[7] High-Throughput Confocal Microscopy for β-Arrestin–Green Fluorescent Protein Translocation G Protein-Coupled Receptor Assays Using the Evotec Opera

By RALPH J. GARIPPA, ANN F. HOFFMAN, GABRIELE GRADL, and ACHIM KIRSCH

Abstract

Ligand-activated G protein-coupled receptors (GPCRs) are known to regulate a myriad of homeostatic functions. Inappropriate signaling is associated with several pathophysiological states. GPCRs belong to a ∼800 member superfamily of seven transmembrane-spanning receptor proteins that respond to a diversity of ligands. As such, they present themselves as potential points of therapeutic intervention. Furthermore, orphan GPCRs, which are GPCRs without a known cognate ligand, offer new opportunities as drug development targets. This chapter describes a systems-based biological approach, one that combines *in silico* bioinformatics, genomics, high-throughput screening, and high-content cell-based confocal microscopy strategies to (1) identify a relevant subset of protein family targets, (2) within the therapeutic area of energy metabolism/obesity, (3) and to identify small molecule leads as tractable combinatorial and medicinal chemistry starting points. Our choice of screening platform was the Transfluor β-arrestin–green fluorescent protein translocation assay in which full-length human orphan GPCRs were stably expressed in a U-2 OS cell background. These cells lend themselves to high-speed confocal imaging techniques using the Evotec Technologies Opera automated microscope system. The basic assay system can be implemented in any laboratory using a fluorescent probe, a stably expressed GPCR of interest, automation-assisted plate and liquid-handling techniques, an optimized image analysis algorithm, and a high-speed confocal microscope with sophisticated data analysis tools.

Choice of Orphans

Introduction

G protein-coupled receptors (GPCRs) are a superfamily of seven transmembrane proteins that have long been established as fertile targets for exploitation within pharmaceutical programs for the development of new

medicinals. In the decade from 1993 to 2002, compounds that had a mode of action relating to direct action on a GPCR (agonism, antagonism, inverse agonism, positive or negative modulation), for marketed compounds and for phase III entities, consistently occupied 23 to 27% of all target types addressed (Gong and Sjogren, personal communication, 2004). This list of target family proteins includes such well-known members such as kinases, proteases, ion channels, and nuclear receptors. GPCRs are involved in numerous physiological processes, and the ligands to these receptors have shown themselves to be clever mechanistic tools and often useful therapeutics. The entire family of known GPCRs has demonstrated an affinity for an unusually wide variety of ligands, from low molecular weight trace amines, to small peptides (<10 amino acids), large peptides (>10 amino acids), even to large phospholipids. Taking into account all of the seven transmembrane sequences identified in the human genome project, it is estimated that there are approximately 600 to 800 GPCRs, of which only 400 to 500 have known ligands. Therefore, the GPCRs with unknown ligands, or orphan GPCRs (oGPCRs), present a significant opportunity for investigators not only to elucidate their involvement in key biological processes, but also to mine them as drug targets. Characterization of oGPCRs would present opportunities for claiming intellectual property and also for developing first-to-market medicinals. However, oGPCR research can be hampered by the lack of literature-based data, making informed decisions on which oGPCR is most relevant within a disease area difficult, and by assay development issues.

We have taken a systems-based approach to identify a set of related oGPCR targets and small molecules with selective activity on these targets. To this end, we have assembled a diverse team of bioinformaticians, cell biologists, and peptide and small molecule chemists. As part of a general strategy for the team, we saw the need to incorporate bioinformatics, genomic, and combinatorial chemistry thinking in the process of generating new targets, leads, reagents, and expertise for a focused drug discovery effort. For example, an extensive list of anabolic and catabolic signaling molecules are known to exert their effect through binding of cell surface GPCRs in central nervous system effector pathways involved in the hypothalamic control of energy balance (Kieffer and Habener, 1999; Shimada *et al.*, 1998). Current pharmacological targets for class A and B known GPCRs include neuropeptide Y, melanocortin, orexin, and bombesin. However, if one examines the clustering of known GPCRs involved in energy balance, the clades within the genomic dendrogram reveal a number of oGPCR sequences with significant homology, that is, with cognate ligands as yet undiscovered. Since it was not technically feasible for us to screen hundreds of potential oGPCR targets for active compounds, as a first phase of the project we utilized an *in silico* bioinformatics approach to focus and filter

the list of potential target candidates. A second phase involved checking these oGPCR sequences for expression in normal human and mouse hypothalamic tissues. A mouse homologue is important to move a compound forward in lead optimization in that rodent models of obesity and energy metabolism are prevalent and accepted within the field. Single target expression profiling (STEP) was accomplished by examining the oGPCR of interest, relevant to GAPDH expression, across ∼50 organ-specific tissue samples. This procedure is well established for estimating the relative expression levels of a particular protein throughout body tissue sample in relatively high-throughput manners (Reidhaar-Olson *et al.*, 2003). In evaluating candidate oGPCRs for this study we identified many with a ubiquitous expression pattern across all tissues queried. For oGPCR38, however, the expression of the receptor was more limited with expression identified in the brain and hypothalamic regions, fulfilling preset criteria for oGPCRs of interest. A handful of candidate oGPCRs fulfilled the criteria of exclusive and specific expression patterns within the brain and hypothalamic regions. Our third and final phase of oGPCR selection involved determining the differential expression of the oGPCR of interest within discrete anatomical nuclei of defined brain regions harvested from diet-induced obese (DIO) mice as compared to nonobese rodents. For this purpose, we employed laser capture microdissection techniques to extract RNA from numerous mouse brain (Kinnecom and Pachter, 2005). The regions of interest included the arcuate nucleus, the paraventricular nucleus, medial amygdala, dentate gyrus, and the anterior pituitary. As precedence, many of the known obesity-related GPCRs are differentially regulated in discrete brain regions (Ludwig *et al.*, 2001). We used quantitative polymerase chain reaction (PCR) methods in order to estimate the degree of up- or downregulation of the oGPCRs of interest (Reidhaar-Olson *et al.*, 2002). Without a priori knowledge of the oGPCR involvement with either the anabolic or the catabolic arm of central energy regulation, we examined both possibilities. At the end of our three-phase selection process, we were able to distill down our list of candidate oGPCRs to 12 out of a possible 200 sequences. Of these 12 oGPCRs, 2 were subsequently deorphanized in the literature: the prolactin-releasing peptide receptor (D'Ursi *et al.*, 2002, Gu *et al.*, 2004) and the neuropeptide W receptor (Shimomura *et al.*, 2002). Both of these GPCRs have ties to energy metabolism, thereby further validating our approach. One oGPCR had considerable intellectual property issues associated with it and another oGPCR showed no evidence of having a mouse homologue. Therefore, we chose 8 oGPCR sequences to move forward in the program. Four of these were high-throughput screened (HTS) using a fluorescent imaging plate reader (FLIPR) assay for calcium flux readout using a promiscuous G-α16 protein background for facilitating the heterotrimeric G protein coupling (data not shown). Four

other oGPCRs were marked for HTS using cell-based, high-speed microscopy-based, green fluorescent protein (GFP) translocation readout. All 8 oGPCR sequences were cloned for the full-length human receptor and subsequently transferred to a pCDNA3.1 cloning vector (see later).

