNA sequence to target a site encoding a residue critical for protein function. For example, nucleotide-interacting residues, and flexible loops (Finger loop, Middle loop and Lariat loop) are critical for function of G protein α subunits (GDP/GTP binding) and arrestins (GPCR binding) [53], respectively. For GPCRs, highly conserved sequence amino acid motifs (e.g. Asp-Arg-Tyr (DRY) at the cytoplasmic face of transmembrane domain 3 and (Asn-Pro-Xaa-Xaa-Tyr (NPxxY) in transmembrane domain 7) are key targets. An in-frame mutation near these residues is expected to ablate protein function. Another potentially effective target site is a splicing acceptor/donor sequence. Since RNA splicing occurs at conserved sequences (GG-GU and AG-G (hyphen denotes a cleavage site) for a donor and an acceptor, respectively), disruption of these nucleotides can result in skipping of proper RNA splicing and loss of mature protein.

Text Box 2

Although widely used for pharmacological studies and able to be transfected with high efficiency HEK293 cells are known to be polyploid (i.e. cells contain more than two full copies of the genome) [54]. Although the precise karyotype varies among sub-clones and probably between individual laboratories HEK293 cells are roughly triploid. From a genome-editing standpoint polyploidy is challenging because all of the alleles encoding the gene of interest must be modified and the probability of obtaining an 'all-allele edited' clone (preferably all with frame-shift mutations) is exponentially less likely as numbers of target alleles or target genes increase. Thus, it is essential to design and validate effective and easy-to-genotype sgRNA constructs. Various online tools to support sgRNA design (minimising scores for potential 'off-target' effects (**Text Box 4**) and increasing on-target scores) are available (for example, CRISPR Design, http://crispr.mit.edu; CRISPRdirect, https://crispr.dbcls.jp; CRISPRko, https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). When multiple-gene targeting is performed, we design three sgRNAs per gene and test mutagenesis activity in a pool of sgRNA/Cas9-expressing cells so that we can select and mix the most highly active sgRNAs to be transfected. For genotyping, we prefer a restriction enzyme (RE)-based method (**Figure I**), in which the RE recognises an unmodified sequence at a DSB but fails to digest it once a random mutation is introduced at the RE site. Compared with other genotyping methods (e.g. mismatched hybrid-recognising nuclease and heteroduplex mobility analysis), the RE-based method is sensitive for discriminating all-allele mutations from heterozygous mutation, although numbers of candidates sgRNA sequences may be quite limited. Since HEK293 cells carry on average three alleles per gene direct sequencing (e.g. Sanger sequencing of a PCR-amplified fragment) is often not feasible because electrographs with more than two waves (distinct mutations) are impractical to deconvolute. We prioritise candidate clones by utilising capillary electrophoresis which has

587 resolution of a few base pairs and can discriminate clones with a single PCR band from those
588 with multiple bands. Clones with a single band are more likely to carry an identical mutation
589 in two or more alleles, in which case sequence can be determined by a direct sequencing
590 method. It is important to validate loss of the targeted gene products in isolated clones as
591 sequencing alone sometimes fails to detect a hypomorphic (functionally residual) allele. Such
592 validation assays include Western blot analysis (protein expression) and functional GPCR
593 assays to monitor downstream signalling. Notably, it is essential to perform re-expression
594 experiments (recovery of an observed phenotype with re-introduction of the targeted gene) to
595 avoid artefacts associated with 'off-target' effects of Cas9 genome-editing (**Text Box 4**).
596 From our experience a highly active sgRNA construct is able to generate an all-allele, frame-
597 shift mutant with probability of 10-20% in HEK293 cells. When two genes are
598 simultaneously targeted, this probability is reduced to approximately 1-4%.

