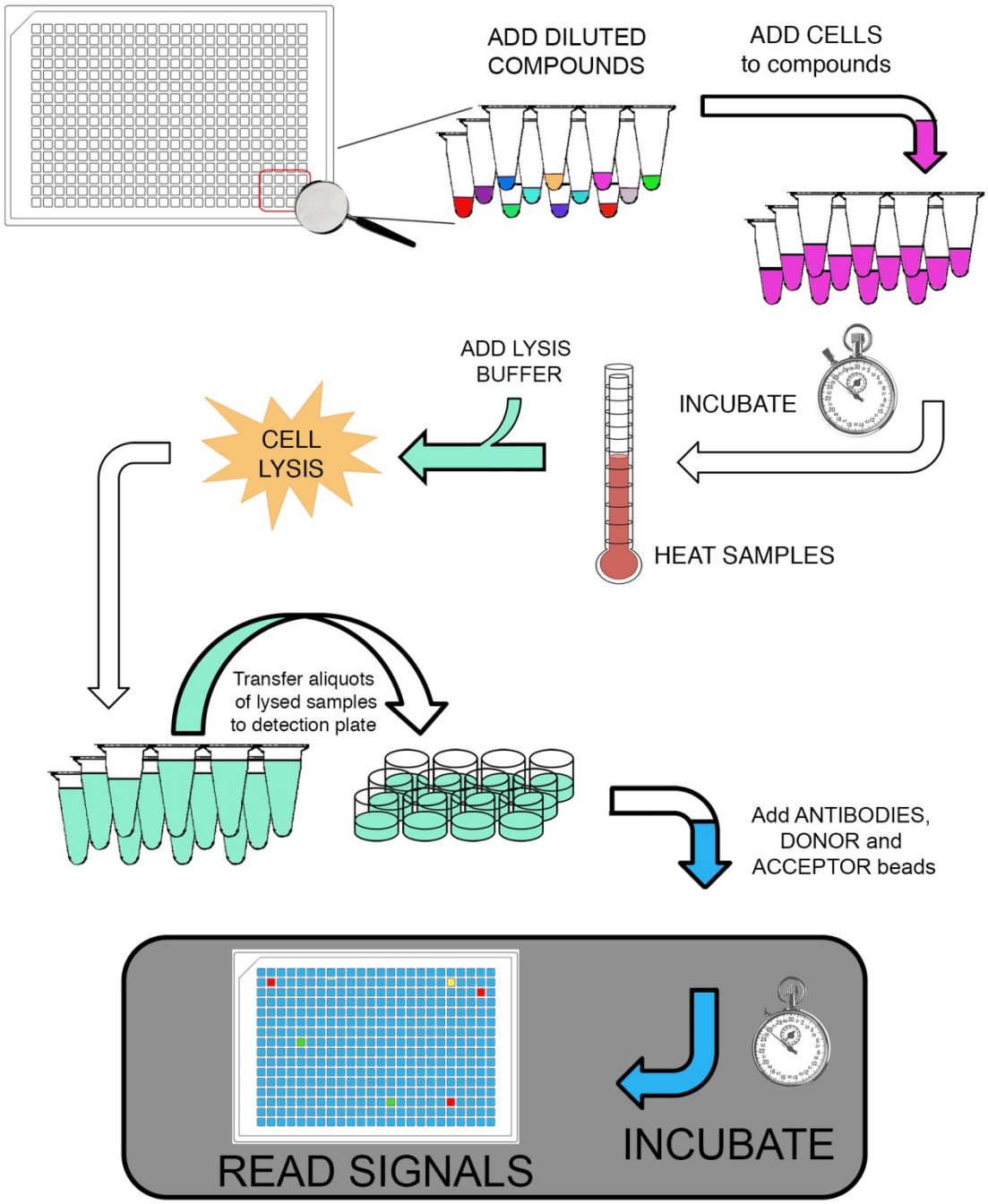


Figure 13: Schematic outline of the assay procedure and screening logistics as applied in the screening campaign on human TS using CETSA. Compounds and controls were predispensed to 384-well plates (Echo 550) and subsequently diluted with cell medium (Multidrop™ Combi). Next, the diluted compounds were transferred to 384-well PCR plates (Bravo liquid handling instrument). Cells were then added using a multipipette followed by preincubation for 2h in a cell incubator at 37°C. After the incubation the plates were transiently heated to 50°C for three minutes followed by a controlled cooling to 20°C (Roche LightCycler®480). The plates were then centrifuged briefly, lysed by addition of lysis buffer (Flexdrop IV). To ensure sufficient lysis the samples were thoroughly mixed using a 384-well tip head (Bravo). A small aliqout of the lysates were then transferred to detection plates followed by the addition of a solution of the detection reagents (Multidrop™ Combi) under subdued light. The plates were incubated overnight at room temperature before detection (Envision plate reader). Reprinted by permission from Macmillan Publishers Ltd: Nat. Commun. *7* (2016), copyright (2016) http://www.nature.com/ncomms/2016/160324/ncomms11040/abs/ncomms11040.html

Halozyme EX2083 (part 4 of 6)
Merck v. Halozyme
PGR2025-00039

Materials and Reagents

The following materials and reagents were used in the development and optimization of a microplate based live cell CETSA assay for TS using AlphaScreen® as the detection format. The assay was developed and validated as outlined above and successfully applied in a screening campaign for known and novel inhibitors (8).

- Human myelogenous leukemia cell line K562 (ATCC No. CCL-243)
- Echo 550 liquid handler (Labcyte)
- Bravo liquid handling platform (Agilent)
- PCR machine (ProFlex™, Applied Biosystems)
- Real-time PCR machine (LightCycler®480 system, Roche)
- EnVision microplate reader with AlphaScreen modality (PerkinElmer)
- Multidrop™ Combi reagent dispenser (Thermo Scientific)
- PlateLoc thermal microplate sealer (Agilent)
- Labcyte 384 LDV plates (LP-0200)
- 384-well polypropylene plates (784201, Greiner)
- 384-well hardshell PCR plate (HSR480, BIORAD)
- Peelable Aluminium seal (24210-001 Agilent)
- Breathable plate seal (3345, Corning)
- AlphaScreen® SureFire® Lysis Buffer (TGRLB100ML, PerkinElmer)
- AlphaLISA buffer (AL000F, PerkinElmer)
- ProxiPlate (#6008280, PerkinElmer)
- AlphaScreen® anti-mouse donor beads (#AS104D, PerkinElmer)
- AlphaScreen® anti-rabbit acceptor beads (#AL104C, PerkinElmer)
- Mouse monoclonal anti-TS IgG sc-376161 (Santa Cruz)
- Rabbit polyclonal anti-TS IgG (15047-1-AP, Proteintech)

Reduction of Assay Variability and Troubleshooting

When transferring an assay into a high-throughput format for screening purposes it is of great importance to reduce assay variability in order to obtain reliable statistics and accurate results. The screen formatting procedure usually also involves an attempt to reduce the number of assay steps to simplify the logistics, which generally results in reduced well-to-well variability and consequently improved assay quality. This section serves to outline some of the most critical aspects in the experimental procedure as we understand them today. Examples of unexpected technical challenges that have been encountered and solutions thereof are also discussed. A brief schematic overview of the different steps, including a practical guidance on how to reduce variability, is available in Figure 14. It is important to remember that the technology is young and is expected to develop and mature with each investigated model system and target protein.

Compound Treatment:

In order to obtain accurate results it is crucial that the cells in each well of the PCR plate are exposed to the intended compound concentration. A somewhat unexpected challenge when moving the assay to a plate-based format was a random lack of contact between compounds predispensed by means of acoustic nanodispensing and cells added with a bulk dispenser such as a MultiDrop™. This problem was caused by the formation of air bubbles trapped in the bottom of the wells in the PCR plate preventing contact between compound and cells. If instead larger volumes (µl) of diluted compounds are transferred to the PCR plates prior to addition of cells, this issue can easily be circumvented. Alternatively the user can consider the inclusion of a brief centrifugation step.

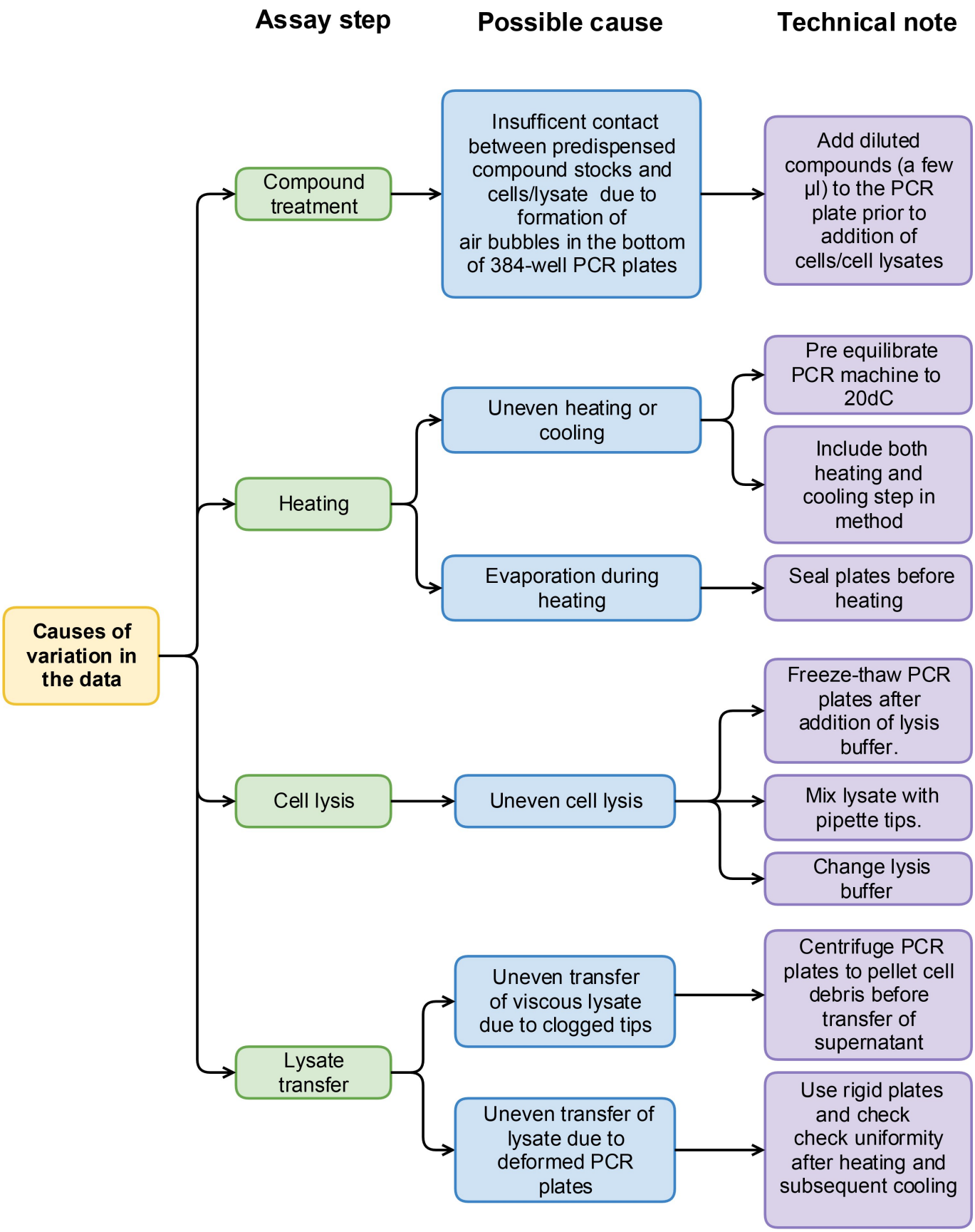


Figure 14. Outline of some critical aspects that influence assay variability in the plate-based CETSA procedure. Figure created using Gliffy software.

Transient Heating, Cooling and Lysis:

For this part of the procedure it is of great importance to identify a PCR plate that will not be distorted after the transient heating and the subsequent cooling to room temperature. This is essential to obtain an even transfer of the heated and lysed samples to the detection plate. In cases where no transfer step is necessary, *i.e.* the detection is done directly in the PCR plate, it will be important for the safe stacking and detection of the plates in a microplate reader.

One challenge in the microplate-formatting of the CETSA assay has been to obtain consistent and reliable cell lysis with little variability between wells. An uneven cell lysis will greatly impact assay quality and it is recommended to keep this step under careful control. The use of lysis buffer with thorough sample mixing will greatly improve the throughput compared to the repeated freeze/thawing procedure applied in the original protocol (6). A possible concern using lysis buffers is that the presence of detergents may resolubilize irreversible aggregates of denatured proteins, and thereby negatively affect the quantification of remaining soluble protein after the heating step. In the assay development process it is therefore advisable to examine this carefully for each individual case. This can for example be done by comparing the results from T_{agg} and $ITDRF_{CETSA}$ experiments with those obtained using the original WB protocol (6). Worth noting is that a recent study reports no signs of resolubilization for a broad range of proteins upon treatment of the irreversible aggregates with a variety of different detergents, with the exception being high concentrations of SDS that thus should be avoided (20). If insufficient lysis is observed with the lysis buffers an inclusion of a freezing and subsequent thawing step may improve the results.

Separation of Soluble and Denatured Proteins:

Separation of the irreversible aggregates of unfolded proteins from the soluble pool of proteins by means of centrifugation and the subsequent transfer of the supernatant (or a filtration step) is not an ideal setup for a microplate-based screening assay. Such steps will have a negative impact on both throughput and assay variability and should if possible be avoided. However, if the aggregated proteins are resolubilized, or if the antibodies cannot discriminate between native and denatured protein, a separation step is necessary.

Detection

One of the biggest challenges in the homogeneous assay setup with immunoassay detection is the observation that ligand binding can cause significant quenching of the signal due to loss of antibody recognition (7,8). This is not related to the CETSA approach per se but is instead associated with the detection step, as it was also found to be the case in the absence of heating. Since this has been observed for all proteins tested so far, the identification of a suitable antibody pair requires careful testing for this effect or screening for conditions to overcome this phenomenon (7,8,23). It is therefore recommended to include known binders of the target protein throughout the assay development process in order to identify and avoid such problems. Cellular systems that are not dependent on recognizing epitopes of native proteins are less likely to be affected in this way, provided the signaling tags are situated in such a way that their accessibility is not affected by ligand binding.

Tentative Applications

The original Science publication from 2013 now has over 100 citations in the scientific literature (as of June 1, 2016). The majority of these are based on studies of target engagement for a few molecules using the WB-based approach, while there are relatively few examples of more systematic studies for libraries or series of compounds in microplate format. Besides its potential utility for looking at patient responses or even for patient stratification we foresee its application in at least three stages in the preclinical drug discovery process.

First, given the abundance of available methodologies for target-based high-throughput screening, it is perhaps most likely that CETSA will be applied as a generic tool for prioritization and cleaning of hit lists in primary cells of relevance for the intended clinical indication. T_m shift assays are generally accepted as robust methods for the purpose of removing misbehaving compounds as they positively select for stabilizing compounds. Furthermore, CETSA in live cells allows for an early assessment of other aspects of the identified hits such as serum protein binding and cell permeability, while still being focused on the intended target protein rather than a proximal biomarker or a phenotypic response that could potentially be influenced by nearby pan-targets or completely different mechanisms.

Additionally, in the case of unknown functions for e.g. an orphan target protein or difficulties in expressing and purifying a particular target protein in a relevant form to support screening, CETSA could provide a viable means for identifying chemical starting points in a screening campaign. Most importantly though CETSA offers, similar to other target engagement assays that can be applied in live cells, a possibility to examine the structure-activity relationships for target engagement and thus allows for comparison with those relationships observed in various functional and phenotypic readouts. Such studies offer a great complement to other target validation tools and should be considered as a prerequisite for the validation of chemical probes.

References

1. Bunnage ME, Chekler ELP, Jones LH. Target validation using chemical probes. Nat Chem Biol. 2013;9(4):195–9. PubMed PMID: 23508172.
2. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov. 2014;13(6):419–31. PubMed PMID: 24833294.
3. Robers MB, Dart ML, Woodroofe CC, Zimprich CA, Kirkland TA, Machleidt T, et al. Target engagement and drug residence time can be observed in living cells with BRET. Nat Commun. 2015;6 10.1038/ncomms10091. DOI. PubMed PMID: 26631872.
4. Schulze J, Moosmayer D, Weiske J, Fernandez-Montalvan A, Herbst C, Jung M, et al. Cell-based protein stabilization assays for the detection of interactions between small-molecule inhibitors and BRD4. J Biomol Screen. 2015;20(2):180–9. PubMed PMID: 25266565.
5. Kaniskan HU, Szewczyk MM, Yu Z, Eram MS, Yang X, Schmidt K, et al. A potent, selective and cell-active allosteric inhibitor of protein arginine methyltransferase 3 (PRMT3). Angew Chemie - Int Ed. 2015;54(17):5166–70. PubMed PMID: 25728001.
6. Martinez Molina D, Jafari R, Ignatushchenko M, Seki T, Larsson EA, Dan C, et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. Science. 2013;341(6141):84–7. PubMed PMID: 23828940.
7. Jafari R, Almqvist H, Axelsson H, Ignatushchenko M, Lundbäck T, Nordlund P, et al. The cellular thermal shift assay for evaluating drug target interactions in cells. Nat Protoc. 2014;9(9):2100–22. PubMed PMID: 25101824.
8. Almqvist H, Axelsson H, Jafari R, Dan C, Mateus A, Haraldsson M, et al. CETSA screening identifies known and novel thymidylate synthase inhibitors and slow intracellular activation of 5-fluorouracil. Nat Commun. 2016;7 10.1038/ncomms11040. DOI. PubMed PMID: 27010513.
9. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen. 2001;6(6):429–40. PubMed PMID: 11788061.
10. Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal Biochem. 2004;332(1):153–9. PubMed PMID: 15301960.