Transfluor

The strategy has been partially validated during the past years because several of the orphan receptors have subsequently been deorphanized by others and demonstrated to regulate feeding behavior, for example, prolactin-releasing peptide receptor and neuromedin receptor (Bouchard *et al.*, 2004; Hanada *et al.*, 2004). We are, however, hampered by the ability to quickly develop high-throughput assays for these orphan GPCRs that couple to an appropriate G protein pathway for a reliable readout of receptor activation, such as calcium flux or changes in cAMP levels. Therefore, we were very interested in the Transfluor technology that is based on the interaction of β-arrestin and an activated GPCR initiating the process of receptor internalization. Multiple HCS instrument platforms are available for cellular imaging that allow for the quantification of the redistribution of β-arrestin–GFP when complexed with the overexpressed GPCR detecting the GPCR signaling. The feasibility of this approach has been shown by Norak Biosciences for over 50 known GPCRs where they have obtained quantitative data for receptor agonists and antagonists that compare with existing technologies. Advantages of this technology include a single readout for all GPCRs; regardless of the type of endogenous G protein coupling, it can be used with all orphan-GPCRs. Without the need to collate out those targets that do not couple to Gq, there is a drastic decrease in assay development time for all GPCR assays, and it is applicable to GPCRs of interest to any therapeutic area without modification. This assay format is HTS compatible and can be used to generate IC_{50}, EC_{50}, and SAR determinations. In addition, there are no expensive reagents required beyond the initial investment in the imaging platform. Therefore the choice of Transfluor offered us this relatively fast HCS/HTS screen on our predefined selected oGPCRs and has positioned us to establish patent claims on these targets.

Technology

For primary screening and secondary follow-up experiments of the potential ''hit'' compounds an Opera confocal imaging plate platform is used. The platform comprises the confocal microscope, the scheduling software, Bernstein, and acquisition and processing customized software for
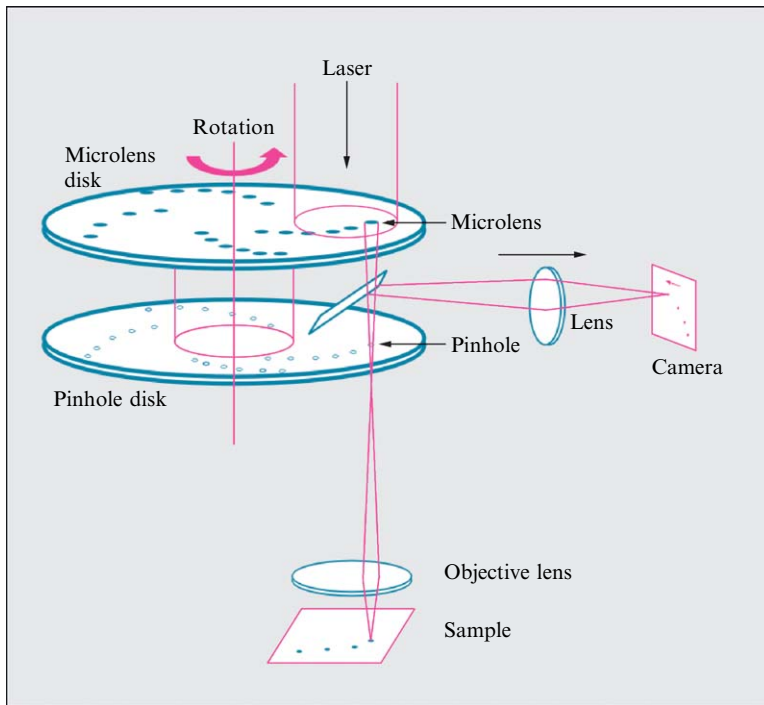
FIG. 1. Confocal imaging based on Nipkow disk technology. A collimated laser beam passes through a spinning disk containing microlenses. Each microlens focuses the incoming laser beam onto a corresponding pinhole, which is imaged into the sample. Fluorescence emitted by the sample in the focal plane is imaged back through the pinhole so that it can pass the pinhole disk and be detected with CCD cameras. Fluorescence from out-of-focus areas of the sample is not focused back through the pinhole but to a large extent is blocked by the pinhole disk. The disk contains approximately 20,000 pinholes with a diameter of approximately 50 μm.

data analysis. This instrument uses a Nipkow spinning disk to create confocal images composed from the approximately 1000 concurrent confocal beams (Fig. 1). The advantage of confocal imaging is that signals from areas of the image sample that are not in the focal plane are strongly attenuated, thus resulting in a very high signal-to-background ratio in the images. Of the four available light sources, 488- and 635-nm confocal excitation lines are used in this HCS assay in order to excite GFP from the β-arrestin–GFP fusion protein, GFP from Renilla mulleri (Prolume, Pinetop, AZ) DRAQ5 (the nuclear dye used to define both nuclear area and cytoplasmic area), respectively. Emission is detected simultaneously using two cameras in spectral bands at 535 and 690 nm. Although enabled, the third camera is

not activated in these experiments. For the primary screening experiments, an air 20×/0.4 NA lens is used. However, for the follow-up of potential agonists accomplished by a 10-point dose–response experiment, a 20×/0.7 NA water immersion lens is implemented. To accomplish this, the immersion water supply is fully automated within the Opera platform, enabling automatic measurement of multiple plates without user intervention. The water immersion lens typically provides higher numeric aperture as compared to the air lenses, which results in higher image quality that is independent of the position of the focal plane within the sample.

To accommodate the high-throughput plate handling and reading requirements of the HCS/HTS assay, greater than 2000 384-well formatted microtiter plates, the Opera is interfaced with a Twister II plate handler (Caliper Life Sciences, Hopkinton, MA) and an external bar code reader (Fig. 2). These are all controlled with Bernstein scheduling software, a complete, dynamic scheduling software allowing the automated robotic plate handling and image acquisition of up to 80 384-well microtiter plates in a single run. This software proves versatile as it also handles the adding and removing of plates during an active processing run. Due to the fact that the time to acquire and perform the on-line image analysis takes 11 to 12 min per 384-well plate, we vary the number of plates per run. This allows us to accommodate (1) the transfer of data and images to the Roche data servers, (2) assay curator time to reevaluate the previous days run, and (3) reimage any out-of-focus wells or plate failures.