599

Text Box 3

Until recent times, beyond gene knock-out strategies in both mice and less complex model organisms, the only practical way to inactivate members of a G protein α subunit subfamily was via the ADP-ribosyltransferase activity of pertussis toxin [38](**Figure II**). Of the G_i subgroup the α subunits of all bar G_z have a Cys residue located 4 amino acids from the C-terminus and pertussis toxin catalyses addition of ADP-ribose to this amino acid [38]. This prevents effective interactions with agonist-occupied GPCRs. Initial introduction of the cell permeable $G_q/G_{11}/G_{14}$ inhibitor YM-254890 [21] (**Figure II**) allowed selective targeting of this class of G proteins. However, YM-254890 is not an effective inhibitor of the GPCR promiscuous G_{15}/G_{16} G proteins from rodent (G_{15})/human (G_{16}) (but these display very limited tissue distribution). Subsequent discovery of the related molecule FR900359 [18], with similar characteristics and G protein selectivity as YM-254890 (**Figure II**), has expanded the tools available. Further G protein subtype selective inhibitors would be exceptionally useful additions but genome-editing offers an alternative and potentially more broad-ranging approach with the potential to allow elimination of multiple G protein α subunits and across subfamilies (see **Main text** for details).

Text Box 4

Although Cas9-induced genome editing is efficacious it is known that Cas9 causes mutations at 'off-target' sites that have sequence similarity with a sgRNA. As such it is important that experimental artefacts associated with potential 'off-target' mutations should be considered. To validate that an observed cellular response in genome-edited cells is mediated by the targeted gene product and not by an 'off-target' effect, it is essential to perform re-expression analysis (see **Main text** for examples). For example, if a GPCR-induced cellular response is lacking in cells from which a G protein subunit has been targeted, recovery of function or phenotype should occur following re-introduction of a plasmid able express the targeted polypeptide. Although in many cases transient expression of such a plasmid is sufficient to rescue a phenotype, it should also be considered whether appropriate and sufficient expression levels are achieved by so doing so.

Unlike genetically modified organisms (e.g. knockout mice), in which 'off-target' mutations can be removed by back-crossing with a wild-type animal, mutations in cell lines are not easily overcome. Furthermore, it is impractical to identify effectively all potential 'off-target' mutations in genome-edited cell lines unless accurate genome sequence is performed. Thus, it is important to avoid the generation of such 'off-target' mutations. Several high-fidelity Cas9 variants have been generated recently that are useful for reducing off-site genome modifications. For example, Cas9 nickase [55] induces a nick at the targeted site and DSB occurs only when two Cas9 nickases target a proximal site in complementary and non-complementary DNA sequences. This requires simultaneous binding of two Cas9-sgRNA complexes and thus reduces substantially the probability of non-specific DSB. Other examples are Cas9 variants [56, 57, 58] designed using rational structural analysis and these have reduced non-specific DNA contacts while maintaining on-target activity. With ongoing

641 efforts next-generation, high-fidelity Cas9 variants will pave a way to generate gene-
642 knockout cells free from off-target-effects.
643

Figure Legends (Main Text)

Figure 1

GPCR effectors and their expression by HEK293 cells

As well as a broad array of GPCRs [5] parental HEK293 cells express at detectable levels the α subunits of members of each of the four broad subgroups [38] of heterotrimeric G proteins. Gα subunits (and other GPCR regulators) in dark colours are not expressed at detectable levels. In addition HEK293 cells also express both arrestin-2 and arrestin-3. Although outwith the topics covered in this review HEK293 cells are also known to express some members of the families of G protein-coupled receptor kinases (GRK)(GRKs 2, 3, 5 and 6) [59] and Regulators of G protein Signaling (RGS) proteins that modify and regulate G protein-mediated signalling cascades. Although $G\alpha_z$ is not a substrate for pertussis toxin-catalysed ADP-ribosylation, all the other G_i-family Gα subunits are. Because this modification prevents GPCR-mediated activation of these G proteins and, therefore, can be used to chemically eliminate signalling via this group, these have not been reported, to date, to have been eliminated in HEK293 cells by genome-editing. By contrast, members of each of the other three subfamilies have been eliminated, sometimes in combinations, from HEK293 cells by CRISPR/Cas9-mediated genome-editing. Alongside HEK293 cells from which both arrestin-2 and arrestin-3 have been eliminated, these cell lines have provided a novel toolbox to explore roles of these polypeptides (see **Main text** for details).