11. Senisterra GA, Markin E, Yamazaki K, Hui R, Vedadi M, Awrey DE. Screening for ligands using a generic and high-throughput light-scattering-based assay. J Biomol Screen. 2006;11(8):940–8. PubMed PMID: 17092916.
12. Epps DE, Sarver RW, Rogers JM, Herberg JT, Tomich PK. The ligand affinity of proteins measured by isothermal denaturation kinetics. Anal Biochem. 2001;292(1):40–50. PubMed PMID: 11319816.
13. Sarver RW, Rogers JM, Epps DE. Determination of ligand-MurB interactions by isothermal denaturation: application as a secondary assay to complement high throughput screening. J Biomol Screen. 2002;7(1):21–8. PubMed PMID: 11897052.
14. Senisterra GA, Soo Hong B, Park H-W, Vedadi M. Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. J Biomol Screen. 2008;13(5):337–42. PubMed PMID: 18448703.
15. Molina DM, Nordlund P. The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target Engagement and Mechanistic Biomarker Studies. Annu Rev Pharmacol Toxicol. 2015;56(30):1–21. PubMed PMID: 26514198.
16. Sinclair I, Stearns R, Pringle S, Wingfield J, Datwani S, Hall E, et al. Novel Acoustic Loading of a Mass Spectrometer: Toward Next-Generation High-Throughput MS Screening. J Lab Autom. 2016;21(1):19–26. PubMed PMID: 26721821.
17. McNolty D, Bonnette L, Ho T, Qi H, Annan R, Holgate M. Revving up CETSA: Evaluating Drug Target Engagement in Living Cells via Cellular Thermal Shift Assay. 2016; Poster at the SLAS 5th Annual International Confer.
18. Menichelli E, Lun A, Ciceri P, Treiber D. HitHunter Pulse: A Novel Biochemical and Cellular Target Engagement Assay Platform for Drug Discovery. 2016; Poster at the SLAS 5th Annual International Confer.
19. Savitski MM, Reinhard FBM, Franken H, Werner T, Savitski MF, Eberhard D, et al. Tracking cancer drugs in living cells by thermal profiling of the proteome. Science. 2014;346(6205) 10.1126/science.1255784. DOI. PubMed PMID: 25278616.
20. Reinhard FBM, Eberhard D, Werner T, Franken H, Childs D, Doce C, et al. Thermal proteome profiling monitors ligand interactions with cellular membrane proteins. Nat Methods. 2015;12(12):1129–31. [Internet]. PubMed PMID: 26524241.
21. Franken H, Mathieson T, Childs D, Sweetman GMA, Werner T, Tögel I, et al. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. Nat Protoc. 2015;10(10):1567–93. doi: 10.1038/nprot.2015.101. Available from: http://dx.doi.org. PubMed PMID: 26379230.
22. Cimmperman P, Baranauskiene L, Jachimoviciūte S, Jachno J, Torresan J, Michailoviene V, et al. A quantitative model of thermal stabilization and destabilization of proteins by ligands. Biophys J. 2008;95(7):3222–31. PubMed PMID: 18599640.
23. Seashore Ludlow B, Axelsson H, Almqvist H, Jenmalm Jensen A, Lundbäck T. Quantitative Interpretation of Target Engagement and Specificity from Cellular Thermal Shift Assays; Manuscript in preparation.
24. Eglen RM, Reisine T, Roby P, Rouleau N, Illy C, Bossé R, et al. The use of AlphaScreen technology in HTS: current status. Curr Chem Genomics. 2008 Jan;1:2–10. PubMed PMID: 20161822.
25. Bembenek ME, Burkhardt A, Ma J, Li Z, Loke HK, Wu D, et al. Determination of complementary antibody pairs using protein A capture with the AlphaScreen assay format. Anal Biochem. 2011;408(2):321–7. PubMed PMID: 20868646.
26. Illy C, St.Pierre JA, Bouchard N, Wenham D. Development of High-Throughput Screening Assays for Kinase Drug Targets Using AlphaScreen^{TM} Technology. In: Minor LK, editor. Handbook of Assay Development in Drug Discovery. Boca Raton, FL, USA: CRC Press; 2006. p. 53–70.

Measurement of cAMP for $G_{\alpha s}$- and $G_{\alpha i}$ Protein-Coupled Receptors (GPCRs)

Tao Wang,[1] Zhuyin Li,[1] Mary Ellen Cvijic,[1] Litao Zhang,[1] and Chi Shing Sum[1]

Created: November 20, 2017.

Introduction

Cyclic adenosine monophosphate (cAMP) is an important intracellular second messenger in GPCR signal transduction. Agonist activation of GPCRs that couple to the $G_{\alpha s}$ protein leads to an increased production of intracellular cAMP levels, whereas activation of GPCRs that couple to the $G_{\alpha i}$ protein leads to reduced production of intracellular cAMP levels. Both of these intracellular cAMP changes are mediated through the modulation of adenylate cyclase activity. cAMP regulates the activity of cAMP-dependent protein kinase A (PKA), which plays an important role in a variety of downstream cellular processes.

A number of reagent kits are available on the market that can be used to measure intracellular cAMP levels. These include the HTRF cAMP kit from Cisbio, the LANCE cAMP kit from PerkinElmer, the HitHunter cAMP kit from DiscoverX and the cAMP Direct Immunoassay Kit from Abcam and BioVision. These assays are all based on the use of antibodies that specifically recognize both intracellular cAMP and an exogenous labeled cAMP conjugate that acts as a competitor, followed by detection of the labeled cAMP conjugate by a variety of detection technologies, including fluorescence resonance energy transfer (FRET) or enzymatic reactions. In addition, the antibody-independent GloSensor cAMP assay from Promega employs semi-split luciferase, which reassembles when bound to cAMP. The advantages and disadvantages of the various cAMP assay technologies are given in Table 1. This chapter focuses on the Cisbio HTRF cAMP kit to illustrate the methodology and considerations for development of assays that measure intracellular cAMP. The assay development principles discussed here can be easily applied to similar cAMP measurement kits or various other detection methodologies.

Table 1. Comparison of popular cAMP assay technologies

Technology	Vendor	Principle	Advantages	Disadvantages
HTRF	Cisbio	• A time-resolved assay technology based on fluorescence resonance energy transfer between a long-lifetime Europium cryptate fluorescent donor (Eu^{3+}) and a fluorescent acceptor (XL665 or d2). • Recognition of a fluorescent acceptor labeled cAMP by a fluorescent donor labeled anti-cAMP leads to energy transfer signal. • Competition from endogenous cAMP for the antibody results in reduced energy transfer signal. • Loss-of-signal measurement: the level of cellular cAMP is inversely related to the signal.	• Homogeneous assay • Ratiometric readout • Anti-cAMP antibodies available with different affinities toward cAMP for desired assay sensitivities • Does not require any recombinant protein expression for detection • Permits absolute quantification of cellular cAMP level by means of a standard curve	• Generally lower signal-to-background compared to other technologies

Author Affiliation: 1 Bristol-Myers Squibb Company, NJ.

Table 1. continued from previous page.

Technology	Vendor	Principle	Advantages	Disadvantages
LANCE TR-FRET	PerkinElmer	• A time-resolved assay technology based on fluorescent resonance energy transfer between a long-lifetime Europium chelate fluorescent donor (Eu^{3+}) and a fluorescent acceptor (U*Light*). • Recognition of a fluorescent donor labeled cAMP by a fluorescent acceptor labeled anti-cAMP leads to energy transfer signal. • Competition from endogenous cAMP for the antibody results in reduced energy transfer signal. • Loss-of-signal measurement: the level of cellular cAMP is inversely related to the signal.	• Homogeneous assay • Ratiometric readout • Does not require any recombinant protein expression for detection • Permits absolute quantification of cellular cAMP level by means of a standard curve	• Generally lower signal-to-background compared to other technologies
HitHunter	DiscoverX	• An enzyme fragment complementation (EFC) technology-based assay that uses two fragments of *E. coli* β-galactosidase (β-gal): a large protein fragment (enzyme acceptor, EA) and a small peptide fragment (enzyme donor, ED). Separately, these fragments are inactive, but in solution they rapidly recombine to form active β-gal enzyme, which can hydrolyze substrate to produce a luminescent signal. • cAMP from cell lysates and ED-labeled cAMP (ED-cAMP) compete for anti-cAMP antibody binding. Unbound ED-cAMP is free to complement EA to form active enzyme, which subsequently produces a luminescent signal. The amount of signal produced is proportional to the amount of cAMP in the cell lysate. • Gain-of-signal assay: increased cAMP level leads to increased signal.	• Homogeneous assay • High signal-to-background • Does not require any recombinant protein expression for detection • Permits absolute quantification of cellular cAMP levels by means of a standard curve	• Compounds that block the recombination or inhibit the activity of β-gal could lead to false positives/negatives

Table 1. continued from previous page.

Technology	Vendor	Principle	Advantages	Disadvantages
Glosensor	Promega	• A luciferase biosensor-based assay that uses genetically encoded biosensor with cAMP binding domain fused to a luciferase. • Binding of cAMP to the cAMP binding domain changes the conformation of the biosensor, leading to the activation of luciferase. The intensity of the luciferase signal is proportional to the amount of cAMP in the cells. • Gain-of-signal assay: increased cAMP level leads to increased signal.	• Homogeneous assay • Real-time live-cell kinetic measurement possible • Two different biosensors with different sensitivities to cAMP are available • Sensitivity may be high enough that activity of $G_{\alpha i}$ receptor can be assessed without the need to first stimulate cAMP production	• Requires expression of cAMP-dependent luciferase in cells • No cAMP calibration curves with an external standard
cAMP Direct Immunoassay Kit	Abcam / Biovision	• A competitive immunoassay by immobilizing anti-cAMP antibody in assay wells. • cAMP-HRP conjugate directly competes with cAMP from samples for binding to the cAMP antibody on the plate. • The signal intensity is inversely proportional to the concentration of cAMP in samples.	• Washing step removes potential interferences from compounds or cell culture medium • Does not require any recombinant protein expression for detection • Permits absolute quantification of cellular cAMP level by means of a standard curve	• Non-homogeneous assay

Flowchart of Assay Development Guidelines

1. Select cell lines or primary cells

- Recombinant cells generally produce greater signal change due to receptor expression .
- Endogenously expressed receptor from cell lines or primary cells can be used for more physiological relevance.

2. Select cAMP assay reagents of desired sensitivity

- cAMP antibody is available as femto2, dynamic2 or HiRange kit for different sensitivity.
- The choice is dependent on the cells and the density of assay format (i.e, 96-, 384-, or 1536-well).

3. Optimize cell density and forskolin concentration (G_i-coupled receptor)

- Cell density and forskolin concentrations should be titrated simultaneously to maximize potential assay window while keeping the measured signals within the linear range of the cAMP calibration curve.

4. Determine agonist dose-response relationships and potency values

- The measured signals are converted to cAMP levels, which are then used to determine the EC_{50} or EC_{80} values.

5. Determine antagonist dose-response relationships and EC_{50} values by stimulating with EC_{50-80} agonist concentration

- The measured signals are converted to cAMP levels, which are then used to determine the EC_{50} or EC_{80} values.

6. Assess assay performance statistics, including Z', signal-to-background ratio, reproducibility of potency data

- Ensure Z' > 0.5 and reproducibility over multiple runs are within acceptable limits.

Overview of the Assay Technology

The Cisbio cAMP assay uses homogeneous time-resolved fluorescence technology (HTRF) to measure cAMP in a non-separation, high throughput format. These kits are based on a competitive immunoassay using Eu^{3+} cryptate-labeled anti-cAMP antibody and d2-labeled cAMP (Figure 1). Detailed guidance on various immunoassay methods are described in the chapter "Immunoassay Methods" within this NIH Assay Guidance Manual (1). In the HTRF cAMP assay, binding of these two molecules will bring the long-lived fluorescent donor Eu^{3+} cryptate and short-lived florescent acceptor d2 into close proximity, which facilitates fluorescence resonance energy transfer to occur between the donor and the acceptor. The long fluorescent lifetime of Eu^{3+} cryptate allows the introduction of a time delay of approximately 50 to 150 microseconds between the system excitation and fluorescence emission measurement. Together with the large Stokes shift of the donor fluorophore (337 nm excitation to 620 nm emission), the Eu^{3+} cryptate-based time-resolved fluorescent technology eliminates non-specific short-lived emissions (e.g. fluorescence from compounds or plastic plates) (2). Furthermore, the use of a ratiometric readout between donor emission at 620 nm and acceptor emission at 665 nm is advantageous, particularly for reducing well-to-well variations or plate edge effects that are often seen in homogeneous assay formats. Cisbio has several cAMP kits that differ in the affinity of the antibody for cAMP and are suitable for detection of intracellular cAMP levels over several relative concentration ranges.

The use of positive and negative controls is important to define the dynamic range of the assay within each plate and for proper quality control from plate to plate and run to run. The inclusion of one or more reference compounds for concentration-response analysis is valuable for assessing the incidence of assay drift. More details can be found in the AGM chapters HTS Assay Validation (3) and Assay Operations for SAR Support (4).

Importance of cAMP Standard Curve

To establish the relationship between the actual cellular response and the assay readout, a cAMP standard curve should be included with each run as an assay control. The level of cAMP is measured indirectly through competition with a labeled cAMP competitor for binding to the cAMP antibody. The more cAMP produced, the lower the measured signal, and vice versa. The actual cAMP level is determined by using a standard curve in which the signals are measured using various known concentrations of cAMP. The cAMP standard curve is usually represented in a semi-log plot of $Em665/Em615 \times 10^4$ versus cAMP concentration (Figure 2A). The measured signal ratio displays an inverse sigmoidal relationship with the log concentration of cAMP. The cAMP level from an assay reaction is determined via interpolation from the standard curve (1).

During assay development, it is important that the assay conditions (e.g. cell density, forskolin concentration, and agonist stimulation) are optimized so that the measured signals fall within the linear range of the standard curve. Due to the sigmoidal relationship between the signal and the cAMP level, signals at the top region of the cAMP standard curve can be too sensitive to changes in cAMP level, whereas signals at the bottom region can be too insensitive to the changes.

Figure 2 illustrates the importance of converting the measured signals to cAMP levels to accurately determine the potency values from concentration-response curves. Figure 2A shows a typical cAMP standard curve. Figure 2B is a series of concentration response curves simulated to closely resemble actual data from the response characteristics of an agonist that mediates inhibition of forskolin-stimulated cAMP production in a $G_{\alpha i}$-coupled receptor. Increasing the concentration of the agonist causes a reduction in intracellular cAMP levels. Each of the nine curves shown in the panel represents different dilutions from the original lysate from 1X to 0.04X dilutions. Since the curves are simply generated by dilution of the same lysates, the agonist EC_{50} values are the same for all the curves (i.e., 100 nM). For illustration purposes, if these agonist curves are converted back to HTRF signal in accordance with the standard curve, another series of curves is generated as shown in Figure 2C. The estimated EC_{50} values from Figure 2C erroneously increase over a 4-fold range but none of these values accurately reflect