FIG. 2. Opera with Twister II plate robot and plate stacks on a Scala. An interface station with keyboard switch for the evaluation network is visible on top of the instrument.

The recorded images are analyzed and stored by a system of three parallel computers, using Acapella data analysis software. Numerical results from the analyzed images are stored in a hierarchical directory structure in XML files. For XML files, many tools for processing the information may be used, thus enabling a simple import of the results into different data processing, storage, and mining systems. For both the primary screening campaign and the follow-up dose–response experiments we use an XML transformation to convert the XML files into comma separated variable text files (csv) to import into ID Business Solutions, Inc. ActivityBase software (IDBS, Bridgewater, NJ). Furthermore, we use APlus screening data analysis software for follow-up dose–response assays that combine numerical screening data with compound information provided by supplementary data bases or imported via compound XML files. The analysis performed by APlus generates EC_{50} curves, percentage maximal responding efficacies, and a scoring based on the dose–response curve and positive/negative controls. We use this analysis to interpret what the specificity of compound agonist action is on one particular orphan GPCR versus others.

Image Analysis Algorithm

As with common HCS assays the image analysis is based on object detection and quantification. The Acapella application in this Transfluor assay is essentially an application quantifying receptor internalization. Representative images are shown defining distribution of the β-arrestin–GFP Prolume-labeled molecules and images depicting the nuclear stain, DRAQ5 (Fig. 3).
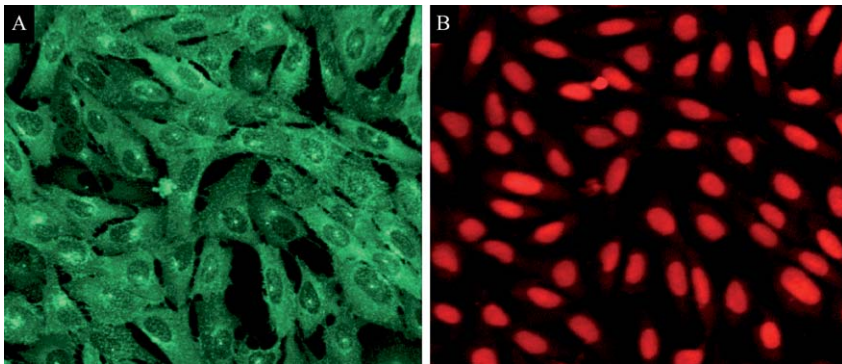

FIG. 3. Typical images showing GFP (A) and DRAQ5(B) fluorescence.

In the first steps of image analysis, nuclei and cell bodies are detected. For nuclei detection, red nuclear fluorescence is used (Fig. 4A). For cell body detection, GFP fluorescence could be used. However, in vesicle forming versions of the Transfluor assay we observed that virtually all of the GFP molecules were recruited to the vesicles so that no signal for cell detection was left. The DRAQ5 staining, however, was not an exclusive stain for DNA but showed some spurious staining in the cytoplasm. As a result, we could define both the nucleus and the cell body based on red nuclear fluorescence (Fig. 4B). The area that was analyzed further was defined using the identified cells. Cells detected using the DRAQ5 signal appeared smaller than cells visible in the GFP channel. Therefore, the cell shapes were dilated (Fig. 4C). The fluorescence distribution in the area covered by the cells was then analyzed in terms of object-oriented image analysis. In Acapella, an ''object'' is a collection of shapes or geometrical attributes combined with a number of numerical attributes, for example, a cell can be represented as an object that contains two geometrical attributes describing where the nucleus and the cell body are and a set of
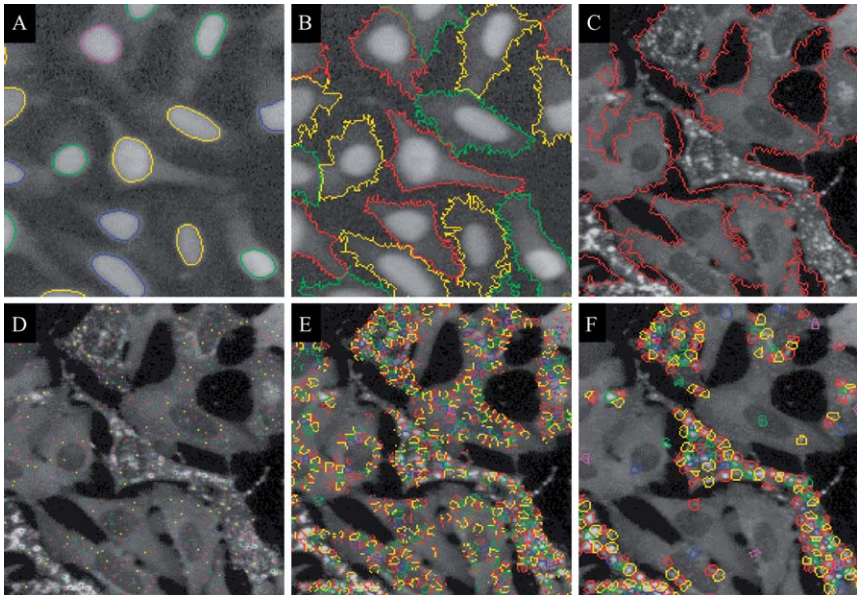

Fig. 4. Image analysis steps for the Transfluor assay. (A and B) Nucleus and whole cell body detection based on the DRAQ5 stain. (C) Area where GFP distribution is analyzed shown on the GFP image. (D and E) Initial set of local intensity maximums and corresponding reference areas for these maximums. (F) Reference regions of the selected maximums. Different colors help in identifying individual objects.

numbers such as area of the nucleus or intensity of GFP fluorescence in the cell body.

In analysis of the GFP fluorescence distribution, first the local intensity maximums were identified (Fig. 4D and E). This was the initial set of objects in the sequence of image analysis steps, and each maximum consisted of a single pixel in the image that was brighter than all surrounding pixels. These maximums were caused either by simple fluctuations in the fluorescence signal or by real aggregations of fluorescent molecules. The absolute intensity of the spots was not suitable for deciding if a spot was caused by a simple intensity fluctuation or not, as the average intensity of the cells containing the spots varied quite drastically. With a fixed intensity limit, some cells would have more maximums simply by the fact that more GFP-labeled molecules were expressed. Therefore, for each maximum, two other criteria were determined, enabling the selection of only those maximums, which were visually recognized as bright spots in the image without the artifacts created by different GFP expression levels in the different cells of the sample. With these two criteria, valid maximums were selected (Fig. 4F). These maximums were regarded as clathrin-coated pits or vesicles. No assumption had to be made in this selection process as to which type of fluorescence aggregation was observed.