Figure 2

Genome-editing of HEK293 cells illustrates the contribution of $G\alpha_q$ and $G\alpha_{12}$ G proteins family subunits to GPCR-mediated shedding of TGFα

Shedding of cell surface expressed alkaline phosphate-linked TGFα (**yellow stick**) from HEK293 cells in a ADAM-17 (**purple**)-dependent manner was introduced as a widely applicable means to identify agonists able to activate a very broad range of GPCRs [39] in wild type HEK293 cells. Initial studies suggested that this endpoint required activation of some combination of $G\alpha_{q}$ and $G\alpha_{12}$ family G protein subunits (**A**). Genome editing to eliminate expression of either both $G\alpha_{12}$ and $G\alpha_{13}$ (**B**) or $G\alpha_{q}$ and $G\alpha_{11}$ (**C**) was insufficient to eliminate this endpoint when activating co-expressed Free Fatty Acid Receptor 2. However, elimination of all four of these G protein α subunits abrogated shedding of TGFα (**D**). Receptor function was restored by re-introduction of any of these four individual G proteins (**E**) or by re-introduction of various chimeric G proteins (**F**). The cartoon is derived from data presented in [34].

Figure I (within Text Box 2)

Generation of HEK293 cells lacking expression of G protein α subunits using CRISPR/Cas9 genome editing

The steps described in Text Box 2 to perform and validate successful genome-editing of HEK293 cells to eliminate expression of the $G\alpha_{12}$ G protein subunit are illustrated. As highlighted in the text, HEK293 cells are polyploid and therefore in the illustration three alelles are indicated, with successful targeting of all three being required to eliminate expression. **A**. Illustrates steps to target a specific gene whilst **B**. and **C**. illustrate steps required to verify isolation of a successfully genome-edited clone.

Figure II (within Text Box 3)

Currently available G protein inhibitors

The identity and target selectivity of currently available inhibitors of various G proteins classes is illustrated. This highlights the dearth of such inhibitors and the potential of genome-editing to overcome this lack of effective tool compounds.

Glossary

Agonist: a chemical species which binds to a receptor and activates it, eliciting a biological response.

Antagonist: a chemical species which binds to a receptor preventing activation by an agonist.

AZ1729: *N*-[3-(2-carbamimidamido-4-methyl-1,3-thiazol-5-yl)phenyl]-4-fluorobenzamide. An allosteric agonist of the Free Fatty Acid 2 receptor. This ligand displays capacity to produce signals via this GPCR that proceed via activation of G_i-family G proteins but not via G_q/G_{11} or G_{12}/G_{13} G proteins [34, 41].

Chimeric G protein α subunit: a G protein α subunit containing elements of sequence from a minimum of two different G proteins. In most examples the C-terminal 5-10 amino acids of one G protein are exchanged for the corresponding sequence from a second G protein [38].

CRISPR/Cas9: Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes. A method for genome-editing (see **Text Box 1** for details).

FR900359: [(1R)-1-[(3S,6S,9S,12S,18R,21S,22R)-21-acetamido-18-benzyl-3-[(1R)-1-methoxyethyl]-4,9,10,12,16-pentamethyl-15-methylidene-2,5,8,11,14,17,20-heptaoxo-22-propan-2-yl-1,19-dioxa-4,7,10,13,16-pentazacyclodocos-6-yl]-2-methylpropyl] (2S,3R)-3-hydroxy-4-methyl-2-(propanoylamino)pentanoate. A synthetic inhibitor of G_q, G_{11} and G_{14} G proteins [16, 17, 18]**.**

G Protein-Coupled Receptor (GPCR): integral membrane protein able to transmit signals across the cell membrane in response to an external stimulus, to activate (usually) heterotrimeric guanine nucleotide binding proteins (G proteins). GPCRs are characterised by a conserved structure comprising 7 helical membrane spanning domains, an extracellular N-terminus, 3 external loops, 3 internal loops and an internal C-terminal domain. Sub-divided into classes based upon structural relatedness.