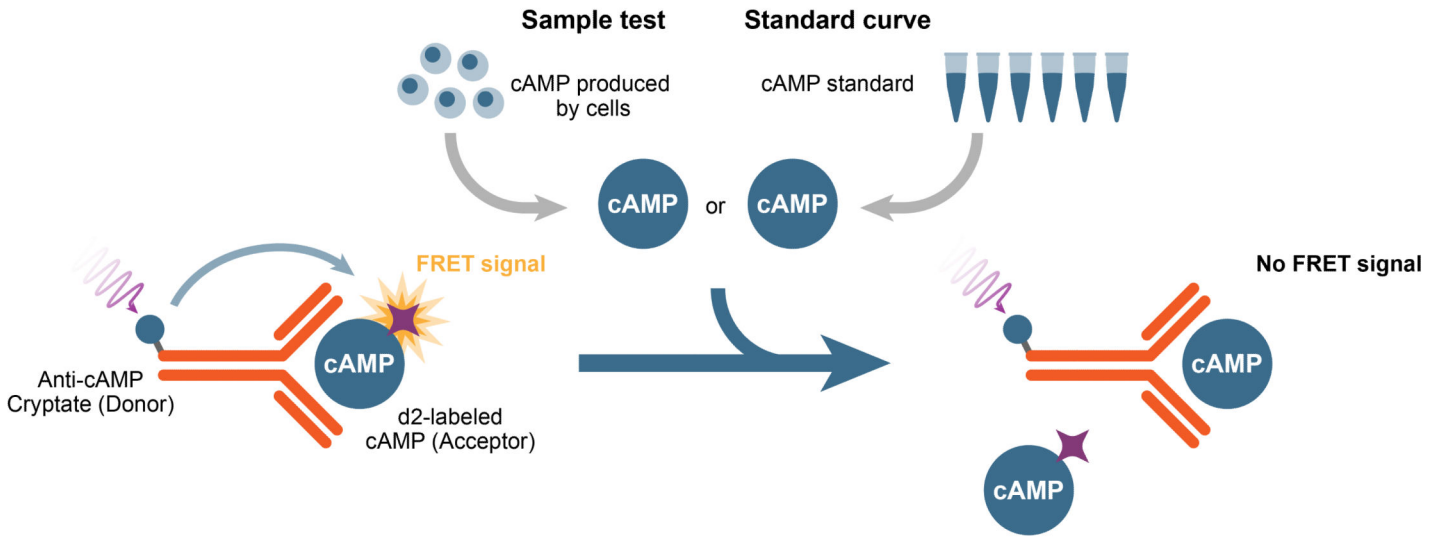


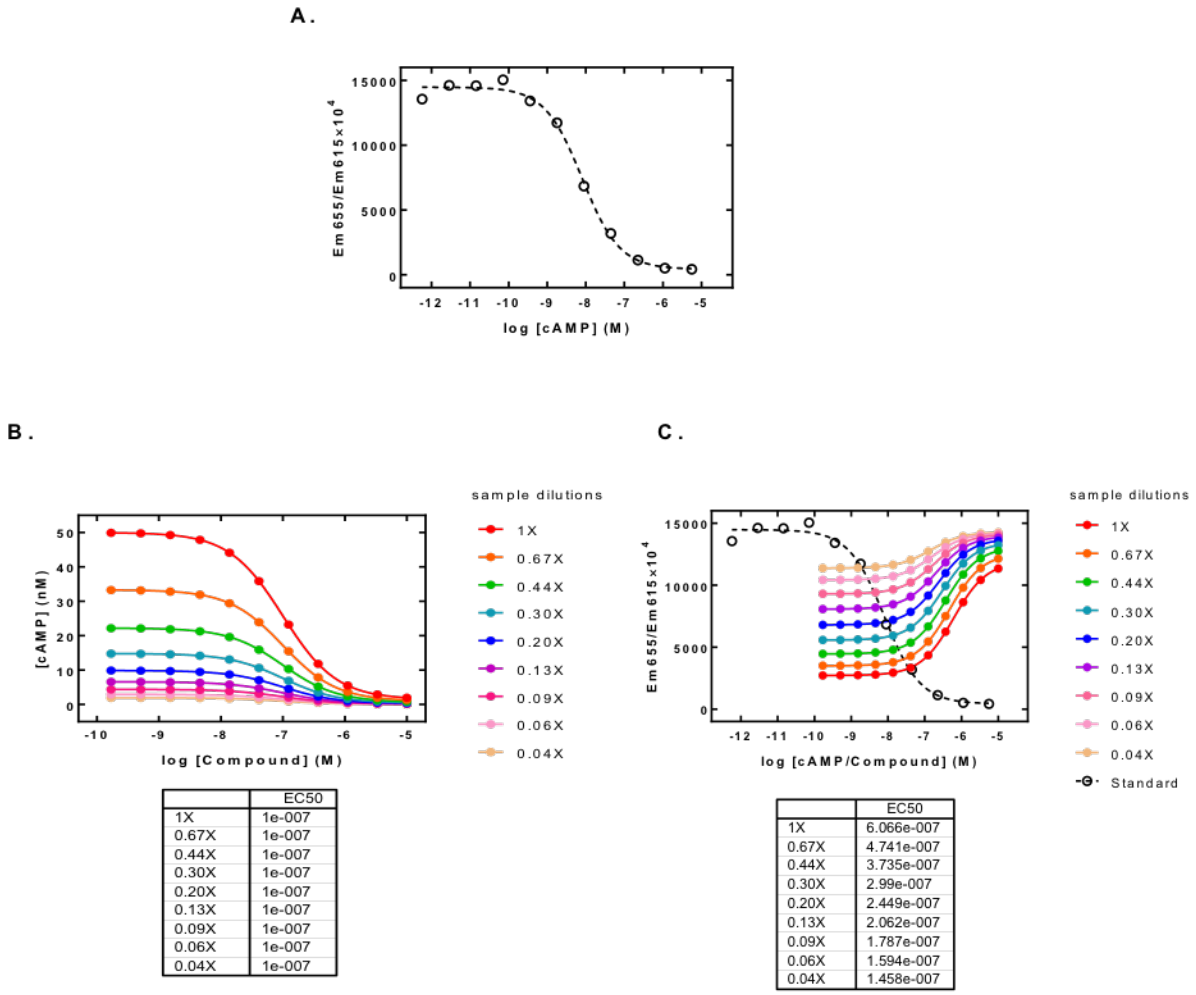
Figure 1. Principles of the HTRF cAMP Assay. Anti-cAMP cryptate and d2-labeled cAMP are the two main components of the detection reagents. cAMP produced intracellularly is detected following cell lysis via competition between the intracellular cAMP and d2-labeled cAMP for the anti-cAMP antibody. An increase in intracellular cAMP leads to disruption of FRET signal, whereas a decrease in intracellular cAMP results in higher FRET signal (2).

the actual EC_{50} value. This phenomenon is a result of the semi-logarithmic relationship between the measured signal and cAMP level. Similar discussion on the importance of using a cAMP standard curve to convert measured signal ratios to levels of cAMP and how basing the activity measurement on HTRF ratio may skew the potency and efficacy estimation are also described in the literature (5,7). Furthermore, the use of a calibration curve and the selection of an appropriate model is thoroughly discussed in the AGM chapter Immunoassay Methods (1).

Assay Development and Optimization Guidelines

As indicated in the above flowchart, assay development should start from selecting the appropriate endogenously expressing cells or transfecting cells to overexpress the target receptor. If cell lines stably expressing the receptors are generated, multiple clones can be selected initially either by FACS to assess the surface protein expression or by qPCR for gene expression. The level of receptor expressed on the cell surface impacts both the potency and efficacy of agonists. Figure 3 illustrates how the estimated potency values increase for a set of agonists as the expression level of target receptor increases. Furthermore, a partial agonist may behave as a full agonist due to a high level of receptor expression, which leads to high receptor reserve. During the early lead discovery phase, it may be advantageous to utilize cell lines of high receptor expression to enhance the sensitivity of detecting hit compounds. Nevertheless, it would be important to assess multiple cell clones expressing receptors at various levels if correlation with other assay readouts is desirable in the lead optimization phase.

Once a cellular model is selected or generated, for a $G_{\alpha s}$-coupled receptor, assay development may begin by determining the concentration-response relationship of a known agonist at various cell densities. Figure 4 shows an example of such agonist and cell density titration curves. As shown in the figure, it is important to note that when there are too many cells, the signal-to-background may be adversely affected. The agonist concentration-response curves are examined to identify concentrations in which the response falls within the linear detection range of the cAMP standard curve. A cell density that produces the best dynamic range and expected EC_{50} value of the agonist is selected.



	EC50
1X	1e-007
0.67X	1e-007
0.44X	1e-007
0.30X	1e-007
0.20X	1e-007
0.13X	1e-007
0.09X	1e-007
0.06X	1e-007
0.04X	1e-007

	EC50
1X	6.066e-007
0.67X	4.741e-007
0.44X	3.735e-007
0.30X	2.99e-007
0.20X	2.449e-007
0.13X	2.062e-007
0.09X	1.787e-007
0.06X	1.594e-007
0.04X	1.458e-007

Figure 2. Illustration of the importance of estimating potency values using concentration-response curves expressed in cAMP levels instead of HTRF signal. Panel A. A typical cAMP standard curve. Panel B. Agonist concentration-response curves expressed in cAMP levels and the estimated EC_{50} values. Panel C. Corresponding concentration-response curves expressed in HTRF signal and the estimated EC_{50} values.

For a $G_{\alpha i}$-coupled receptor, its activation needs to be assessed under a detectable cAMP level. This can be achieved using forskolin or a known agonist of a $G_{\alpha s}$-coupled receptor to stimulate cAMP production, as shown in Figure 5. Similarly, a combination of appropriate forskolin concentration and cell density that can stimulate cAMP production to a level that is within the linear detection range and generate optimal assay dynamic range would be selected. Subsequently, the agonist response can be assessed at the chosen conditions to ensure its potency agrees with anticipated values.

For either $G_{\alpha s}$- or $G_{\alpha i}$-coupled receptors, following the determination of agonist response, the potency of antagonist can be assayed at a selected agonist concentration between EC_{50} to EC_{80}. A summary of assay optimization and troubleshooting guidelines is given in Table 2.

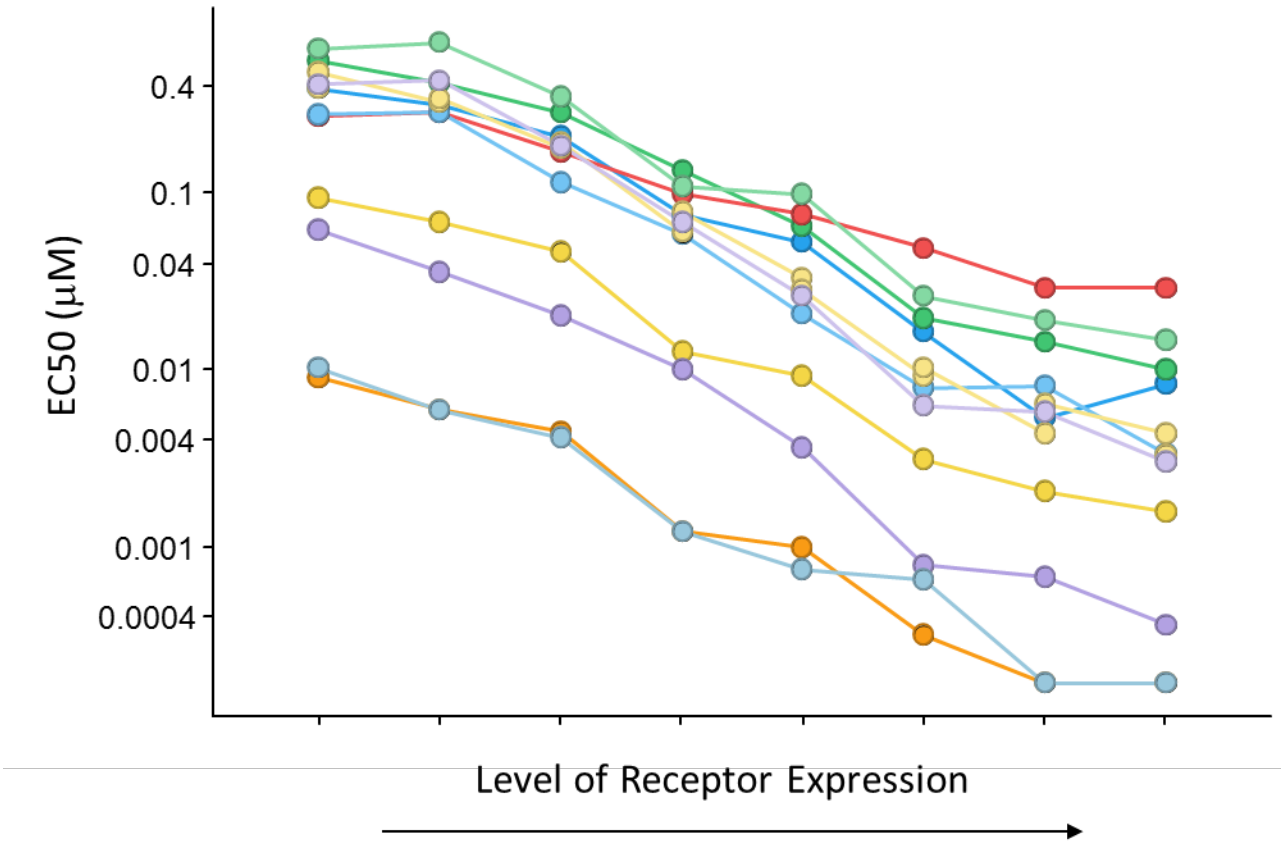
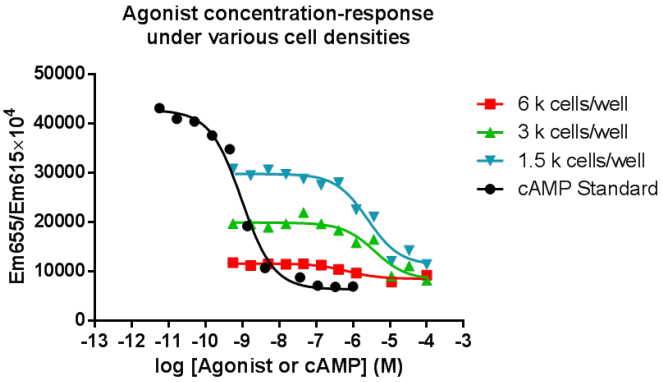
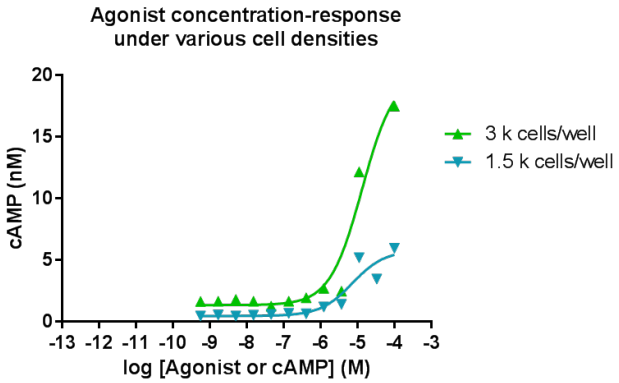


Figure 3. Dependence of agonist potency (EC_{50} values) on the level of receptor expression. Shown are individual compound potency values that were measured in cells lines expressing a given level of receptor. Different compounds are shown in symbols and lines of different colors.

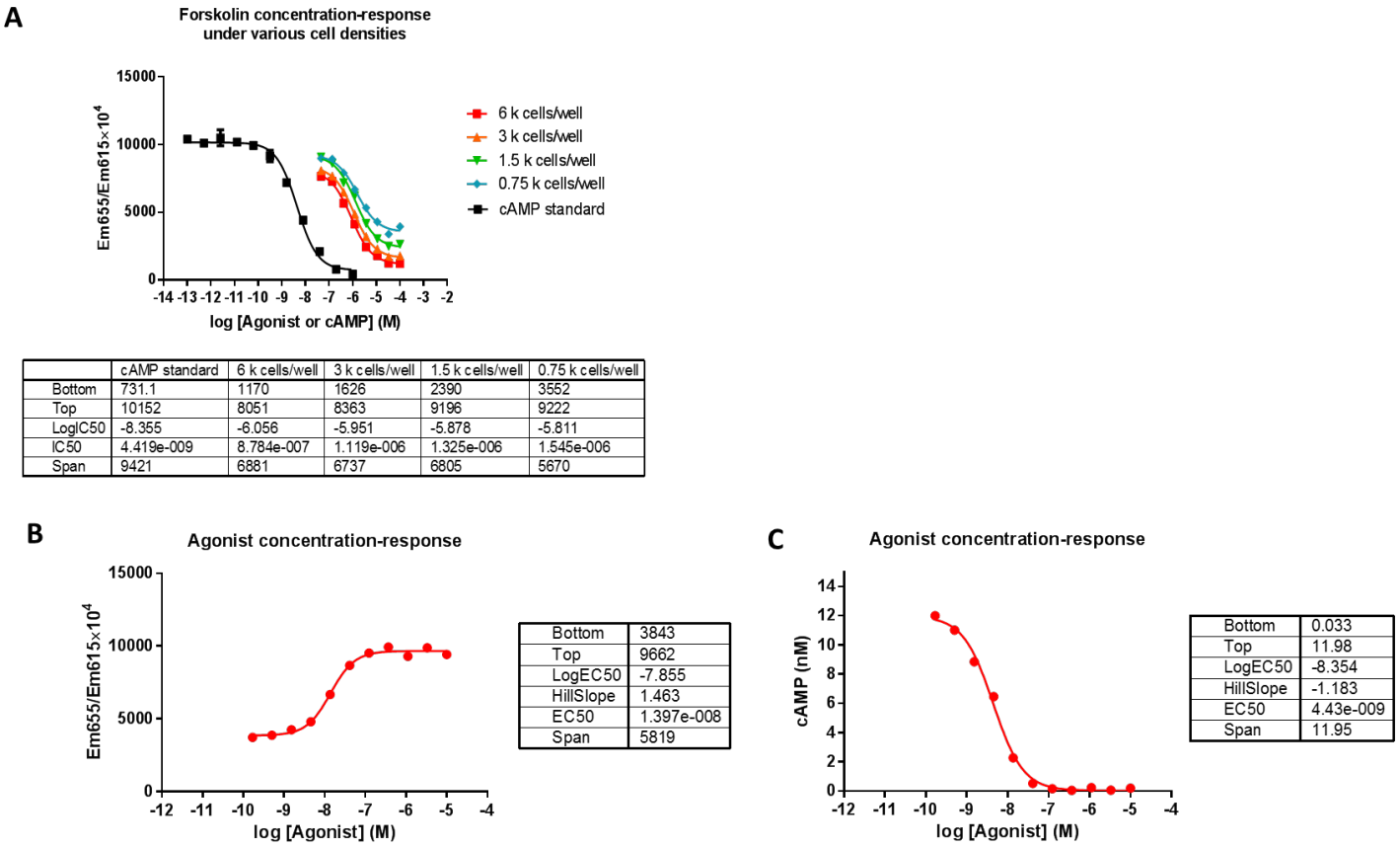


	6 k cells/well	3 k cells/well	1.5 k cells/well	cAMP Standard
Bottom	8450	8255	11249	6304
Top	11535	19842	29739	42671
LogIC50	-6.152	-5.409	-5.601	-9.017
IC50	7.047e-007	3.899e-006	2.508e-006	9.622e-010
Span	3085	11587	18490	36367



	3 k cells/well	1.5 k cells/well
Bottom	20.13	5.803
Top	1.365	0.48
LogIC50	-4.876	-5.135
IC50	1.33e-005	7.323e-006
Span	-18.77	-5.323

Figure 4. Examples of (A) agonist-concentration response curves at various cell densities for a $G_{\alpha s}$-coupled receptor and (B) agonist-concentration response curves expressed in cAMP level. In this example, when the cell density is too high as with 6 k/cells, the signals produced exceeded the optimal range of cAMP level (panel A) and therefore the curve cannot be converted to cAMP value (panel B).



	cAMP standard	6 k cells/well	3 k cells/well	1.5 k cells/well	0.75 k cells/well
Bottom	731.1	1170	1626	2390	3552
Top	10152	8051	8363	9196	9222
LogIC50	-8.355	-6.056	-5.951	-5.878	-5.811
IC50	4.419e-009	8.784e-007	1.119e-006	1.325e-006	1.545e-006
Span	9421	6881	6737	6805	5670

Bottom	3843
Top	9662
LogEC50	-7.855
HillSlope	1.463
EC50	1.397e-008
Span	5819

Bottom	0.033
Top	11.98
LogEC50	-8.354
HillSlope	-1.183
EC50	4.43e-009
Span	11.95

Figure 5. Examples of (A) forskolin concentration response-curves at various cell densities for a $G_{\alpha i}$-coupled receptor, (B) an agonist concentration-response assayed at 3 μM forskolin and 1,500 cells/well as selected from Panel A, and (C) the agonist concentration-response from Panel B expressed in cAMP level.