Based on the selected maximums, the total number of qualified bright spots, the number of qualified bright spots normalized to the number of cells, or the total fluorescence contained in spots normalized to the total fluorescence in the cells was determined. All these parameters were used to determine the typical state of the whole cell population.

For the experiments described more outputs were added in order to get a better understanding of the processes in the assay. Adding such outputs required only a few lines of Acapella script code once the objects were identified. The first additional output employed quantifying nuclear area to add information about the compound effects on the cells. Nuclei of dead or apoptotic cells are typically smaller than nuclei of normal cells. Therefore the average size of nuclei was added to the list of outputs. Also the fluorescence signals in the nuclear channel are usually not influenced by the assay. Therefore the fluorescence intensity in this channel was used to determine if a compound was autofluorescent or not. Since two channels were being acquired, this translated into identifying whether compounds displayed fluorescence in the red or green channel comparing the same intensities over the nuclear area to the basal controls.

The second more assay-specific output added was the fraction of qualified maximums in the nuclear area. During assay development where a test library of compounds was screened it was observed that some compounds led to the formation of GFP aggregates in the nucleus (Fig. 5). Therefore
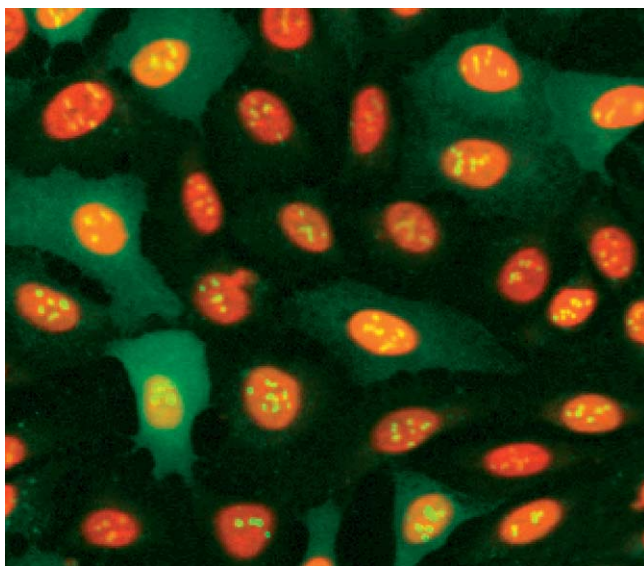
FIG. 5. Sample with GFP aggregates in the nucleus.

the fraction of the spots in the nucleus was added as an additional output. This output was subsequently used for flagging compounds that caused this kind of unusual GFP distribution. It was determined after subsequent experiments that this phenotype consequently led to cell death. These additional outputs demonstrated the virtues of high-content assays. While samples were analyzed with respect to the internalization of the receptor: β-arrestin complex quantifying the Transfluor assay, the simultaneous information on compound-specific properties, such as nuclear area, apoptosis, cytotoxicity, autofluorescence, and the induction of an unusual phenotype of GFP distribution, was obtained.

EC_{50} Fitting and Scoring

Numerical results of the dose–response confirmation screen were analyzed further using the APlus screening data analysis software. This is a client/server system for combining screening results with compound information from additional data sources. In the case of this experiment the compound information was provided using an XML file defining compound concentration per well. The location of the high and low control wells was derived from the plate layout. This information could also have been provided via the compound information path.

At first several screening metrics such as Z' or the signal-to-background ratio were determined (Zhang *et al.*, 1999). From the positive controls for each plate, G protein-related kinase (GRK) transfections, the full activity (100%) of the assay was determined and based on this, the activity of the individual wells was calculated. These activities were the basis for further analysis. Referencing to the controls on the same plate allowed for the compensation of potential drifts and so on.

In the next step for each compound the dose–response curve was fitted using various models. The best-fitting model was selected automatically by the software. Available models used were as follow.

- sigmoidal: classical dose response model
- linear: central part of the sigmoidal model
- exponentially converging/diverging: higher/lower end of the sigmoidal model
- constant positive/negative: constant high/low measurements
- unclassified: curves that cannot be described with any of the other models

For data generated, the sigmoidal, linear, and exponential models of EC_{50} values were generated.

In addition to the fitting model, the dynamic range was also reported. Dynamic range classes were none, medium, and full to extreme, where these classes describe how a dose–response curve compares to the positive and negative controls. As the potential compound's agonist response could be classified as full or partial, this means of categorizing the effects was practically constructive.

From all this information—the fitting model, the EC50 value, and the dynamic range—a score for each compound tested in a dose–response fashion was generated. These scores are in the range of 1 to 30 and describe the potential of a compound in a single value. The mathematical model for this scoring has been determined empirically before and was used successfully in many screening campaigns. Using this dynamic range measurement, compounds may be classified as to varying characteristic responses on the receptor internalization and β-arrestin–GFP redistribution process.

Assay Methods

Tissue Preparation

Three diet-induced obese (DIO) and three lean mice (C57Bl/6J) brains are used (Van Heek *et al.*, 1997). Using RNase-free conditions, the brain is removed rapidly, placed in a cryomold, covered with OCT compound

(Tissue Tek frozen tissue embedding medium Sakura Finetek, Torrance, CA), frozen in liquid nitrogen, and sectioned serially at 7 to 10 μm in a cryostat. Sections are mounted on plain (noncoated) clean microscope slides and placed on dry ice immediately. The sections are stored at –70°. The slides are fixed in 70% ethanol, rinsed in RNaase-free distilled water, stained in cresyl violet acetate (to identify the nuclei of interest), and dehydrated in graded ethanol (95% and absolute) with a final 5 min in xylene as the last step in the procedure. Slides are then placed in a desiccator for drying. Completely dried slides are used for laser capture microscopy.

Following the manufacturer's protocol, nuclei from the following hypothalamic regions are microdissected for further study: arcuate nucleus, dentate gyrus, dorsomedial hypothalamus, lateral hypothalamus, medial amygdala, median eminence, paraventricular nucleus, and ventromedial hypothalamus. These hypothalamic regions are identified as referenced in Franklin and Paxions (1997). After performing laser capture microdissection, the cap (with captured cells) is placed in a reagent tube containing 200 μl RNA denaturing buffer (GITC) and 1.6 μl β-mercaptoethanol. The tube is then inverted for 2 min to allow tissue to be digested off the cap. The reagent containing microdissected nuclei is then ready for RNA extraction, isolation, amplification, etc.