Knock-down: Introduction into cells of an oligonucleotide sequence that is complementary to specific mRNA sequences. This results in destruction of the corresponding mRNA. By preventing new protein synthesis from the mRNA this, over time, results in reduction in level of the corresponding protein.

Knock-out: a model organism (frequently mouse) in which one or more specific genes have been modified such that expression of corresponding mRNA/protein is lacking.

Null background: cells, frequently model systems used for ligand screening, that are known not to express endogenously the molecular target of interest.

Orphan (GPCR): A receptor for which endogenous activating ligands remain unidentified or where suggested ligands are not fully accepted by the research community.

Shedding assay: Release of an alkaline phosphatase-linked form of transforming growth factor α (TGFα) from the surface of transfected cells in a ligand (usually GPCR agonist)-dependent manner [39].

TALEN: Transcription activator-like effector nucleases. An alternate approach to genome-editing based on restriction enzymes engineered to cut specific sequences of DNA.

YM-254890: A cyclic peptide isolated initially from the soil bacterium *Chromobacterium* sp. QS3666 strain that selectively inhibits G_q, G_{11} and G_{14} G proteins [21].

Figure 1

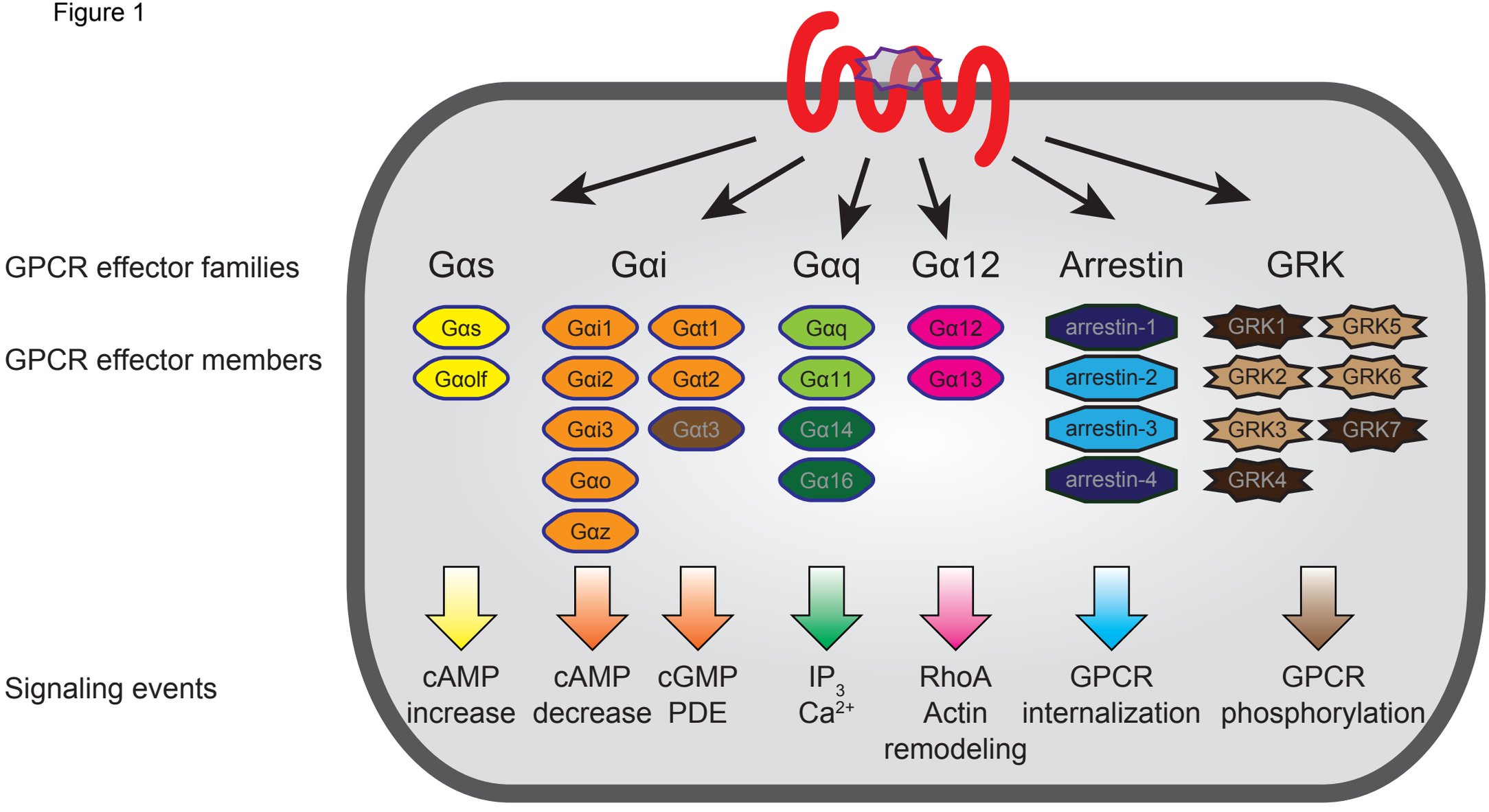

GPCR effector families
GPCR effector members
Signaling events
Gαs
Gαi
Gαq
Gα12
Arrestin
GRK
Gαs
Gαolf
Gαi1
Gαi2
Gαi3
Gαo
Gαz
Gαt1
Gαt2
Gαt3
Gαq
Gα11
Gα14
Gα16
Gα12
Gα13
arrestin-1
arrestin-2
arrestin-3
arrestin-4
GRK1
GRK5
GRK2
GRK6
GRK3
GRK7
GRK4
cAMP increase
cAMP decrease
cGMP PDE
IP3 Ca2+
RhoA Actin remodeling
GPCR internalization
GPCR phosphorylation