Table 2. Key assay optimization parameters and troubleshooting guidelines

Problem	Possible reason	Solution
$G_{\alpha s}$ agonist assay		
Basal cAMP falls out of the lowest measurable limit in the linear part of the cAMP standard curve (i.e., cAMP level too low)	Insufficient number of cells	Increase the number of cells in the assay reaction
Basal cAMP exceeds the highest measurable limit in the linear part of the cAMP standard curve (i.e., cAMP level too high)	Too many cells	Decrease the number of cells in the assay reaction
Agonist-stimulated cAMP level undetectable (i.e., cAMP level too low)	cAMP rapidly degraded by endogenous phosphodiesterase	Include phosphodiesterase inhibitors such as IBMX
	Cells not expressing sufficient quantities of receptor	Select for higher expressing clones as determined by qPCR or by FACS
	Compounds not efficacious	Test other compounds known to be efficacious in modulating cAMP production
	Insufficient compound binding to the receptor	Allow for longer incubation of the assay reactions

Table 2. continued from previous page.

Problem	Possible reason	Solution
Agonist-stimulated cAMP level exceeds the linear part of the cAMP standard curve (i.e., cAMP level too high)	Phosphodiesterase inhibition too strong	Test the presence and absence of phosphodiesterase inhibitor
	cAMP detection antibody too sensitive	Select for an anti-cAMP antibody with lower affinity for cAMP
	Too many cells	Decrease the number of cells in the assay reaction
Gαi agonist assay		
Basal cAMP falls out of the lowest detection limit in the cAMP standard curve (i.e., cAMP level too low)	Insufficient number of cells	Increase the number of cells in the assay reaction
Forskolin-stimulated cAMP not achieving high enough level (i.e., cAMP level too low)	Not enough cells	Increase the number of cells in the assay reaction
	cAMP rapidly degraded by endogenous	Include phosphodiesterase inhibitors such as IBMX
	Forskolin concentration too low	Increase the concentration of forskolin as determined in forskolin concentration-response curves
Forskolin stimulated cAMP exceeded the linear part of the cAMP standard curve (i.e., cAMP level too high)	Too many cells	Decrease the number of cells in the assay reaction
	Forskolin concentration too high	Reduce the concentration of forskolin as determined in forskolin concentration-response curves
Agonist-stimulated change in cAMP level undetectable	Cells not expressing sufficient quantities of receptor	Select for higher expressing clones as determined by qPCR or by FACS
	Compounds not efficacious	Test other compounds known to be efficacious in modulating cAMP production
	Insufficient compounds binding to the receptor	Allow for longer incubation of the assay reactions
$G_{\alpha s}$ antagonist assay		
Antagonist compounds do not antagonize the agonist-stimulated response or not giving expected potency	Insufficient antagonist compounds binding to the receptor	Allow for longer pre-incubation of the antagonist compounds
	Compounds not efficacious	Test other compounds known to be efficacious in modulating cAMP production
	Agonist stimulation too strong	Reduce agonist concentration used for stimulation
$G_{\alpha i}$ antagonist assay		
Antagonist compounds do not antagonize the agonist-stimulated response or not giving expected potency	Insufficient antagonist compounds binding to the receptor	Allow for longer pre-incubation of the antagonist compounds
	Compounds not efficacious	Test other compounds known to be efficacious in modulating cAMP production
	Agonist stimulation too strong	Reduce agonist concentration used for stimulation
	Accumulated cAMP not degraded	Test presence or absence of phosphodiesterase inhibitor

cAMP Assays for Receptors Coupled to $G_{\alpha s}$ Protein

The following guidelines were developed for assays using suspension cells for high throughput mode; however, the assay can be adapted easily for adherent cells. Volumes are given for a 384-well assay (20 µL assay reaction) and can be adjusted proportionally to other plate densities.

$G_{\alpha s}$-Coupled Agonist Format

The activation of adelnylate cyclase by $G_{\alpha s}$, in the presence of an agonist, the generation of intracellular cAMP and its measurement by HTRF technology is shown in Figure 6.

Sample Protocol for $G_{\alpha s}$-Coupled Agonist Assays

Step 1: Grow cells in tissue culture flasks for 1-3 days in the corresponding cell culture medium.

- Select a cell line containing the target GPCR of interest.
- A negative control cell line as a counter screen may either be the parental cell line devoid of the target or a cell line expressing an unrelated GPCR that couples to the same G protein.
- Make sure the cells are healthy and active in the log phase of growth and are maintained in the same way in all experiments, as the state of the cells may affect the receptor response.
- Limit cell passage in concordance with assay performance statistics. Ideally, to avoid variation in cellular response due to changes in the state of the cells or receptor expression, expand the cell culture to the quantities required for the entire screening campaign in one large batch of a single passage and prepare cryopreserved cells as one batch. Various manufacturers of cell culture reagents offer different cryopreservation reagents and methods. During screening, recover the cells and use them as needed.

Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes.

- Use of Trypsin or other enzymatic/non-enzymatic cell dissociation agents should be tested prior to screening.
- Some receptor proteins may tolerate trypsinization, but other enzymatic reagents such as TrypLE (Thermo) or non-enzymatic reagents such as CellStripper (Corning) may be used if needed.

Step 3: Aspirate the medium and re-suspend the cells in assay buffer. Count and dilute the cells to the proper densities in assay buffer.

- In most cases, HBSS or Dulbecco's PBS (DPBS) buffers can be used: 1X HBSS/20mM HEPES with or without 0.1% BSA (fat acid-free) or DPBS with Ca^{2+}/Mg^{2+} with or without 0.1% BSA.
- Cell densities should be determined during assay development in cell titration experiments such that the measured signal falls within the linear range of the cAMP standard curve while yielding the best assay window determined with a reference agonist.
- Using too many cells, and hence the presence of too many receptors, may reduce the effective free concentrations of ligands and cause the assay to bottom-out and therefore limit the ability to differentiate potent compounds. Too many cells may also saturate the cAMP assay reagents.

Step 4: Add 3-isobutyl-1-methylxanthine (IBMX) or other phosphodiesterase inhibitor to the cell suspension. IBMX is a competitive nonselective *phosphodiesterase inhibitor* which inhibits the degradation intracellular cAMP. It is generally used at a final concentration of 0.1 mM in the assay, but the impact of phosphodiesterase inhibitor on compound potency and the required concentration should be determined according to the assays and cell lines being used. Dispense 10 µL of cells to assay plates pre-dotted with compounds. *(Note that depending on the extent of activity of phosphodiesterase in the cells, the use of inhibitor is optional.)*

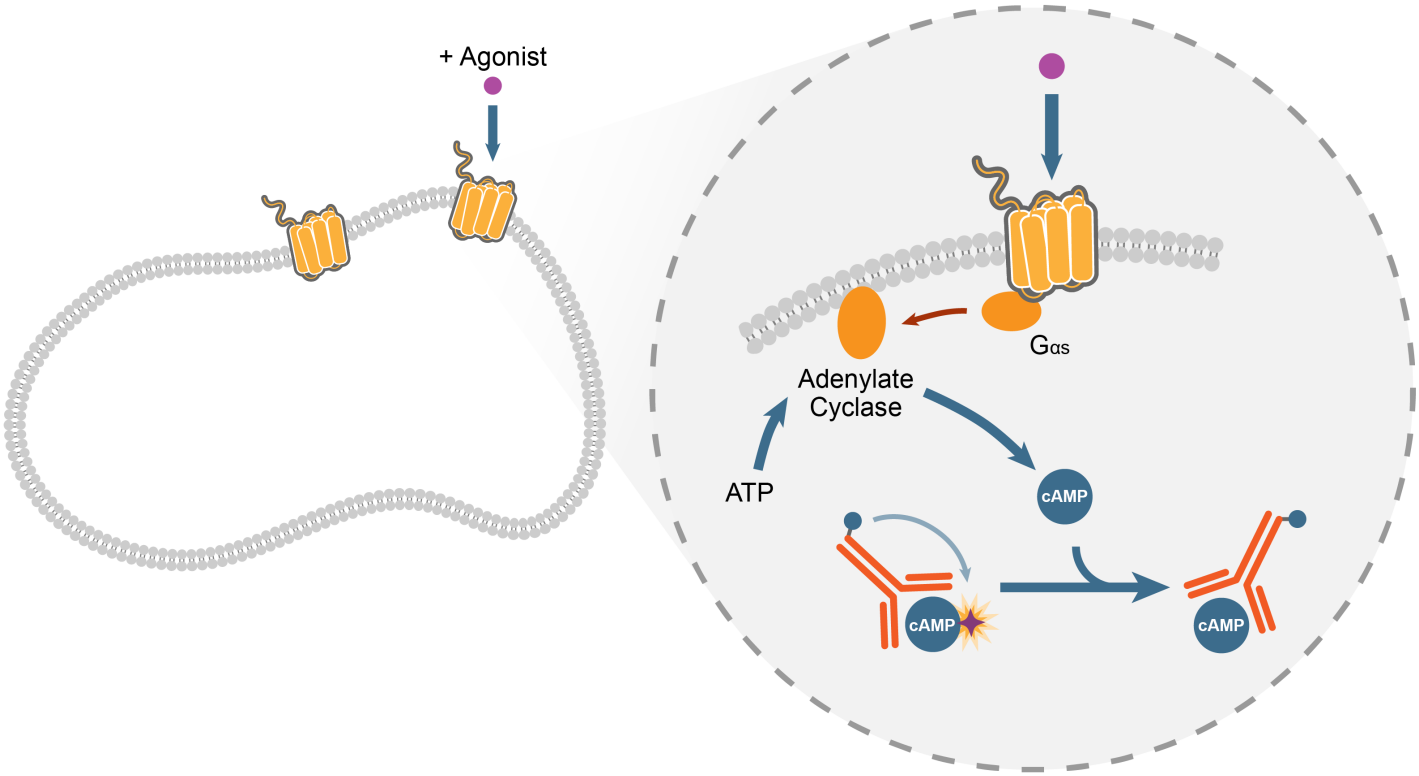


Figure 6. cAMP Measurement for Agonists of a $G_{\alpha s}$-Coupled Receptor. Activation of the receptor by an agonist causes an exchange of GDP for GTP in the G protein complex that subsequently dissociates into a $G_{\beta\gamma}$ dimer and a $G_{\alpha s}$ monomer. The activated $G_{\alpha s}$ subunit binds to adenylate cyclase resulting in the generation of intracellular cAMP, which is measured using the specific Eu^{3+}-labeled anti-cAMP antibody and d2-labeled cAMP in a competitive HTRF assay.

- IBMX can be made at high concentration (0.5 M) and aliquoted and stored at -20°C to avoid multiple freeze-thaw cycles.
- Different assay plates may be tested. Please refer to Cisbio website for microplate recommendations (http://www.cisbio.com/usa/drug-discovery/htrf-microplate-recommendations).
- The cell suspension is best prepared only when it is ready for dispensing to avoid deterioration of cellular response. In a large screening campaign, longevity of the cell suspension should be determined from the concentration-response curves.
- Compounds dissolved in DMSO can be pre-dotted into assay plates in sub-microliter volume using an acoustic dispenser, such as Labcyte Echo or EDC Biosytems ATS, or pre-diluted in an assay buffer prior to addition. Tolerance of the cellular response to DMSO should be investigated during assay development. Most cells can tolerate up to 1% DMSO in the assay. An example of a DMSO tolerance test is shown in Figure 7.
- Cells and other reagents can be dispensed by various peristaltic pump-based liquid handlers, such as Thermo Combi-drop or the BioTek Washer Dispenser, or any of the widely available tip-based liquid handlers.

Step 5: Cover the assay plate and incubate for 30 minutes at room temperature (RT).

- The duration of assay incubation depends on the kinetic properties of the compounds. Potency values can be underestimated if the incubation is insufficient to achieve steady state conditions.

Step 6: Prepare the cAMP standard curve according to the manufacturer's instructions.

Step 7: Dispense 5 µL/well of diluted d2-labeled cAMP conjugate followed by 5 µL/well of cryptate-labeled anti-cAMP antibody to the cell plate and standard curve plate. Both MultiDrop and Tempest liquid handling instruments are compatible for HTRF reagents.

- d2-labeled cAMP conjugate and cryptate-labeled anti-cAMP antibody are reconstituted per manufacturer's instructions.
- Working solutions are prepared by diluting the concentrated stock 20-fold using lysis buffer provided in the kit.

Step 8: The plates are incubated for 1 hour at RT and read using instruments such as a PerkinElmer Envision with a protocol that is set and optimized for HTRF detection.

- Fluorescence intensity is measured at emission of 665 nm and 615 nm (Em665 and Em615), with excitation at 350 nm.
- Plates can be read repeatedly to ensure that the antibody binding reaches steady state.

Step 9: Convert the results from the reader into cAMP levels using the cAMP standard curve.

- Readings are generally expressed as Em665/Em615×10^4, but may vary depending on the exact detection instrument being used.

Step 10: Data analysis. See Figure 8.

$G_{\alpha s}$-Coupled Antagonist Format

In this format, an antagonist displaces an agonist for a $G_{\alpha s}$ coupled receptor and the resulting inhibition of adenylate cyclase is measured using the HTRF assays, as shown in Figure 9.

Sample Protocol for $G_{\alpha s}$-Coupled Antagonist Assays

The procedure for determination of antagonist activity of a $G_{\alpha s}$-coupled receptor using cAMP measurement is essentially the same as the procedure described above for detection of agonist activity. In the antagonist format, an agonist is added at concentrations that trigger 50-80% of the maximum response ($EC_{50\text{-}80}$) of the agonist. The rationale and precautions for each step can be referred to in the above agonist assay protocol, except for those specific to antagonist assays. All compounds of interest should also be tested in the agonist mode to ensure the absence of agonist activities.

Step 1: Grow cells in flask for 1-3 days in the corresponding cell culture medium.

Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes.

Step 3: Aspirate the medium and re-suspend the cells in assay buffer. Count and dilute the cells to the desired concentration in assay buffer.

- The working concentration of the cell suspension is twice as much as in agonist mode due to the volume difference. For example, if 1×10^6 cell/ml is used in the agonist assay, then 2×10^6 cells/ml is prepared for the antagonist assay.

Step 4: Add IBMX, which is generally used at a final concentration of 0.1 mM in the assay, but the impact of phosphodiesterase inhibitor on compound potency and the required concentration should be determined according to the assays and cell lines being used. Dispense 5 µL of cells into assay plates that have been pre-dotted with test compound in DMSO or pre-diluted in an assay buffer. *(Note that depending on the extent of activity of phosphodiesterase in the cells, the use of inhibitor is optional.)*

Step 5: Cover the plate and incubate for 15-30 minutes at room temperature (RT).