Laser Capture Microscopy

The PixCell II LCM system from Arcturus Engineering (Mountain View, CA) is used for laser capture microdissection. This system utilizes a low-power class IIIb invisible, infrared laser. Specially developed Cap-Sure Macro Caps (Arcturus Engineering, Mountain View, CA) coated with thermoplastic film are placed over the tissue section. The cells of interest are positioned in the center of the field. A focused laser beam is pulsed over the cells of interest and the transfer film fuses to the selected cells. When the cap is lifted off, the selected cells remain adherent to the cap/film surface. The film on the under side of the CapSure is placed in direct contact with the reagent buffer. The cellular contents detach from the film into the reagent buffer and are ready for further processing.

Single Target Expression Profiling: Real Time Quantitative PCR

The snap-frozen human tissues used in the STEP analysis are obtained from Cooperative Human Tissue Network (CHTN, Eastern Division, NCI, Philadelphia, PA). Tumor tissues have been trimmed previously by licensed pathologists to limit potential cross-contamination. Total RNA is extracted from snap-frozen human tissues using ''Ultraspecl RNA isolation kits'' (Biotecx Laboratories, Houston, TX, BL10100) and purified further

using RNeasy mini kits (Qiagen, Valencia, CA). Fifteen milligrams of total RNA is converted into double-stranded cDNA by reverse transcription (GIBCO BRL Life Technologies, Grand Island, NY) using the T7–T24 primer [5′-GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG (dT24)] and cleaned up by phenol/chloroform/isoamyl extraction using phase lock gel.

Master 384-well plates are generated containing 5 ng/ml double-stranded cDNA from 40 different tissue types (300 tissues total). Daughter plates are produced [final cDNA concentration: 40 pg/ml (200 pg/well)] via robotics. Duplex real-time PCR (target gene and GAPDH as reference gene) on 384-well optical plates is performed using TaqMan technology (Roche Molecular Systems, Inc., Alameda, CA) and analyzed on an ABI Prism PE7900 sequence detection system (Perkin-Elmer Applied Biosystems, Lincoln, CA), which uses the 5′ nuclease activity of *Taq* DNA polymerase to generate a real-time quantitative DNA analysis assay. PCR mix per well (25 ml) consists of commercially available, premixed GAPDH TaqMan primers/probe (PE), 900 nM each of 5′ and 3′ primers, and 200 nM TaqMan probe from each target gene, 200 pg cDNA and TaqMan universal PCR master mix (Perkin-Elmer Applied Biosystems). The following PCR conditions are used: 50° for 2 min and 95° for 10 min, followed by 40 cycles at 95° for 15 s and 62° for 1 min. The expression levels of target genes are normalized to reference gene levels and represented as relative expression (E) where DCt is the difference between reference and target gene cycles at which the amplification exceeds an arbitrary threshold. For reference gene values, GAPDH Ct values for each tissue are adjusted based on expression values of a panel of eight housekeeping genes in order to further improve normalization.

Development of Stable Cell Line

After the choice of the orphan candidate is made, oGPCR38 cDNA is cloned into the modified pCDNA3.1/Zeo(+) vector (Invitrogen Life Technologies, Carlsbad, CA). A PCR-generated IVS-IRES fragment using the pIRESpuro vector (Invitrogen Life Technologies, Carlsbad, CA) as a template is inserted into the *Xba*I-*Bst*107I sites. This vector is reused in a cassette-like manner to develop other oGPCR cell lines by insertion of the individual cDNAs. To construct the vector the following is performed.

1. Digest pCDNA3.1/Zeo(+) with *Xba*I and *Bst*1107 I. Isolate the fragment with the CMV promoter.
2. Using an IRES-containing vector (pIRESpuro), PCR amplify the IVS-IRES portion of this vector. Engineer an *Xba* I site at the 5′ end and incorporate about 20 5′ bases of the Zeo gene at the 3′ end.

3. Using the pCDNA3.1/Zeo vector as a template, PCS amplify the Zeocin-SV40 polyadenylation fragment incorporating about 20 3′ bases of IRES at the 5′ end and including a *Bst*1107I site at the 3′ end.

4. Set up overlapping PCR using amplicons generated from steps 2 and 3, digest with appropriate enzymes, and ligate with fragment from step 1.

5. The following restriction sites will no longer be unique when the construct is completed: *Hind*III, *Kpn*I, *Pst*I, *Apa*I, and 5′*Pme*I.

6. The remaining unique restriction sites in the MCS will be *Nhe*I, *Pme*I, *Afl*II, *Bam*HI, *Est*XI, *Eco*RI, *Eco*RV, *Bst*XI, *Not*I, *Xho*I, and *Xba*I.

7. The resulting vector is CMV–MCS–IVS–IRES–Zeo–SV40 PA.

The oGPCR DNA is inserted into the vector as a FLAG-oGPCR38 with a modified cytoplasmic tail region, referred to as ''E'' for enhanced activation. This modification of adding phosphorylation sites to the carboxy-terminal tail of the receptor has been shown to enhance the affinity of the receptor:β-arrestin interaction and has been described by Oakley *et al.* (2005a). The cells are maintained in DMEM 10% FCS (heat inactivated), 10 μg/ml gentamicin, 10 mM HEPES, 2 mM L-glutamine, 0.4 mg/ml Zeocin (to maintain selection of the oGPCR38), 0.4 mg/ml G418 (to maintain selection of the β-arrestin-GFP-Prolume tag), and penicillin/streptomycin. To identify the receptor expression during the cloning, propagation, and screening processes, cell aliquots are monitored for FACS analysis. To monitor receptor expression it is critical to remove the monolayers from the tissue culture flasks or the cell factories using EDTA. Although we use typical trypsinization procedures for propagating the cells, when cells are trypsinized 24 h prior to assay, FACS analysis results show an unexpectedly low receptor number due to the trypsin degrading the receptor at the surface of the cells. We monitor receptor expression on a twice-weekly basis using the Guava Technologies personal cell analysis instrument (Guava Technologies, Hayward, CA) employing an anti-FLAG M2 (mouse monoclonal) from Sigma and a secondary goat F(ab′)$_2$ anti-mouse IgG-PE from Guava Technologies. This Express assay uses proprietary Guava Express software to quantify receptor expression along with a viability assessment. During the process of producing clones of oGPCR38, a minimum of 10 independent cell clones are characterized prior to identifying the clone for the HCS/HTS assay. We chose a clonal line (#216) for three reasons. One was due to the fact that the FACS analysis confirmed it as a highly responding population of receptor expressing cells. Second, we also identified #216 as the most responsive vesicle/pit forming clone

responding to the GRK LITe assay (Oakley *et al.*, 2005b), achieving a S/B >10. Third, it also fulfilled the criteria of being a most stable clonal cell line during the course of assay development by reproducibly retaining receptor expression levels and maintaining the response to the GRK LITe assay over 22 passages. The propagated cells were maintained under the afore-mentioned standard culture conditions and used up to passage 21 for the screening assay and did not exceed this. We monitored receptor expression on a twice-weekly basis over the course of the screen, which confirmed that the receptor was maintained at >90% expression over the 20 days of screening runs.