Figure 2

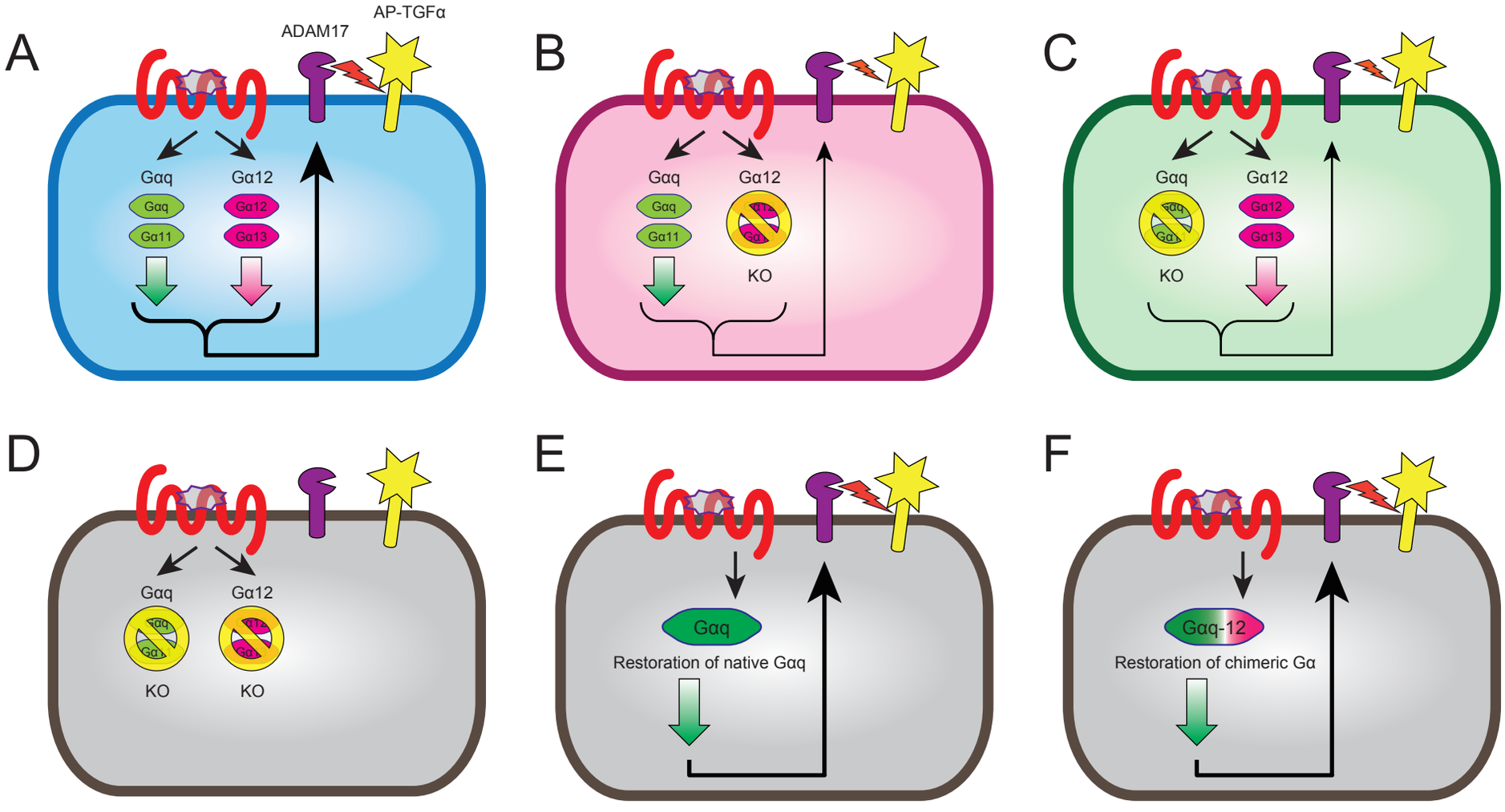

A
ADAM17
AP-TGFα
Gαq
Gα12
Gαq
Gα11
Gα12
Gα13
B
Gαq
Gα12
Gαq
Gα11
KO
C
Gαq
Gα12
KO
Gα12
Gα13
D
Gαq
Gα12
KO
KO
E
Gαq
Restoration of native Gαq
F
Gαq-12
Restoration of chimeric Gα

Figure I

A

sgRNA target construction (~1 weeks)

Designing sgRNA seqeunces (typically, three constructs per gene)

↓

Ordering synthelized oligonucleotides

↓

Inserting an oligonucletide duplex (sgRNA) into a Cas9-GFP-encoded plasmid

↓

Sequencing the inserted sgRNA

Single clone isolation and genotype screening (~3 weeks)

Transfecting a sgRNA/Cas9 plasmid

↓

FACS-isolating GFP-positive cells (sgRNA/Cas9-transfected cells)

↓

Seeding FACS-isolated cells in a 96-well plate in a limiting dilution (1 cell per well for 10 plates and 2 cells per well for 10 plates)

↓

Adding medium at least every 1 week

↓

Checking colony growth

↓

Passaging and genotyping colonies
Splitting harvested cells into a 6-well plate and a microtube

Validation of gene-knockout clones and a re-expression experiment (~3 weeks)

Passaging and splitting cells into a 10-cm dish (for stock) and desired dishes (assays)

↓ 10-cm dish

Harvesting and freezing cells

assay dishes

Western blot (protein expression)
Downstream signaling (protein function)
TA cloning (target allele sequence)