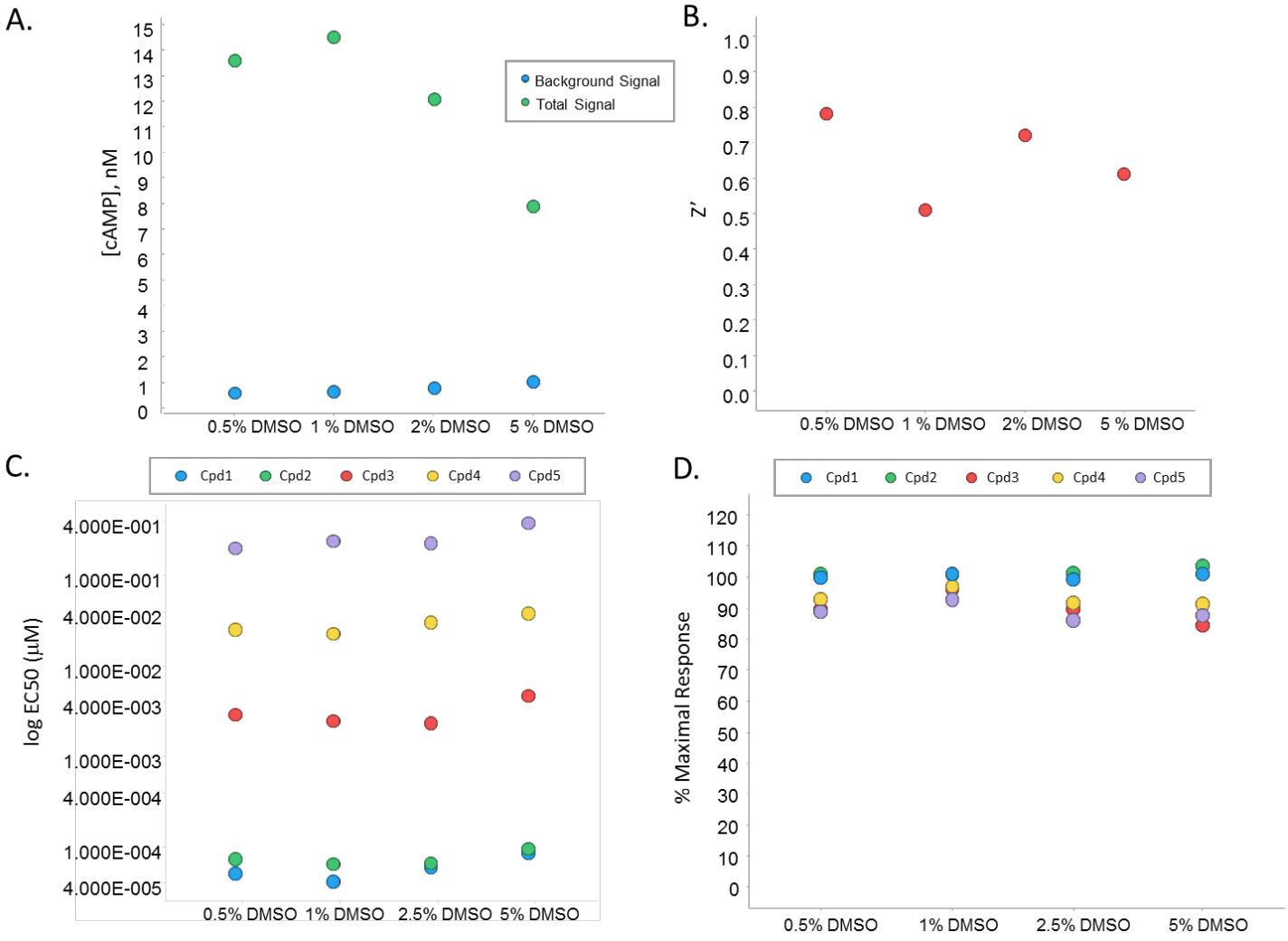


Figure 7. An example of a DMSO tolerance test carried out by conducting the cAMP assay in the presence of different concentrations of DMSO. Various parameters are assessed to identify the acceptable range of DMSO without adversely affecting the performance of the assay and the pharmacological measurements. The parameters are: (A) levels of cAMP in the presence of an agonist (total signal) and absence of the agonist (background signal), (B) Z' as described in (6), (C) agonist EC_{50} values, and (D) % maximal agonist response. In panels C and D, symbols of different colors denote different agonists.

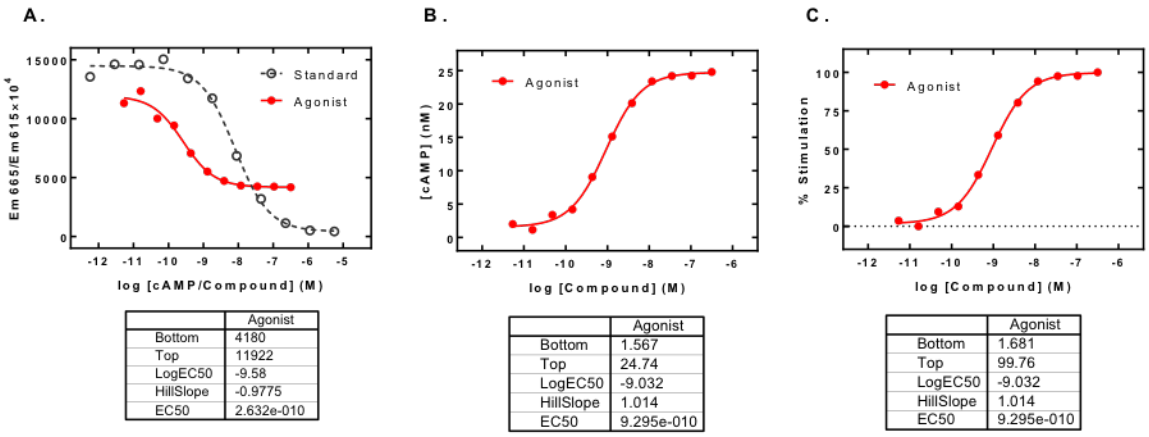
- Depending on the kinetic properties of test compounds relative to the agonist used for stimulation, this step can be adjusted from no incubation to an extended incubation time as needed.

Step 6: Add 5 µL of the agonist solution diluted in assay buffer. Incubate the plates at RT for 30 minutes.

- The final concentration of the agonist chosen should result in 50 to 80% of the maximum response of the agonist, as determined during development of the agonist assay.
- The working concentration is prepared at 2X the desired final concentration.
- At the chosen $EC_{50\text{-}80}$ concentration of agonist, the measured agonist response should not exceed the linear range of the cAMP standard curve.

Step 7: Prepare the standard curve according to the manufacturer's instructions.

Step 8: Dispense 5 µL of diluted d2-labeled cAMP conjugate followed by 5 µL of cryptate-labeled anti-cAMP antibody into the cell plate and the standard curve plate.



	Agonist
Bottom	4180
Top	11922
LogEC50	-9.58
HillSlope	-0.9775
EC50	2.632e-010

	Agonist
Bottom	1.567
Top	24.74
LogEC50	-9.032
HillSlope	1.014
EC50	9.295e-010

	Agonist
Bottom	1.681
Top	99.76
LogEC50	-9.032
HillSlope	1.014
EC50	9.295e-010

Figure 8. An example of an agonist concentration-response curve in $G_{\alpha s}$-coupled receptors. Panel A. The agonist response is plotted as the HTRF ratio (Em665/Em615×10^4) along with the cAMP standard curve. Note that the agonist produces signals that are within the linear portion of the standard curve. Panel B. The agonist response is plotted as the cAMP level to determine the EC_{50} values. Panel C. The same agonist response is plotted as % stimulation. Data are analyzed using a four-parameter logistic regression.

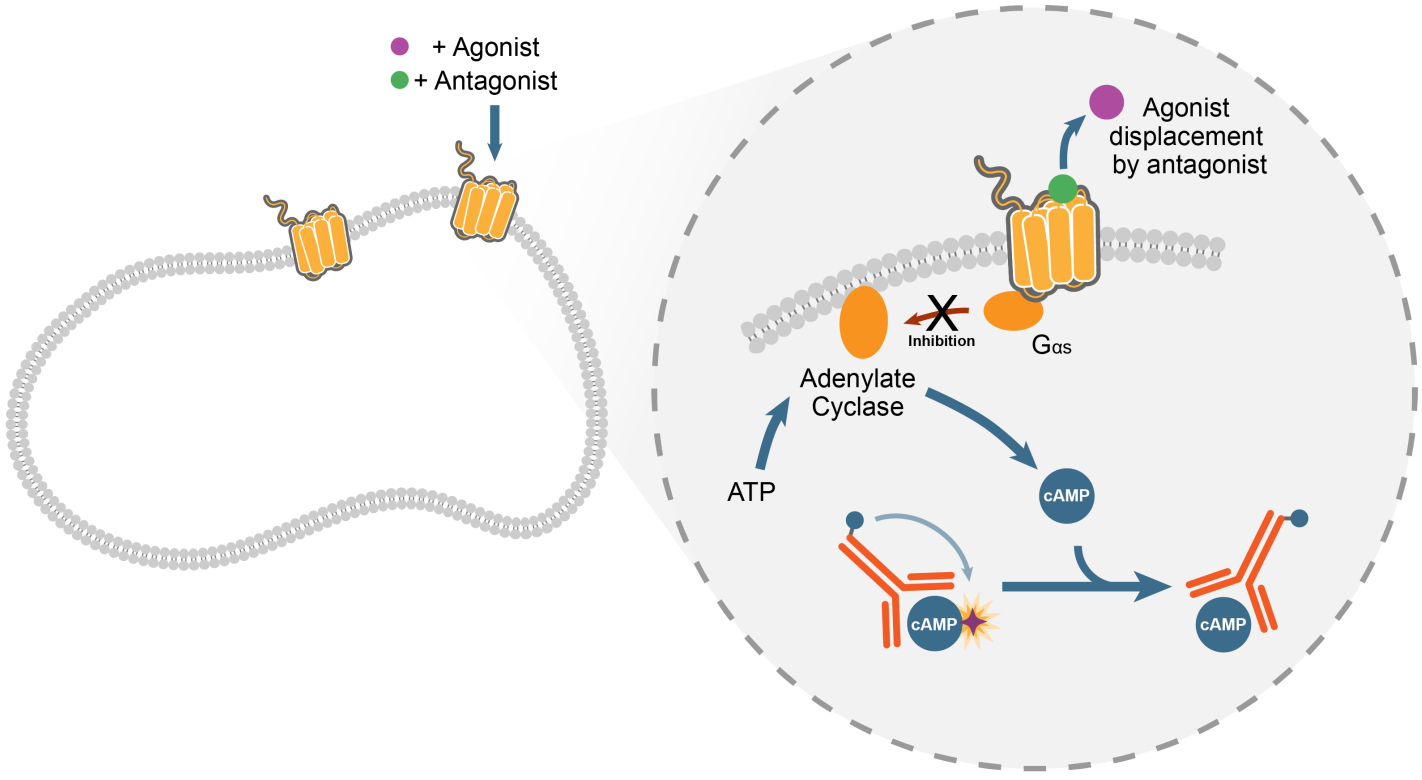


Figure 9. cAMP Measurement for Antagonists of a $G_{\alpha s}$-Coupled Receptor. An antagonist prevents the activation of the receptor by an agonist which prevents G protein complex dissociation. The adenylate cyclase is not activated and the generation of intracellular cAMP does not occur.

Step 9: Incubate the plates for 1 hour at RT and measure the signal with a PerkinElmer Envision reader (or equivalent) using a protocol that is optimized for HTRF measurements.

Step 10: Convert the results from the reader into cAMP levels using the cAMP standard curve.

- Readings are generally expressed as Em665/Em615×10^4.

Step 11: Data analysis. See Figure 10.

cAMP Assays for Receptors Coupled to $G_{\alpha i}$ Protein

The following guidelines are developed for assays carried out using suspension cells for high-throughput mode, although the assay can be adapted easily for adherent cells. Volumes are given for a 384-well assay (20 µL assay reactions) and can be adjusted proportionally to other plate density formats.

$G_{\alpha i}$-Coupled Agonist Format

In assaying a $G_{\alpha i}$ -coupled receptor, forskolin is typically used to increase the level of cAMP so that the lowering of the cAMP level due to receptor negative regulation of adenylate cyclase can be observed (Figure 11). Alternatively, the cells can be stimulated to produce cAMP by an agonist to a $G_{\alpha s}$-coupled receptor that is also present in the cells.

Sample Protocol for $G_{\alpha i}$-Coupled Agonist Assays

Step 1: Grow cells in tissue culture flasks for 1-3 days in the corresponding cell culture medium.

- Select a cell line containing the target GPCR of interest.
- A negative control cell line as a counter screen may either be the parental cell line devoid of the target or a cell line expressing an unrelated GPCR that couples to the same G protein.
- Make sure the cells are healthy and active in the log phase of growth and are maintained in the same way in all experiments, as the state of the cells may affect the receptor response.
- Control cell passage number based on assay performance. Ideally, to avoid variation in cellular response due to changes in the state of the cells or receptor expression, expand the cell culture to the quantities required for the entire screening campaign in one large batch of a single passage and prepare cryopreserved cells as one batch. Various manufacturers of cell culture reagents offer different cryopreservation reagents and methods. During screening, recover the cells and use them as needed.

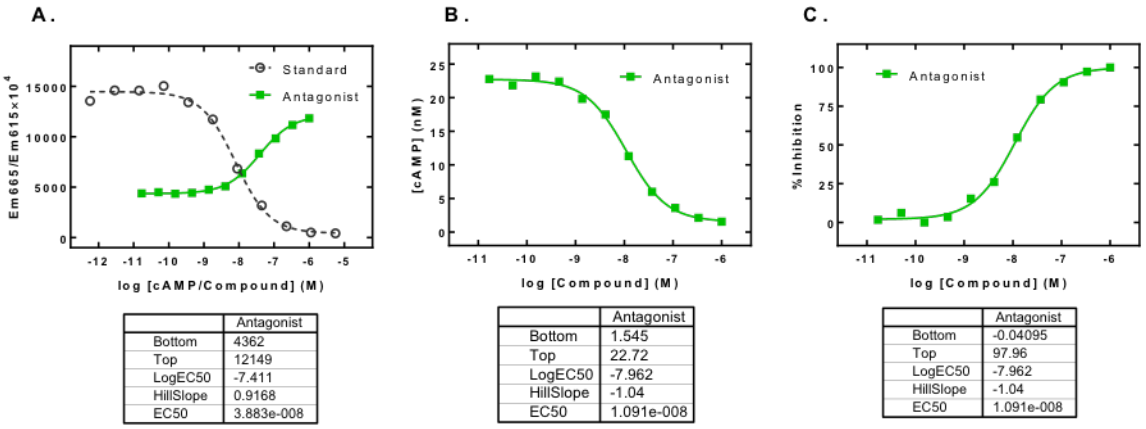
Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes.

- Use of Trypsin or other enzymatic/non-enzymatic cell dissociation agents should be tested prior to screening.
- Some receptor proteins may tolerate trypsinization, but other enzymatic reagents such as TrypLE (Thermo) or non-enzymatic reagents such as CellStripper (Corning) may be used if needed.

Step 3: Aspirate the cell culture medium and re-suspend the cells in assay buffer. Count and dilute the cells to the proper densities in assay buffer.

- In most cases, HBSS or Dulbecco's PBS (DPBS) buffers can be used: 1X HBSS/20mM HEPES with or without 0.1% BSA (fat acid-free) or DPBS with Ca^{2+}/Mg^{2+} with or without 0.1% BSA.
- Cell densities should be determined during assay development in cell titration experiments such that the measured signal falls within the linear range of the cAMP standard curve while yielding the best assay window.
- Using too many cells, and hence the presence of too many receptors, may reduce the effective free concentrations of ligands and cause the assay to bottom-out and therefore limit the ability to differentiate potent compounds. Too many cells may also saturate the cAMP assay reagents.

Step 4: Prepare a solution containing the desired amount of forskolin (to stimulate cAMP production so that the inhibitory effect of agonist on the cAMP level can be detected.) Add IBMX, generally at a final concentration of



	Antagonist
Bottom	4362
Top	12149
LogEC50	-7.411
HillSlope	0.9168
EC50	3.883e-008

	Antagonist
Bottom	1.545
Top	22.72
LogEC50	-7.962
HillSlope	-1.04
EC50	1.091e-008

	Antagonist
Bottom	-0.04095
Top	97.96
LogEC50	-7.962
HillSlope	-1.04
EC50	1.091e-008

Figure 10. An example of antagonist concentration-response curves in $G_{\alpha s}$-coupled receptors. Panel A. The antagonist response is plotted as the HTRF ratio ($Em665/Em615\times10^4$) along with the cAMP standard curve. Note that the antagonist produces signals that are within the linear portion of the standard curve. Panel B. The antagonist response is plotted as the cAMP level to determine the IC_{50} values, respectively. Panel C. The same antagonist response is plotted as % Inhibition. Data are analyzed using four-parameter logistic regression.