HCS Assay Procedure

Materials

- Transfluor cell line(s)
- GRK LITe BacMam
- MEM media without phenol red
- HEPES
- HBSS
- Phosphate-buffered saline (PBS)
- Dimethy sulfoxide (DMSO)
- DRAQ5
- Formic acid

A generic description of the Transfluor assay methods used has been published by Hoffman (2005) among other literature and commercial methods (Ghosh *et al.*, 2005; Hudson *et al.*, 2002; Oakley *et al.*, 2002). The following methods for this screening campaign are similar with the following changes. We had been able to put into place a ligand-independent control using the GRK in a transiently expressed BacMam vector, referred to as the LITe assay (Ames *et al.*, 2004). This allowed us to monitor the 100% full activation of the assay with this transiently ex-pressed protein. Maximal transcription resulting in receptor activation and internalization occurred over the course of 16 to 20 h, allowing the β-arrestin:receptor internalization process to occur. The GRK virus is added 4 to 6 h after the double stable transfected U 2-OS oGPCR 38 cell line is plated onto 384-well microtiter plates. The multiplicity of infection (MOI) is optimized to 60, which translates into the addition of a 1- to 2-μl aliquot per control well depending on the viral stock potency. We chose to proceed with 16 wells of the positive GRK controls and 16 wells of the basal or untreated cells exposed to a maximum of 0.75% DMSO as the vehicle, as this is what the compounds were diluted in. Our choice of

compound concentration was based upon experience running other cell-based screening assays, as well as the running the previous Transfluor assays where the results of hits and leads were quite feasible to follow up.

Compound plates are prepared previously as single-use plates containing a 2-μl DMSO aliquot of a 1 mM stock compound solution. Plates are diluted with 50 μl of HBSS containing 20 mM HEPES on a screening day using a Multidrop microplate dispenser with Titan stackers (Titertek, Huntsville, AL) for dispensing. These 384-well plates are bar coded to map the individual compounds per plate in IDBS's ActivityBase software. Typically, compounds are prealiquoted in only 352 of the 384 wells, leaving columns 1 and 2 open for control solutions. In compound plates, columns 1 and 2 are filled with 50 μl HBSS containing 20 mM HEPES and 4% DMSO.

Cell plates are formatted such that the first two columns are reserved for the controls: column 1 are GRK–LITe-treated positive controls and column 2 are basal vehicle controls. The 384-well cell plates (BD Falcon 384 well black/clear plates, Tissue Culture treated, reference #350504) are plated with 4000 Transfluor cells per well 20 h prior to assay. For the 16 wells in column 1, the 1- to 2-μl aliquot of the GRK BacMam is added 4 h after cell plating. Procedurally, 80 cell plates and 80 compound plates are matched up on a given daily screening day for executing the campaign. Four consecutive rounds of 20 compound and 20 cell plates are prepared and processed consecutively. Twenty compound plates are diluted as described at the same time as conditioned growth medium from 20 cell plates is replaced manually with 25 μl of prewarmed 37° MEM (without phenol red) containing 20 mM HEPES using a Multidrop from Titertek (Huntsville, AL). This replacement step is required to be done in a relatively quick time frame to avoid drying of the cells in the cell plates. Hereafter 6 μl of the compounds and controls is aliquoted into the cell plates using a programmed Tomtec Quadra 384 workstation (TOMTEC, Hamden, CT) in which tips are washed with DMSO followed by three washes in water and sonication to clean the tips between plates. The compound-containing cell assay plates are then left for 1 h incubation at room temperature. These plates are then quenched using a second dedicated stacker Mulitdrop where a 25-μl aliquot of 4% formic acid prepared in HBSS containing 20 mM HEPES and 2 μM DRAQ5 stain is dispensed. Cell assay plates are then lidded and remain in a laminar flow hood overnight. The following day, the plates are then washed with three 50-μl aliquots on a Titertek MAP-C2 liquid-handling assay processor with PBS. The final processing step dispenses a 25-μl aliquot to each well. The now-completed plates are heat sealed with plastic black sealers, and the run is placed into the Twister II Plate Handler stacker.

Image and Data Acquisition

The acquisition of 80 plates takes 12 h after which the assay curator reviews the plate-by-plate analysis to ensure that there were no plate errors in reading either whole plates due to occasional nonstandard plate specifications or errors due to no data acquisition, which occasionally is due to nonfocusing issues. These identified plates are rerun using the same algorithm collecting another field of view, which in 90% of the situations proves successful. Nonsuccessful acquisitions are then fully repeated in subsequent assays. A daily examination of the preliminary findings is done to verify that the images resulted in the expected phenotype defining the hit compounds.

Image Analysis Algorithm

The image analysis is implemented as Acapella script code. The relevant part of the algorithm is shown here:

```
nuclei_detection_g(Image = reference)
if (nuclei.count != 0)
   Cytoplasm_detection_b(cytoplasm = reference)
   WholeObjectRegion()
   set(WholeCells = Objects)
   set(SearchMask = WholeCells.ObjectRegion)

   set(NumberOfCells = nuclei.count)
   CalcIntensity(Image = signal, Cells = nuclei)
   set(GreenSignal = cells.intensity.mean)
   CalcIntensity(Image = reference, Cells = nuclei)
   set(RedSignal = cells.intensity.mean)
   set(NucleusArea = nuclei.area.mean)
   if (SearchMask.area != 0)
      spot_detection_c(signal, "ObjectRegion," ShowOutputParameters = no)
      // determine the number of the nucleus at this location
      calcintensity(SpotCenters, nuclei.index, Objects = spots)
      objectfilter(SpotCenters_intensity != 0)
      if (spots.count != 0)
         set(NuclearCenterFraction = objects.count/spots.count)
      end()
   end()
end()

output(IntegratedSpotSignalPerCellularSignal_BackgroundSubtracted,
   "Counts from centers per total counts")
output(NumberOfCells,"Number of cells")
```

output(NucleusArea, ''Size of Nuclei'')
output(GreenSignal, ''Green Nuclear Fluorescence'')
output(RedSignal, ''Red Nuclear Fluorescence'')
output(NumberOfSpots, ''Number of translocation centers'')
output(NuclearCenterFraction, ''Fraction of translocation centers in the nucleus'')
output(SpotsPerObject, ''Centers per cell'')
output(SpotsPerArea, ''Centers per area'')

First nuclei are detected using a single cell analysis procedure call. If nuclei are present, cells are detected using the reference signal channel (DRAQ5) and the complete cell is defined as mask where spots are searched. Some initial results are determined, which are based on the detected nuclei: the number of cells, intensity of the green and red signal in the nuclear area, and the average size of the nuclei.