↓

Passaging cells into assay dishes

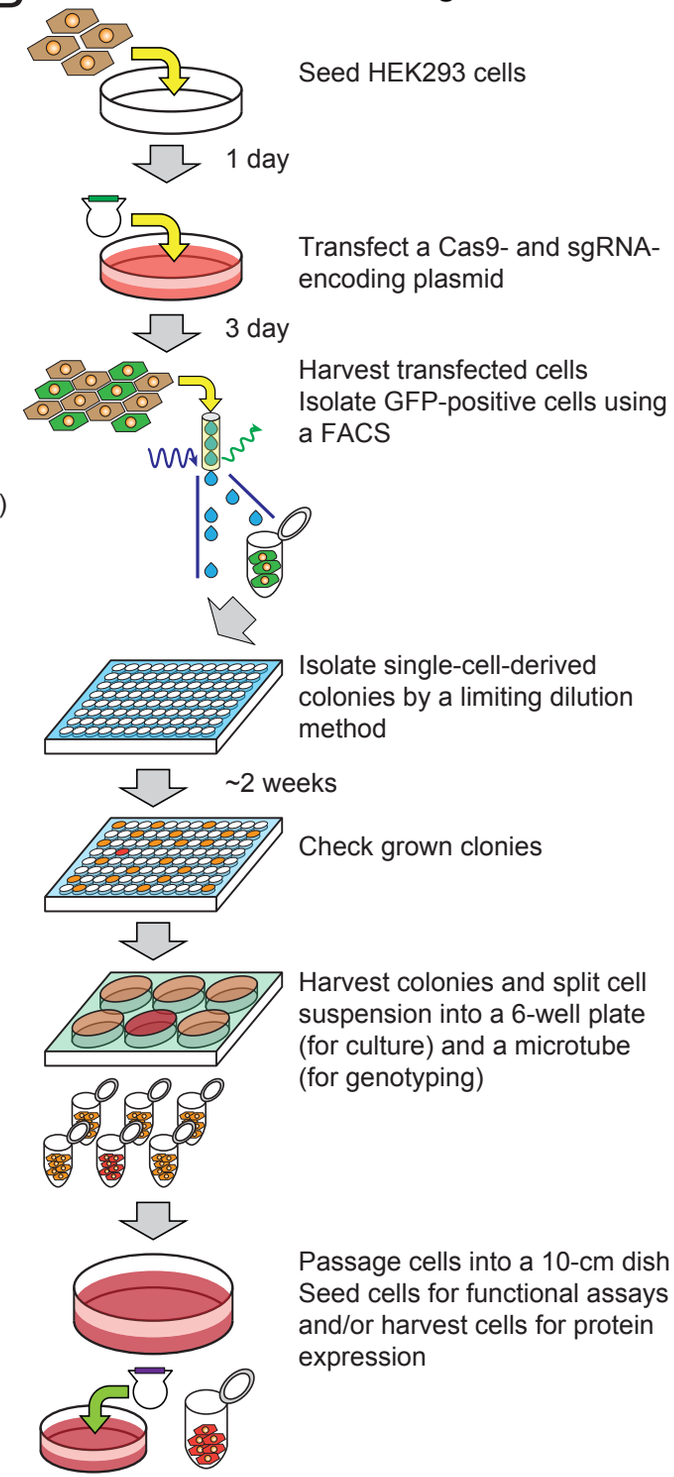
↓

Transfecting an expression plasmid encoding an ORF of the targeted gene

↓

Testing phenotypic “rescue” upon re-expression of the targeted gene

B



C

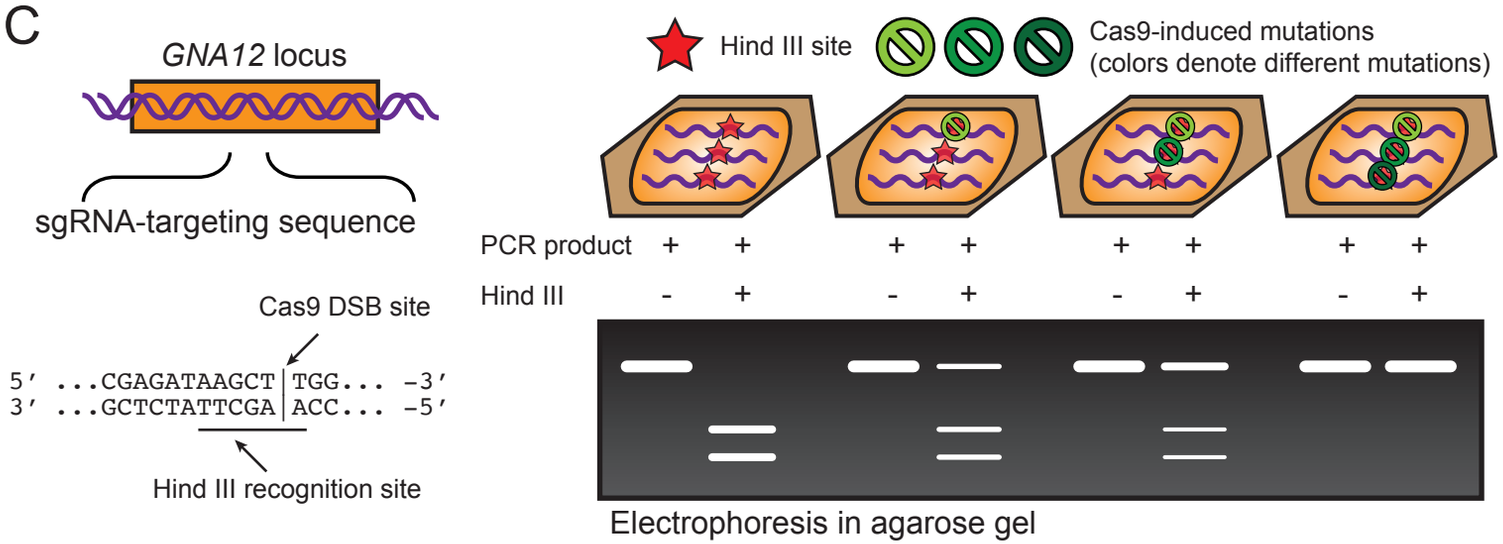


Figure II

General issues with G protein inhibitors

- Inhibitors affect only subsets of G protein subfamily members
- Washout and functional recovery/rescue experiments are difficult to perform

G protein inhibitors	Targets	Disadvantages/concerns
Pertussis toxin	G_i family $G\alpha$ subunits (except for G_z)	Long-term treatment is required (typically overnight)
YM-254890, FR900359 and analogues	G_q family inhibitors (except for $G_{15/16}$)	Compound may be rapidly degraded

Outstanding Questions

Can cell lines be developed that lack expression of all G protein α subunits?

As HEK293 cells express detectable amounts of 13 distinct G protein α subunits (**Figure 1**) the only practical means to approach total elimination is in a step-wise fashion. Given the key roles G proteins play in transmembrane signalling, elimination of further family members may result in very poor growth characteristics of such cells.

Elimination of expression of G protein β and/or γ subunits?

Given the complexity of the potential number of distinct G β/γ pairings and that older studies using antisense-based 'knock-down' suggested distinct functions for many such pairs, the generation of such lines and their analysis would be a substantial task.

Elimination of expression of G protein-coupled receptor kinases?

The contribution of members of the G protein-coupled receptor-(GRK) (**Figure 1**) and second-messenger-regulated kinases to ligand-induced phosphorylation of GPCRs has been controversial and challenging to assess, due in part to a lack of useful and selective inhibitors of GRK subtypes. Genome-editing would appear to offer a means to (re)-assess these issues.

Use of genome-editing to introduce point mutations/polymorphisms into GPCRs and signalling pathway components?

Although transfection of cDNAs mutated to encode point-mutants and non-synonomous single nucleotide polymorphisms is a standard approach, this frequently results in high level over-expression of the variant protein.

Introduction of such variants into cells that endogenously express a receptor or signalling protein of interest via genome-editing may ensure more native expression levels in the absence of the wild type protein sequence.