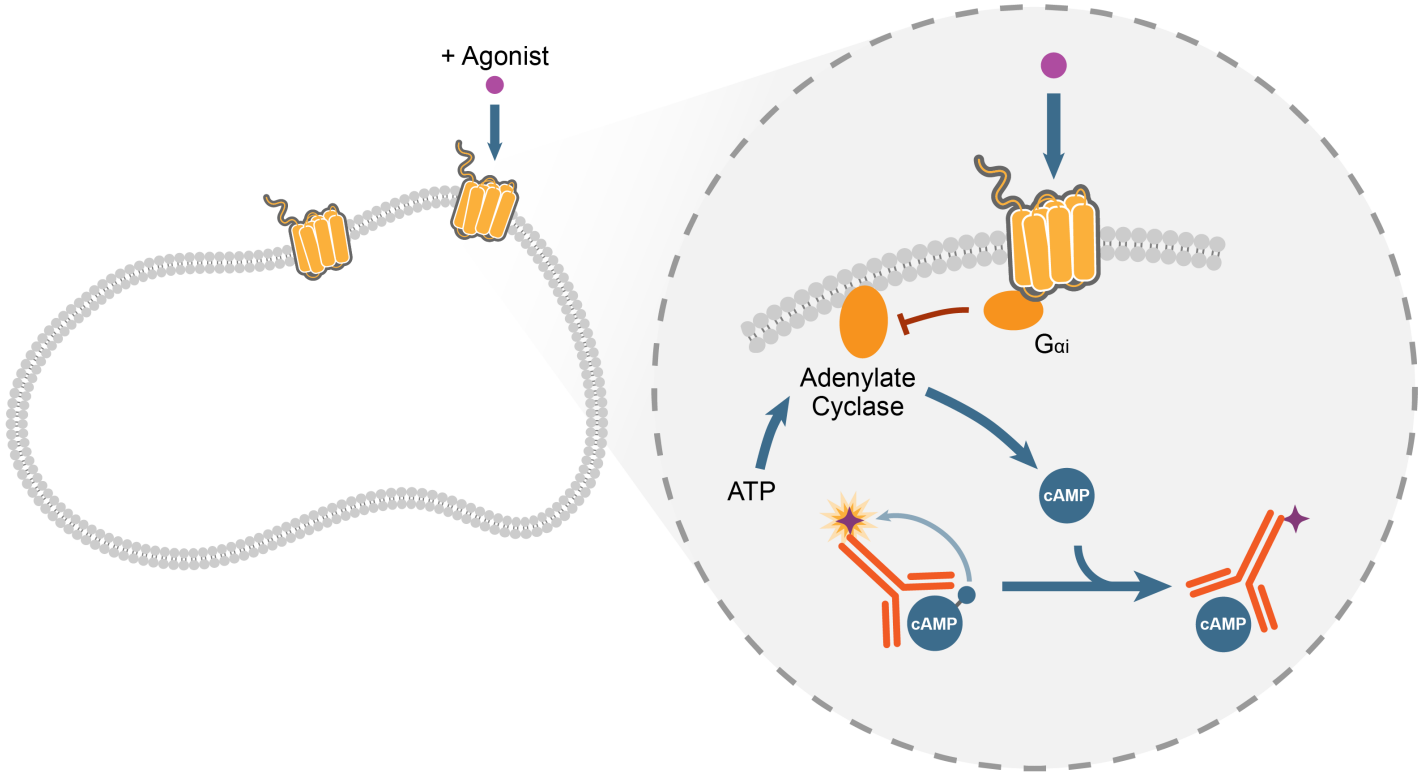


Figure 11. cAMP Measurement for Agonists of a $G_{\alpha i}$-Coupled Receptor. Treatment of cells with forskolin results in an increase in cellular cAMP levels. The addition of an agonist decouples the G protein complex that subsequently dissociates into a $G_{\beta\gamma}$ dimer and a $G_{\alpha i}$ monomer. The activated $G_{\alpha i}$ subunit binds to adenylate cyclase resulting in the negative regulation of intracellular cAMP (no intracellular cAMP is generated).

0.1 mM, or other phosphodiesterase inhibitor and dispense 5 µL of the mixture into assay plates pre-dotted with compounds, followed by dispensing 5 µL of cell suspension. *(Depending on the extent of activity of*

phosphodiesterase in the cells, the use of an inhibitor is optional. The impact of a phosphodiesterase inhibitor on compound potency and the required concentration should be determined according to the assays and cell lines being used.)

- Forskolin can be prepared as a 10 mM stock in DMSO in a glass vial and kept at RT.
- The concentration of forskolin to be used in the assay should be optimized to ensure that the measured signal falls within the linear range of the cAMP standard curve while yielding the best assay window. Cell density and forskolin concentration can be optimized simultaneously by a three-way titration that includes cells and forskolin at different concentrations in the absence and presence of a reference agonist.
- IBMX can be made at high concentration (0.5 M), aliquoted and stored at -20°C to avoid multiple freeze-thaw cycles.
- Different assay plates may be tested. Please refer to the Cisbio website for microplate recommendations (http://www.cisbio.com/usa/drug-discovery/htrf-microplate-recommendations).
- The cell suspension is best prepared only when it is ready for dispensing to avoid deterioration of cellular response. In a large screening campaign, longevity of the cell suspension should be determined from the concentration-response curves.
- Compounds dissolved in DMSO can be pre-dotted into assay plates in sub-microliter volume using an acoustic dispenser, such as Labcyte Echo or EDC Biosystem ATS, or pre-diluted in an assay buffer prior to addition. Tolerance of the cellular response to DMSO should be investigated during assay development. We found that most cells can tolerate up to 1% DMSO in the assay.

Step 5: Cover the assay plate and incubate for 30 minutes at room temperature (RT).

- The duration of assay incubation depends on the kinetic properties of the compounds. Potency values can be underestimated if the incubation is insufficient to achieve equilibrium.

Step 6: Prepare the cAMP standard curve according to the manufacturer's instructions.

Step 7: Dispense 5 µL/well of diluted d2-labeled cAMP conjugate followed by 5 µL/well of cryptate-labeled anti-cAMP antibody to the cell plate and cAMP standard curve plate.

- d2-labeled cAMP conjugate and cryptate-labeled anti-cAMP antibody are reconstituted per manufacturer's instructions.
- Working solutions are prepared by diluting the concentrated stock 20-fold using lysis buffer provided in the kit.

Step 8: The plates are incubated for about 1 hour at RT and read using an instrument such as the PerkinElmer Envision with a protocol that is set and optimized for HTRF detection.

- Fluorescence intensity is measured at Em665 and Em615, with excitation at 350 nm.
- Plates can be read repeatedly to ensure that the antibody binding reaches steady state.

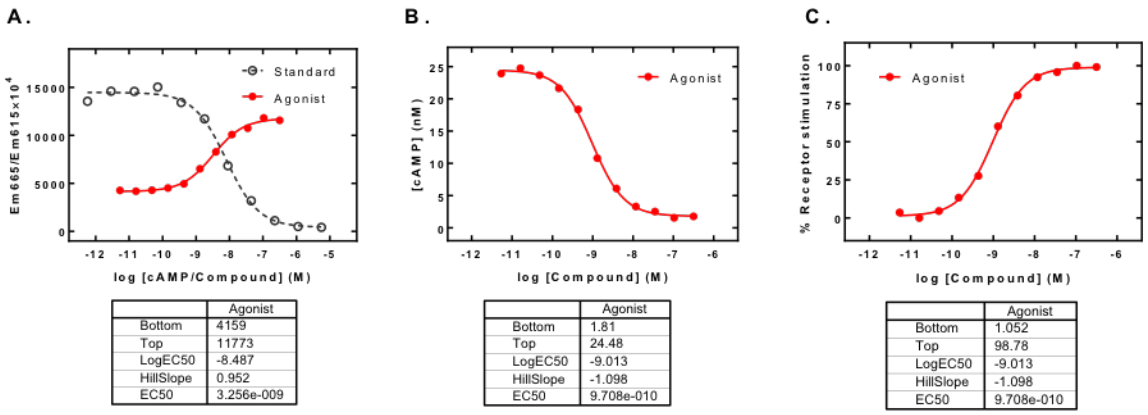
Step 9: Convert the results from the reader into cAMP levels using the cAMP standard curve.

- Readings are generally expressed as $Em665/Em615 \times 10^4$.

Step 10: Data analysis. See Figure 12.

$G_{\alpha i}$-Coupled Antagonist Format

In this format, an antagonist displaces an agonist for a $G_{\alpha i}$ coupled receptor and the resulting inhibition of adenylate cyclase is measured using the HTRF assays, as shown in Figure 13.



	Agonist
Bottom	4159
Top	11773
LogEC50	-8.487
HillSlope	0.952
EC50	3.256e-009

	Agonist
Bottom	1.81
Top	24.48
LogEC50	-9.013
HillSlope	-1.098
EC50	9.708e-010

	Agonist
Bottom	1.052
Top	98.78
LogEC50	-9.013
HillSlope	-1.098
EC50	9.708e-010

Figure 12. An example of agonist concentration-response curves in $G_{\alpha i}$-coupled receptors. Panel A. The agonist response is plotted as the HTRF ratio ($Em665/Em615\times10^4$) along with the cAMP standard curve. Note that the agonist produces signals that are within the linear portion of the standard curve. Panel B. The agonist response is plotted as the cAMP level to determine the IC_{50} values, respectively. Panel C. The same agonist response is plotted as % Inhibition. Data are analyzed using a four-parameter logistic regression.

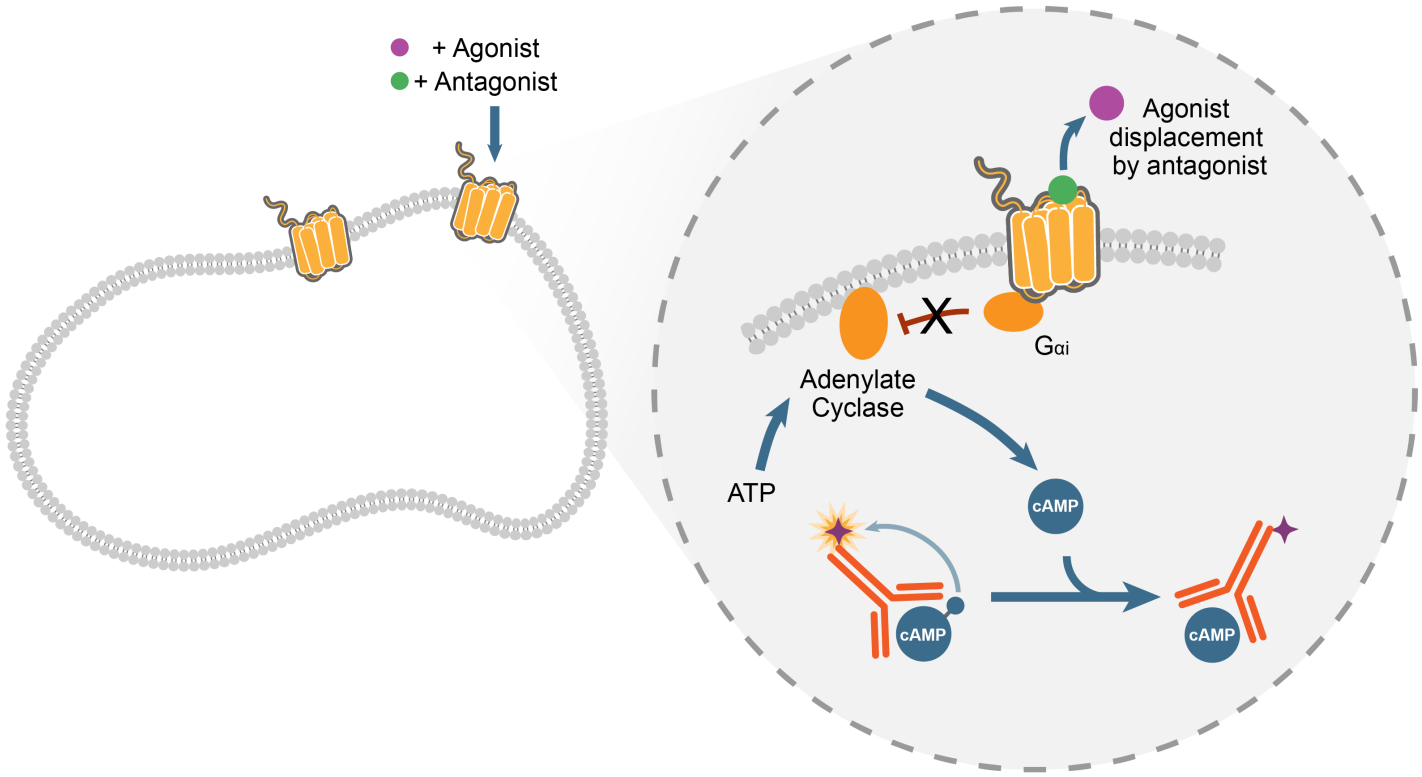


Figure 13. cAMP Measurement for Antagonists of a $G_{\alpha i}$-Coupled Receptor. Treatment of cells with forskolin results in an increase in cellular cAMP levels. The addition of an agonist decouples the G protein complex that subsequently dissociates into a $G_{\beta\gamma}$ dimer and a $G_{\alpha i}$ monomer. The activated $G_{\alpha i}$ subunit binds to adenylate cyclase resulting in the negative regulation of intracellular cAMP (no intracellular cAMP is generated). The presence an antagonist causes the release of the cAMP inhibition through competition between the antagonist and the agonist. Consequently, the level of cAMP level increases.

Sample Protocol for $G_{\alpha i}$-Coupled Antagonist Assays

The procedure for determination of antagonist activity of a $G_{\alpha i}$-coupled receptor using cAMP measurement is essentially the same as the procedure described above for detection of agonist activity. In the antagonist format, an agonist is added at concentrations that trigger 50-80% of the maximum response (EC_{50-80}) of the agonist. The rationale and precautions for each step can be referred to in the above agonist assay protocol, except for those specific to antagonist assays. All compounds of interest should also be tested in the agonist mode to ensure the absence of agonist activities.

Step 1: Grow cells in flask for 1-3 days in the corresponding cell culture medium.

Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes.

Step 3: Aspirate the medium and re-suspend the cells in assay buffer. Count and dilute the cells to the desired concentration in assay buffer. Dispense 5 µL of cells into assay plates that have been pre-dotted with compounds in DMSO or pre-diluted in an assay buffer. Incubate the plate at RT for 15-30 minutes.

- Depending on the kinetic properties of the test compounds relative to the agonist used for stimulation, this step can be adjusted from no incubation to an extended incubation time as needed.

Step 4: Prepare a solution containing the desired concentration of forskolin, IBMX (generally at a final concentration of 0.1 mM) and agonist, and dispense 5 µL of the mixture into the assay plates pre-dotted with compounds. *(Depending on the extent of activity of phosphodiesterase in the cells, the use of inhibitor is optional. The impact of phosphodiesterase inhibitor on compound potency and the required concentration should be determined according to the assays and cell lines being used.)*

- The final concentration of the agonist chosen should result in 50 to 80% of the maximum response of the agonist, as determined during development of the agonist assay.
- The working concentration of forskolin, IBMX and agonist solution mixture is prepared at 2X the desired final assay concentration.
- At the chosen EC_{50-80} concentration of agonist, the measured agonist response should not exceed the linear range of the cAMP standard curve.

Step 5: Cover the plate and incubate for 30 minutes at room temperature (RT).

Step 6: Prepare the standard curve according to the manufacturer's instructions.

Step 7: Dispense 5 µL diluted d2-labeled cAMP conjugate followed by 5 µL of cryptate-labeled anti-cAMP antibody into the cell plate and the standard curve plate.

Step 8: Incubate the plates for 1 hour at RT and measure the signal with a PerkinElmer Envision using a protocol that is optimized for HTRF detection.

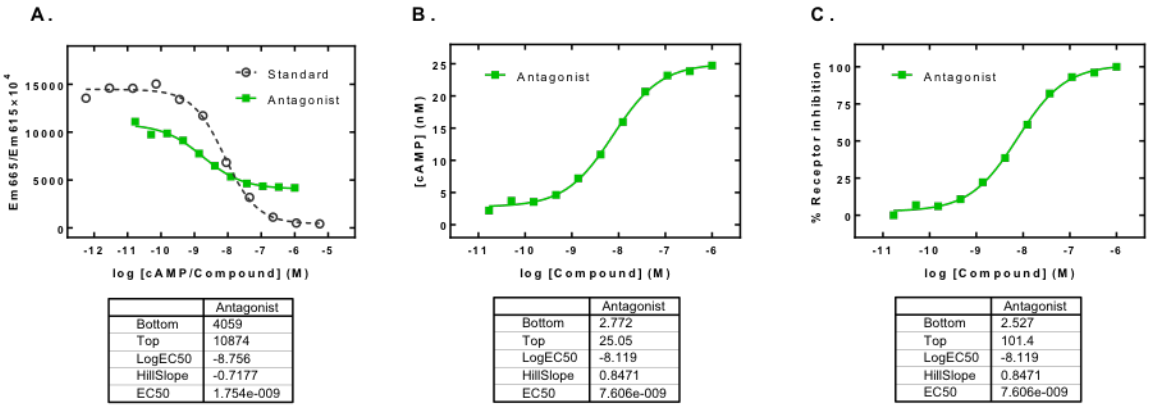
Step 9: Convert the results from the reader into cAMP levels using the cAMP standard curve.

- Readings are generally expressed as $Em665/Em615 \times 10^4$.

Step 10: Data analysis. See Figure 14.

Acknowledgements

The authors would like to thank Nathan Cheadle, Jie Pan and Melissa Yarde for the review of this guidance.



	Antagonist
Bottom	4059
Top	10874
LogEC50	-8.756
HillSlope	-0.7177
EC50	1.754e-009

	Antagonist
Bottom	2.772
Top	25.05
LogEC50	-8.119
HillSlope	0.8471
EC50	7.606e-009

	Antagonist
Bottom	2.527
Top	101.4
LogEC50	-8.119
HillSlope	0.8471
EC50	7.606e-009

Figure 14. Examples of agonist and antagonist concentration-response curves in $G_{\alpha i}$-coupled receptors. Panel A. The agonist and antagonist responses are plotted as the HTRF ratio (Em665/Em615×10^4) along with the cAMP standard curve. Note that both types of response produce signals that are within the linear portion of the standard curve. Panel B. The agonist and antagonist responses are plotted as the cAMP level to determine the EC_{50} and IC_{50} values, respectively. Data are analyzed using four-parameter logistic regression.