If the search mask is not empty, spots are searched and then those spots that are in the nuclei area are selected and their fraction of all the spots is calculated. Finally, all results are output to the calling application. The main result is *counts from centers per total counts*. The result parameters *number of cells, size of nuclei, green nuclear fluorescence, red nuclear fluorescence,* and *fraction of translocation centers in the nucleus* are used to flag compounds and as additional information sources.

As per the previous section, the comma-separated variable files are processed by Abase software where the identification of potential compound hits is tabulated. Once identified, experiments are prepared to test the potential hit compounds under 10-point dose–response curves in orphan GPCR38, 7, and 105. For these experiments, compounds are prepared from 5 mM DMSO stock solutions and diluted in HBSS containing 20 mM HEPES at final concentrations in the assay of 30 μM to 1 nM in duplicate. In these cases, the compounds are evaluated on the oGPCR38 as well as two additional oGPCRs propagated and prepared in the same manner. This allowed us to evaluate the specificity of the potential agonist compounds in the same Transfluor system. Thereafter, other assays used to characterize GPCRs such as cAMP, calcium flux as measured by luminescence, FLIPR, and other HCS internalization assays were employed.

Screen Metrics and Results

Example of Experimental Findings

Using the aforementioned experimental paradigm, we were able to achieve a plate throughput of 80–100 clear-bottom 384-well plates per screening day. This allowed us to finish the entire primary screen of ~750,000

compounds within 5 weeks. The primary hit rate was 0.27%, that is, 2073 qualified hits based on *counts from centers per total counts* (see earlier discussion). We were confident that the BacMam GRK positive control was not overestimating the degree of vesiculation that would be seen in the Transfluor cell assay in the presence of active small molecule compounds. On a relative scale accounting for the translocated/internalized β-arrestin–GFP oGPCR complexes, the $\sim$748,000 inactive compounds were clustered around 0.005 units, the BacMam GRK positive control displayed an average of 0.100 activation units, and active compounds ranged between 0.025 and 0.250 activation units. Therefore, in retrospect, the constitutively active GRK positive control was not overestimating the amount of receptor translocation in the presence of active compounds. Typically, at this point in a high-content, high-throughput cell-based screen, we conduct confirmation assays of the primary screening hits by reordering and rerunning each compound twice on three different cell plates. Compounds that maintain activity in two out of three retests are subjected to 10-point concentration response curves and selectivity assays using the oGPCR of interest (oGPCR38) and two other stably transfected oGPCR U-2 OS Transfluor cell lines (oGPCRs 7 and 105). The signal-to-background (S/B) ratio for oGPCR38 was 35, and the S/B ratio for oGPCRs 7 and 105 was 8.4 and 4.5, respectively. The S/B ratio for the latter two receptor screens was lower due to a level of residual basal activation in the resting state. The Z prime values for all three assays, the primary screen and the two counter screens, was >0.6. After identification of validated selective hit compounds, other mechanistic GPCR signal transduction assays, such as calcium flux and cAMP determination assays, were employed to investigate the specific heterotrimeric G protein-coupled pathway that was activated by the agonist compounds.

Preliminary findings on the validated hit compounds were verified by manual curation of the phenotypes displayed in the confocal images. Careful examination revealed hit compounds with the expected phenotype of β-arrestin–GFP:receptor complexes robustly distributed among a mixed population of clathrin-coated pits and internalized vesicles. However, a number of other expected and unexpected phenotypes were also identified.

i. Compounds that caused an activation event but exhibited a preference for either clathrin-coated pit morphology or internalized vesicle morphology via an unknown compound-specific mechanism.

ii. Compounds that elicited the expected pit or vesicle activation phenotype but also produced evidence of cytotoxicity, via either plasma membrane ruffling or cell shrinkage of overt cell loss.

iii. Compounds that caused a nuclear condensation event, as registered in the DRAQ5 (Biostatus Limited, Leicestershire, UK) nuclear stain channel for the object characteristic for *size of nuclei.*

iv. Compounds that produced *translocation centers* for the GFP β-arrestin that reside within the nuclear mask area, an unexpected event given the fact that a nuclear localization sequence was not factored into the experimental design.

v. Compounds that precipitated out of solution and the resulting particulate, aggregate, or crystalline-like structures were visualized

vi. Compounds flagged for autofluorescence in the red channel. These compounds, if genuinely active, may still be potential drug development candidates; however, a different mode of detection in the secondary assay may be appropriate.

vii. Compounds flagged for autofluorescence in the green channel. There was a great variety among the exact cellular localization of these compounds, ranging from compounds that were distributed broadly within the cytoplasm, to nuclear localization, to defined compartment/organellar localization (presumably due to selective cellular uptake mechanisms).

viii. Compounds that caused spurious phenotypes, such as the complete translocation of the β-arrestin–GFP to the actin fibrils within the cell and a single characteristic foci in the nucleus. These compounds may eventually become tools to help elucidate mechanism of β-arrestin involvement in the cytoskeleton or in cell cycle regulation.

Of the validated active compounds, we found incidences where a compound was monoselective for only the oGPCR of interest. In other cases, a compound was active on two out of three oGPCRs tested. There were also compounds that activated all three of the Transfluor oGPCR cell lines. There were even cases of oGPCR-specific cytotoxicity events, which may have been due to a receptor-based mechanism or simply differences in the U- 2 OS cell background due to stable clonal selection. Most of the active validated compounds displayed EC_{50} values in the low micromolar range; however, a number of compounds were potent in the high nanomolar range. The exact nature of the mechanism of each compound-induced GFP β-arrestin activation event remains to be discovered. Notably, there are literature reports of molecules that act through nonheterotrimeric G protein-coupled pathways, possibly through extracellular signal-regulated kinase activation or via arrestin-induced clustering of intracellular signaling molecules (Ren *et al*., 2005).

Concluding Remarks

We have described protocols for conducting high-throughput, fluorescent image-based confocal cell-based screens for the identification of surrogate small molecule chemistry as a starting point for identifying mechanistic

tools or for initiating medicinal drug discovery chemistry efforts toward a subset oGCPRs related to obesity and energy metabolism in humans. These selected oGPCRs were identified using a systems-based approach that combined the power of *in silico* bioinformatics, molecular biological techniques, genomics, stable green fluorescent protein expression, laser capture micro-dissection, quantitative expression profiling, high-speed automation-assisted laser confocal microscopy, and state-of-the-art live cell-based assay development. Such experimentation would not have been possible without the convergence of recent significant advancements within each of these fields of study. The use of GFP translocation events as a bona fide high-throughput readout opens the possibility of using similar translocation or activation events for addressing a wide variety of target classes in drug development and academic settings alike. In our hands, the Transfluor assay has been particularly effective in assisting in the elaboration of compounds to address oGPCRs and their signaling mechanisms, thus presenting a new avenue for receptor deorphanization strategies.