References

1. Cox KL, Devanarayan V, Kriauciunas A, Manetta J, Montrose C, Sittampalam S. Immunoassay Methods. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance Manual. Bethesda (MD)2004.
2. Available from: https://en.wikipedia.org/wiki/Stokes_shift.
3. Iversen PW, Beck B, Chen YF, Dere W, Devanarayan V, Eastwood BJ, et al. HTS Assay Validation. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance Manual. Bethesda (MD)2004.
4. Beck B, Chen YF, Dere W, Devanarayan V, Eastwood BJ, Farmen MW, et al. Assay Operations for SAR Support. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance Manual. Bethesda (MD)2004.
5. Hill SJ, Williams C, May LT. Insights into GPCR pharmacology from the measurement of changes in intracellular cyclic AMP; advantages and pitfalls of differing methodologies. British journal of pharmacology. 2010;161(6):1266–75. PubMed PMID: 21049583.
6. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. 1999;4(2):67–73. PubMed PMID: 10838414.
7. Burford NT, Watson J, Alt A. Standard Curves Are Necessary to Determine Pharmacological Properties for Ligands in Functional Assays Using Competition Binding Technologies. Assay and Drug Development Technologies. 2017;15(7):320–329. PubMed PMID: 29120673.

Measurement of β-Arrestin Recruitment for GPCR Targets

Tao Wang,[1] Zhuyin Li,[1] Mary Ellen Cvijic,[1] Carol Krause,[1] Litao Zhang,[1] and Chi Shing Sum[1]

Created: November 20, 2017.

Introduction

β-Arrestins are ubiquitously expressed in all cell types, and function in the desensitization of G- protein coupled receptors (GPCRs), the control of GPCR intracellular trafficking, and the activation of GPCRs to multiple signaling pathways (1-4). Therefore, β-arrestin-mediated signaling constitutes an important part of GPCR signaling in addition to G protein-mediated signaling. As many GPCRs are found to recruit β-arrestin, the β-arrestin recruitment assay has found important use in drug discovery, especially in the discovery of ligands for orphan GPCRs and in situations where the second messenger signaling is unknown (5,6). Furthermore, the discovery of biased GPCR ligands and the findings that distinct G-proteins versus β-arrestin signaling preferences may offer therapeutic advantages over conventional ligands imply that a screening campaign should be designed to focus on the most disease relevant pathways (7-10). In this aspect, the β-arrestin recruitment assay has added an important piece to the repertoire of assay tools in drug discovery.

There are four major *in vitro* assay technologies available on the market that are capable of measuring ligand-induced β-arrestin recruitment: PathHunter β-arrestin Assay (DiscoverX) (11), Tango GPCR Assay (Thermo Fisher Scientific) (12), LinkLight GPCR/ β-arrestin Signaling Pathway Assay (BioInvenu) (13), and Transfluor Assay (Molecular Devices) (14). The PathHunter β-arrestin Assay, Tango GPCR Assay System and LinkLight GPCR/ β-arrestin Signaling Pathway Assays are homogenous, high throughput assays while the Transfluor Assay is a fluorescence image-based assay. All four assays involve the expression of the β-arrestin as a fusion protein with another protein or fragment, while the PathHunter, Tango and LinkLight assays require fusion of the GPCR to another peptide or protein moiety as well. Table 1 compares the principles of these technologies as well as their advantages and limitations.

This guideline uses PathHunter β-arrestin from DiscoverX to illustrate the concepts for performing GPCR β-arrestin recruitment assay**.** The principle behind this guideline can be applied to all the β-arrestin assay technologies.

Table 1. Comparison of various β-arrestin assay technologies as well as their advantages and limitations.

Technologies	Vendor	Principles	Advantages	Limitations
PathHunter β-arrestin Assay	DiscoverX	A β-galactosidase enzyme fragment complementation technology-based assay. β-Arrestin recruitment brings about reconstitution of an active β-galactosidase, which is split into a PK fragment fused to the GPCR and an enzyme acceptor protein fused to the β-arrestin. The active enzyme catalyzes a chemiluminescence reaction from an exogenous substrate to generate a chemiluminescence readout.	• Signal is specific to the GPCR of interest tagged to the PK fragment off-target signal is minimized. • High signal-to-background from chemiluminescence readout.	• Requires construction of fusion proteins for both the GPCR and β-arrestin proprietary detection reagents.

Author Affiliation: 1 Bristol-Myers Squibb Company, NJ.

Table 1. continued from previous page.

Technologies	Vendor	Principles	Advantages	Limitations
Tango GPCR Assay	Thermo Fisher Scientific	A reporter gene assay that is based on the use of a non-native transcription factor linked to the GPCR via a protease cleavage site. β-Arrestin recruitment causes the irreversible cleavage and release of the transcription factor mediated by a protease-tagged β-arrestin. The freed transcription factor promotes the expression of a β-lactamase reporter, which in turns catalyzes the turnover of an exogenous substrate to generate a fluorescent signal.	• Signal is specific to the GPCR of interest tagged to the transcription factor and off-target signal is minimized. • Two color ratiometric FRET readout can reduce experimental noise.	• Requires the construction of fusion proteins for both the GPCR and β-arrestin, and a cell line harboring the reporter gene. • Longer assay duration due to the need for expression of the reporter gene.
LinkLight GPCR/ β-arrestin Signaling	BioInvenu	β-Arrestin is fused to a permutated luciferase. Recruitment causes the irreversible cleavage of the luciferase mediated by a non-native protease fused to the GPCR. Subsequent refolding of the luciferase into an active enzyme can be detected by an exogenous substrate to generate a chemiluminescence readout.	• Signal is specific to the GPCR of interest tagged to the transcription factor and off-target signal is minimized. • Compatible with any existing luciferase substrates.	• Requires construction of fusion proteins for both the GPCR and β-arrestin.
Transfluor Assay	Molecular Devices	Recruitment of a GFP-tagged β-arrestin and the concomitant intracellular translocation is visualized and quantified by bioimaging method.	• Amenable to studies of the intracellular spatial-temporal relationship of GPCR activation and β-arrestin recruitment.	• Requires the use of GFP fusion protein of β-arrestin.

Flowchart of Assay Development Guidelines

1. Select a GPCR target cell line and a control cell line expression another GPCR, preferably with similar total and background.
   - The control cell line allows for detection of off-target activity.
2. Determine agonist dose-response relationships and potency values while optimizing cell density and determining the time course and desired temperature of β-arrestin recruitment
   - The three parameters (cell density, temporal profile and temperature) for β-arrestin recruitment can be optimized simultaneously.
   - The measured chemiluminescent signals are used to determine the EC_{50} or EC_{80} values.
   - Since the assay is an end-point assay, multiple reactions need to be set up for all combinations.
3. Determine antagonist dose-response relationships and IC_{50} values by stimulating with EC_{50-80} agonist concentration
   - The measured chemiluminescent signals are used to determine the IC_{50} value.
4. Assess assay performance statistics, including Z', signal-to-background ratio, reproducibility of potency data
   - Ensure Z' > 0.5 and that reproducibility over multiple runs are within acceptable limits

Overview of the β-Arrestin Recruitment Assay

The PathHunter GPCR β-arrestin assay is based on proprietary Enzyme Fragment Complementation technology from DiscoverX. PathHunter β-arrestin GPCR cells are engineered to co-express the ProLink (PK) tagged GPCR

and the Enzyme Acceptor (EA) tagged β-arrestin. Activation of the GPCR-PK induces β-arrestin-EA recruitment, facilitating complementation of the two β-galactosidase enzyme fragments (EA and PK). The resulting functional enzyme hydrolyzes the Galacton Star substrate to generate a chemiluminescent signal (Figure 1).

PathHunter β-arrestin GPCR cells can be purchased from DiscoverX (https://www.discoverx.com) directly as stable expressing cell lines, or generated using PathHunter parental cells engineered to express β-arrestin1 or β-arrestin2-EA fusion protein. Also, retroviral particles are available to generate the parental cells in various cellular backgrounds. Various types of ProLink vectors can be chosen to create the GPCR construct carrying the ProLink (PK) tagged GPCR. Specific detection reagents can be purchased from DiscoverX.

The use of positive and negative controls is important to define the dynamic range of the assays in each plate and monitor the overall assay quality. Inclusion of one or multiple reference compounds for concentration-response analysis is valuable for assessing the incidence of assay drift and robustness. More details can be found in the AGM chapters HTS Assay Validation (15) and Assay Operations for SAR Support (16). As in other assays, performing a counter-screen assay with an unrelated GPCR in the same technology readout is crucial to ensure on-target interaction. The signal may be subject to interference from compounds that inhibit the β-galactosidase enzymatic activity, or those that inhibit the GPCR β-arrestin complex formation.

Guidelines and Sample Protocol for Agonist Assays

Step 1: Grow cells in tissue culture flasks for 1-3 days in the corresponding cell culture medium.

- Select a DiscoverX PathHunter cell line expressing the tagged GPCR target of interest.
- A negative control cell line as a counter screen may be either the DiscoverX β-arrestin parental cell line devoid of the target or a DiscoverX β-arrestin cell line expressing an unrelated GPCR.
- Make sure the cells are healthy and active in the log phase of growth and are maintained the same way in all experiments, since the growth state of the cells may affect the receptor response.
- Cell passage number should be tightly controlled based on assay performance statistics. Ideally, to avoid variation in cellular response due to changes in the state of the cells or level of receptor expression, expand the cell culture to the quantities required for the entire screening campaign in one large batch of a single passage, "assay-ready" cryopreserved cells in appropriate aliquot sizes (17). DiscoverX has a suggested cryopreservation method and cryopreservation reagent for its cell lines outlined in the instruction manual. Other cryopreservation methods and reagents are available from various manufacturers of cell culture reagents. These can be tested as well for suitability to identify a method that provides high viability and receptor functionality. Frozen cell aliquots can be stored for some time and recovered as needed during a screen.

Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes to pellet the cells.

- Use of Trypsin or other enzymatic/non-enzymatic cell dissociation agents should be tested prior to screening. Some receptor proteins may tolerate trypsinization, but other enzymatic reagents such as TrypLE (Thermo Fisher) or non-enzymatic reagents such as Cellstripper (Corning) may be used if the system under study does not tolerate treatment with proteases.

Step 3: Aspirate the cell culture medium and re-suspend the cells in cell plating medium. Count and dilute the cells to the proper densities in cell plating medium.

- Cells should be dissociated well enough so that no cell clumps are observed during cell counting. For troublesome cell lines, optimize the dissociation condition and filter the cells through an appropriate strainer (e.g. BD Cell Strainer).

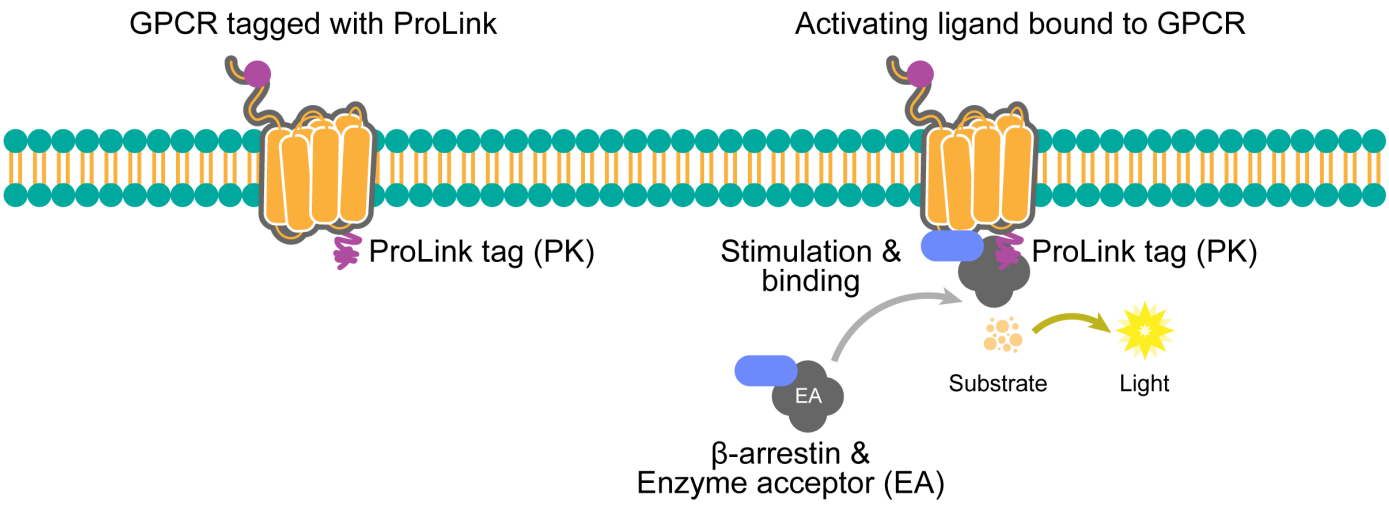


Figure 1. The GPCR β-arrestin recruitment assay principle.

- In most cases, basal medium supplemented with 0.5 to 10% FBS can be used. Sensitivity of the cells and compound activity to serum should be tested.
- Cell densities should be determined during assay development in cell titration experiments. Generally, a range of 5,000–20,000 cells/well will be suitable for 384-well assay.

Step 4: Dispense cells in plating medium into 384-well microplate and incubate the cells overnight at 37°C at 95% O_2/5% CO_2.

- Cells and other reagents can be dispensed by various peristaltic pump-based liquid handlers, such as Thermo Combi-drop or the BioTek Washer Dispenser, or any of those widely available tip-based liquid handlers.
- Tissue culture-treated or poly-D-lysine plates may be used depending on the cell type.
- The PathHunter β-arrestin assay may be run as a suspension assay. In that case, the cells diluted in assay buffer can be seeded directly into assay plates and used immediately for assays.

Step 5: Remove plating medium and replace with serum-free or low serum assay medium if needed.

- At this step, cells can be serum-starved for 4-16 hours depending on application. Serum starvation helps minimize basal signaling, which is especially important if agonists for the GPCR target are present in the serum.

Step 6: Dispense test compounds into the plates.

- Compounds dissolved in DMSO can be pre-dotted into assay plates in sub-microliter volume using an acoustic dispenser such as the Labcyte Echo, or pre-diluted in an assay buffer prior to addition.
- Tolerance of the cellular response to DMSO should be tested. Maintaining DMSO concentration under 1% is preferred. An example of DMSO tolerance test is shown in Figure 2.

Step 7: Incubate the cell plates at 37°C for 90 minutes (or other optimized condition).

- GPCR-β-arrestin interaction is known to exhibit two distinguishable characteristics, known as Class A and Class B interactions. Class A interaction refers to the formation of transient GPCR-β-arrestin complexes which dissociate during receptor trafficking. Class B interaction is more sustained, and the complex co-localizes to subcellular compartments during receptor translocation (18).

- Incubation time of the assay should be optimized for each GPCR. For some GPCRs, the assay can also be performed at room temperature.
- With the PathHunter β-arrestin assay, the PK tag on the engineered GPCR may have some degree of intrinsic affinity for the EA fragment. This may influence the GPCR-β-arrestin interaction. A time-course experiment may be performed to assess how this may affect Class A and Class B profiles.

See figure 3 for determination of compound incubation time.

Step 8: Prepare a solution of PathHunter Detection Reagent (1 part Galacton Star, 5 parts Emerald II Solution and 19 parts PathHunter Cell Assay Buffer according to manufacturer's instruction).

- If cells are incubated with compounds at 20 µl of volume, then 10 µl of detection reagent is needed for each well.
- The detection reagent can be added directly to the compound-treated cells using any of the peristaltic pump-based liquid handlers (e.g., Combi-drop from Thermo Fisher or BioTek Washer Dispenser).
- If compound interference is suspected, a wash step using DPBS with Ca^{2+}/Mg^{2+} can be carried out prior to the addition of detection reagent.

Step 9: Incubate for 60 minutes at room temperature and measure chemiluminescent signal using plate readers such as the PerkinElmer EnVision. The output signal is relative light units (RLU). The absolute RLU scale may vary depending on the detector.

The final cell density, temperature and the selection of optimal time point is based on the Z' and the expected potency (15).

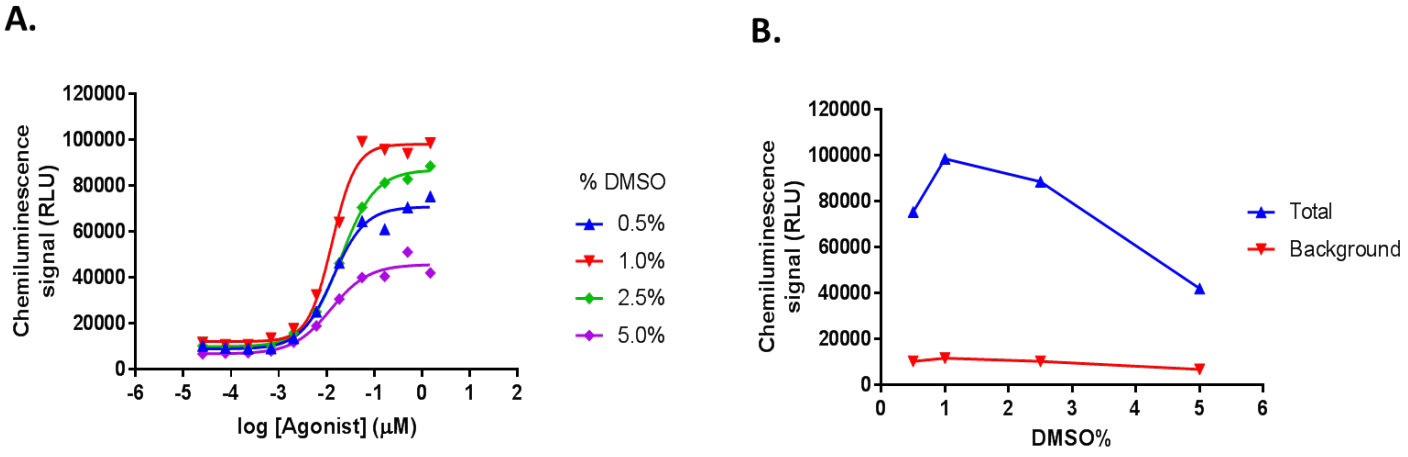
Guidelines and Sample Protocol for Antagonist Assays

Prior to setting up an antagonist-characterization assay, the agonist to be used for stimulating the cells during the antagonist mode should be tested to determine the potency. The $EC_{50\text{-}80}$ of the agonist (determined from a concentration response curve) can be selected for use in the antagonist assay. The assay procedure is similar to the agonist mode with an additional step for agonist stimulation ($EC_{50\text{-}80}$ value) after incubation of cells with test compounds. All test compounds should also be assessed in the agonist mode to examine any agonist activities.