References

Ames, R., Fornwald, J., Nuthulaganti, P., Trill, J., Foley, J., Buckley, P., Kost, T., Wu, Z., and Romanos, M. (2004). BacMam recombinant baculoviruses in G protein-coupled receptor drug discovery. *Recept. Channels* **10,** 99–107.

Bouchard, L., Drapeau, V., Provencher, V., Lemieux, S., Chagnon, Y., Rice, T., Rao, D. C., Vohl, M., Tremblay, A., Bouchard, C., and Perusse, L. (2004). Neuromedin β: A strong candidate gene linking eating behaviors and susceptibility to obesity. *Am. J. Clin. Nutr.* **80,** 1478–1486.

D'Ursi, A., Albrizio, S., Di Fenza, A., Crescenzi, O., Carotenuto, A., Picone, D., Novellino, E., and Rovero, P. (2002). Structural studies on Hgr3 orphan receptor ligand prolactin-releasing peptide. *J. Med. Chem.* **45,** 5483–5491.

Franklin, K. B. J., and Paxions, G. (1997). ''The Mouse Brain in Stereotaxic Coordinates.'' Academic Press, San Diego, CA.

Ghosh, R. N., DeBiasio, R., Hudson, C. C., Ramer, E. R., Cowan, C. L., and Oakley, R. H. (2005). Quantitative cell-based high-content screening for vasopressin receptor agonists using Transfluor technology. *J. Biomol. Screen* **10,** 476–484.

Gong and Sjogren, Personal communication.

Gu, W., Geddes, B. J., Zhang, C., Foley, K. P., and Stricker-Krongrad, A. (2004). The prolactin-releasing peptide receptor (GPR10) regulates body weight homeostasis in mice. *J. Mol. Neurosci.* **22,** 93–103.

Hanada, R., Teranishi, H., Pearson, J. T., Kurokawa, M., Hosoda, H., Fukushima, N., Fukue, Y., Serino, R., Fujihara, H., Ueta, Y., Ikawa, M., Okabe, M., Murakami, N., Shirai, M., Yoshimatsu, H., Kangawa, K., and Kojima, M. (2004). Neuromedin U has a novel anorexigenic effect independent of the leptin signaling pathway. *Nature Med.* **10,** 1067–1073.

Hoffman, A. F. (2005). The preparation of cells for high-content screening. *In* ''Handbook of Assay Development in Drug Discovery'' (A. F. Minor, ed.), pp. 227–242. Dekker, New York.

Hudson, C. C., Oakley, R. H., Cruickshank, R. D., Rhem, S. M., and Loomis, C. R. (2002). Automation and validation of the Transfluor technology: A universal screening assay for G protein-coupled receptors. *Proc. SPIE Int. Soc. Optic. Eng.* **4626,** 548–555.

Kieffer, T. J., and Habener, J. F. (1999). The glucagon-like peptides. *Endocr. Rev.* **20,** 876–913.

Kinnecom, K., and Pachter, J. S. (2005). Selective capture of endothelial and perivascular cells from brain microvessels using laser capture microdissection. *Brain Res. Protocols* **16,** 1–9.

Ludwig, D. S., Tritos, N. A., Mastaitis, J. W., Kulkarni, R., Kokkotou, E., Elmquist, J., Lowell, B., Flier, J. S., and Maratos-Flier, E. (2001). Melanin-concentrating hormone overexpression in transgenic mice leads to obesity and insulin resistance. *J. Clin. Invest.* **107,** 379–386.

Oakley, R. H., Barak, L. S., and Caron, M. G. (2005a). Real-time imaging of GPCR-mediated arrestin translocation as a strategy to evaluate receptor-protein interactions. *In* "G Protein Receptor-Protein Interactions" (M. G. George and S. R. Dowd, eds.). Wiley, New York.

Oakley, R. H., Cowan, C. L., Hudson, C. C., and Loomis, C. R. (2005b). Transfluor provides a universal cell-based assay for screening G protein-coupled receptors. *In* "Handbook of Assay Development in Drug Discovery" (C. R. Minor, ed.), pp. 435–457. Dekker, New York.

Oakley, R. H., Hudson, C. C., Cruickshank, R. D., Meyers, D. M., Payne, R., Jr., Rhem, S. M., and Loomis, C. R. (2002). The cellular distribution of fluorescently labeled arrestins provides a robust, sensitive, and universal assay for screening G protein-coupled receptors. *Assay Drug Dev. Technol.* **1,** 21–30.

Ren, X., Reiter, E., Ahn, S., Kim, J., Chen, W., and Lefkowitz, R. J. (2005). Different G protein-coupled receptor kinases govern G protein and β-arrestin-mediated signaling of V2 vasopressin receptor. *Proc. Natl. Acad. Sci. USA* **102**(5), 1448–1453.

Reidhaar-Olson, J. F., Braxenthaler, M., and Hammer, J. (2002). Process biology: Integrated genomics and bioinformatics tools for improved target assessment. *Targets* **1,** 189–195.

Reidhaar-Olson, J. F., Ohkawa, H., Babiss, L. E., and Hammer, J. (2003). Process biology: Managing information flow for improved decision making in preclinical R&D. *Preclinica* **1,** 161–169.

Shimada, M., Tritos, N. A., Lowell, B. B., Flier, J. S., and Maratos-Flier, E. (1998). Mice lacking melanin-concentrating hormone are hypophagic and lean. *Nature* **396,** 670–674.

Shimomura, Y., Harada, M., Goto, M., Sugo, T., Matsumoto, Y., Abe, M., Watanabe, T., Asami, T., Kitada, C., Mori, M., Onda, H., and Fujion, M. (2002). Identification of neuropeptide W as the endogenous ligand for orphan G-protein-coupled receptors GPR7 and GPR8. *J. Biol. Chem.* **277,** 35826–35832.

Van Heek, M., Compton, D. S., France, C. F., Tedesco, R. P., Fawzi, A. B., Graziano, M. P., Sybertz, E. J., Strader, C. D., and Davis, H. R., Jr. (1997). Diet-induced obese mice develop peripheral, but not central, resistance to leptin. *J. Clin. Invest.* **99,** 385–390.

Zhang, J., Chung, T. D. Y., and Oldenburg, K. R. (1999). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4,** 67–73.