Step 1: Grow cells in tissue culture flasks for 1-3 days in the corresponding cell culture medium.

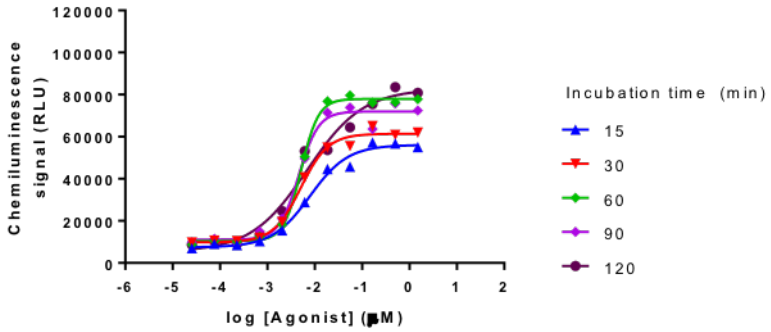
- Select a DiscoverX PathHunter cell line expressing the tagged GPCR target of interest.
- A negative control cell line as a counter screen may be either the DiscoverX β-arrestin parental cell line devoid of the target or a DiscoverX β-arrestin cell line expressing an unrelated GPCR.
- Make sure the cells are healthy and active in the log phase of growth and are maintained in the same way in all experiments, as the growth state of the cells may affect the receptor response.
- Cell passage number should be tightly controlled based on assay performance statistics. Ideally, to avoid variation in cellular response due to changes in the state of the cells or level of receptor expression, expand the cell culture to the quantities required for the entire screening campaign in one large batch of a single passage, "assay-ready" cryopreserved cells in appropriate aliquot sizes (17). DiscoverX has a suggested cryopreservation method and cryopreservation reagent for its cell line that is available in the instruction manual. Other cryopreservation methods and reagents are available from various manufacturers of cell culture reagents. These can be tested as well for suitability to identify a method that provides high viability and receptor functionality. Frozen cell aliquots can be stored for some time and recovered as needed during a screen.

Step 2: Dissociate the cells from the flask and centrifuge at 1000 rpm for 5 minutes to pellet the cells.



	0.5%	1.0%	2.5%	5.0%
Bottom	8887	12130	9881	6742
Top	70952	98197	87026	45732
LogEC50	-1.869	-1.884	-1.702	-1.908
HillSlope	1.282	1.734	1.201	1.098
EC50	0.01353	0.01305	0.01985	0.01237

Figure 2. DMSO tolerance test. (A) The effect of DMSO on agonist-stimulated response is shown in relationship to the agonist concentration-dependent responses. In this example, DMSO concentrations at 5% or higher reduce the maximal responses by over 50% without affecting the EC50 values. (B) The effect of DMSO is examined in terms of the effect on total and background signals. The total signal is seen to decline dramatically beyond 2.5% DMSO. Similar analysis can be performed to determine DMSO effects on Z'.



	15	30	60	90	120
Bottom	7442	9812	10427	11077	5339
Top	56104	61391	77964	71995	82733
LogEC50	-2.102	-2.298	-2.288	-2.338	-2.162
HillSlope	1.085	1.496	2.357	2.086	0.7347
EC50	0.007916	0.005038	0.005149	0.004589	0.006881

Figure 3. Determination of compound incubation time in an agonist assay. A known agonist is used to assess the time it takes to achieve steady-state cellular response by incubating the agonist for increasing periods of time. In this example, the maximal response elicited by the agonist is slightly higher upon longer incubation; however, the potency remains largely the same. If available, agonists with different potencies may be used to assess the kinetic behavior of various compounds in order to determine the best timing that yields the expected potency.

- Use of Trypsin or other enzymatic/non-enzymatic cell dissociation agents should be tested prior to screening.
- Some receptor proteins may tolerate trypsinization, but other enzymatic reagents such as TrypLE (Thermo Fisher) or non-enzymatic reagent such as Cellstripper (Corning) may be used if needed.

Step 3: Aspirate the cell culture medium and re-suspend the cells in cell plating medium. Count and dilute the cells to the proper densities in cell plating medium.

- Cells should be dissociated well enough so that no cell clumps are observed during cell counting. For troublesome cell lines, optimize the dissociation and filter the cells through an appropriate strainer (e.g. BD Falcon Cell Strainer).
- In most cases, basal medium supplemented with 0.5 to 10% FBS can be used. Sensitivity of the cells and compound activity to serum should be tested.
- Cell densities should be determined during assay development in cell titration experiments. Generally, a range of 5,000–20,000 cells/well will be suitable for 384-well assay.

Step 4: Dispense cells in plating medium into 384-well microplate and incubate the cells overnight at 37°C at 95% O_2/5% CO_2.

- Tissue culture-treated or poly-D-lysine plates may be used depending on the cell type.
- The PathHunter β-arrestin assay may be run as a suspension assay. In that case, the cells diluted in assay buffer can be seeded directly into assay plates and used immediately for assays.

Step 5: Remove plating medium and replace with serum-free or low serum assay medium if needed.

- At this step, cells can be serum-starved for 4-16 hours depending on application. Serum starvation helps minimize basal signaling, which is especially important if agonists for the GPCR target are present in the serum.

Step 6: Dispense test compounds into the plates.

- Tolerance of the cellular response to DMSO should be tested. Maintaining DMSO concentration under 1% is preferred.

Step 7: Incubate the cell plates at 37°C for 15-30 minutes (or other optimized condition), then dispense the proper amount of the $EC_{50\text{-}80}$ concentration of agonist.

Step 8: Incubate the cell plate at 37°C (or other optimized condition) for the desired length of time according to the time-course experiment.

- GPCR-β-arrestin interaction is known to exhibit two distinguishable characteristics. Class A interaction is transient, and the GPCR- β-arrestin complex may dissociate during receptor trafficking. Class B interaction is more sustained, and the complex co-localizes during receptor translocation (18).
- Appropriate incubation time of the assay can be optimized for each GPCR. For some GPCRs, the assay can also be performed at room temperature.
- With the PathHunter β-arrestin assay, the PK tag on the engineered GPCR may have some degree of intrinsic affinity for the EA fragment. This may influence the GPCR-β-arrestin interaction. Tests may be performed to assess how this may affect the Class A and Class B profiles.

Step 9: Prepare a solution of PathHunter Detection Reagent (1 part Galacton Star, 5 parts Emerald II Solution and 19 parts PathHunter Cell Assay Buffer according to manufacturer's specification).

- The detection reagent can be added directly to the compound-treated cells using any of the peristaltic pump-based liquid handlers (e.g., Combidrop from Thermo Fisher or BioTek Washer Dispenser).
- If compound interference is suspected, a wash step using DPBS with Ca^{2+}/Mg^{2+} can be carried out prior to addition of detection reagent.

Step 10: Incubate for 60 minutes at RT and measure chemiluminescent signal using plate readers such as EnVision.

The final cell density, temperature and the selection of optimal time point is based on the Z' and the expected potency (15).

See figure 4 for examples of agonist and antagonist concentration-response curves in GPCR-β-arrestin recruitment assays.

An overall guideline of the assay workflow is shown in the following table.

Assay Step	Typical Range	Instrumentation
Grow cells		CO_2 incubator
Harvest cells		Centrifuge
Plate cells	5,000 – 20,000 cells per well (384w)	Combi-drop (Thermo Fisher), BioTek Washer Dispenser, or tip-based liquid handler
Dispense compounds	0.1 nM – 10 μM	Echo, ATS Gen5, or tip-based liquid handler
Incubate cells	Room temperature or 37°C	CO_2 incubator
Dispense agonist (for antagonist assay)	EC_{50} –EC_{80}	Combi-drop (Thermo Fisher), BioTek Washer Dispenser, or tip-based liquid handler
Wash cells with PBS (if compound interference is suspected)		BioTek Washer, BlueWasher (BlueCatBio), or manually inverted the plate to remove the liquid
Stop reaction by addition of lysis and detection reagents		Combi-drop (Thermo Fisher), BioTek Washer Dispenser, or tip-based liquid handler
Incubate detection reagents	Room temperature	
Measure chemiluminescence		Envision, or other luminometer

Troubleshooting

Key assay optimization parameters and troubleshooting guidelines are available in Table 2.

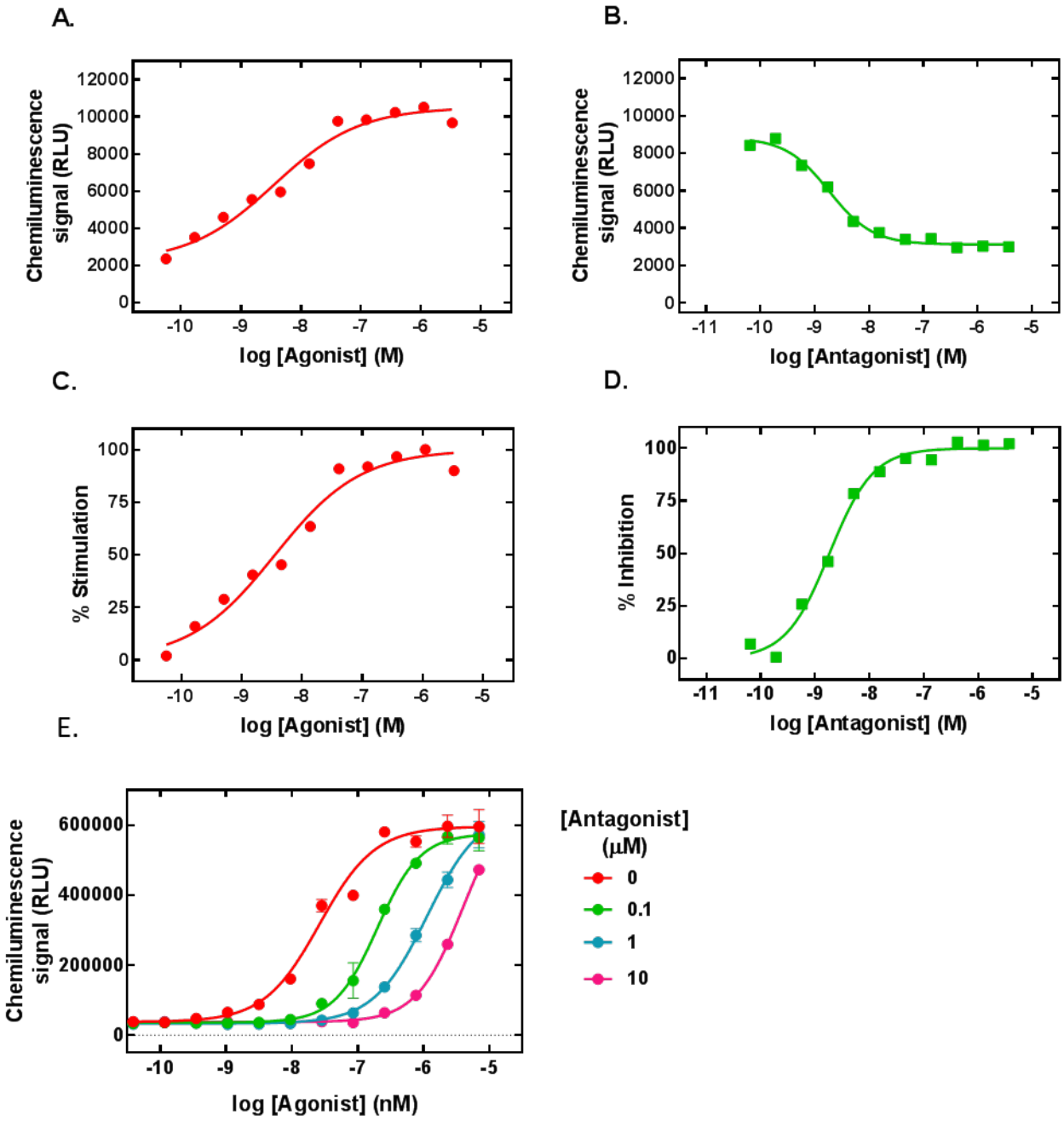


Figure 4. Examples of agonist and antagonist concentration-response curves in GPCR-β-arrestin recruitment assays in agonist mode (A, C), antagonist mode (B, D) and Schild dose-shift assay (E). Raw data from chemiluminescence readout are illustrated in Panels A and B, with the normalized data given in Panels C and D. % Stimulation = [(RLU signal – Min RLU)/(Max RLU – Min RLU)] x 100. % Inhibition = [1-(RLU signal – Min RLU)/(Max RLU – Min RLU)] x 100. Panel E. Schild dose-shift assay can be performed to assess the surmountability of antagonism and to determine the affinity of the antagonist (19), as implemented in the GraphPad Prism software. Data are analyzed using four-parameter logistic regression.

Table 2. Key assay optimization parameters and troubleshooting guidelines

Problem	Possible reason	Solution
Background signal higher than expected	Presence of endogenous agonist(s) in serum	Perform serum starvation for different periods of time.
	Constitutive receptor signaling	Examine multiple clones of different receptor expression levels.

Table 2. continued from previous page.

Problem	Possible reason	Solution
Agonist-stimulated response low or undetectable	Receptor intrinsically not recruiting β-arrestin or the subtype of β-arrestin being used	Assess β-arrestin recruitment with various subtypes by other methods such as co-immunoprecipitation if feasible.
	Cells not expressing sufficient quantities of receptor or β-arrestin	Examine multiple clones of different receptor and β-arrestin expression levels.
	Compound is not efficacious	Test other compounds known to be efficacious in triggering β-arrestin recruitment.
	Insufficient compound binding to receptors	Perform time-course and temperature experiments for optimal activity.
Agonist/antagonist potency not as expected	Insufficient compound binding to receptors	Perform time-course and temperature experiments for optimal activity.
	Dynamics of receptor-β-arrestin recruitment may be transient	Perform time-course and temperature experiments for optimal β-arrestin recruitment.
	Intrinsic signaling bias	Confirm signaling bias with an orthologous β-arrestin recruitment assay.
	Influence of β-galactosidase activity in detection step	Perform orthologous β-arrestin recruitment assay or counter screen assay with another cell line expressing a GPCR known to recruit β-arrestin; incorporate a wash step to remove compounds from cells prior to detection.

Acknowledgements

The authors would like to thank Nathan Cheadle, Jie Pan and Melissa Yarde for the review of this guidance.

References

1. Gurevich EV, Gurevich VV. Arrestins: ubiquitous regulators of cellular signaling pathways. Genome Biol. 2006;7(9):236doi: 10.1186/gb-2006-7-9-236. PubMed PMID: 17020596.
2. Lefkowitz RJ, Shenoy SK. Transduction of receptor signals by beta-arrestins. Science. 2005;308(5721):512–7.doi: 10.1126/science.1109237. PubMed PMID: 15845844.
3. Moore CA, Milano SK, Benovic JL. Regulation of receptor trafficking by GRKs and arrestins. Annu Rev Physiol. 2007;69:451–82.doi: 10.1146/annurev.physiol.69.022405.154712. PubMed PMID: 17037978.
4. Tian X, Kang DS, Benovic JL. beta-arrestins and G protein-coupled receptor trafficking. Handb Exp Pharmacol. 2014;219:173–86.doi: 10.1007/978-3-642-41199-1_9. PubMed PMID: 24292830.
5. Kroeze WK, Sassano MF, Huang XP, Lansu K, McCorvy JD, Giguere PM, et al. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. Nat Struct Mol Biol. 2015;22(5):362–9.doi: 10.1038/nsmb.3014. PubMed PMID: 25895059.
6. Southern C, Cook JM, Neetoo-Isseljee Z, Taylor DL, Kettleborough CA, Merritt A, et al. Screening beta-arrestin recruitment for the identification of natural ligands for orphan G-protein-coupled receptors. J Biomol Screen. 2013;18(5):599–609.doi: 10.1177/1087057113475480. PubMed PMID: 23396314.
7. Rominger DH, Cowan CL, Gowen-MacDonald W, Violin JD. Biased ligands: pathway validation for novel GPCR therapeutics. Curr Opin Pharmacol. 2014;16:108–15.doi: 10.1016/j.coph.2014.04.002. PubMed PMID: 24834870.
8. Violin JD, Crombie AL, Soergel DG, Lark MW. Biased ligands at G-protein-coupled receptors: promise and progress. Trends in pharmacological sciences. 2014;35(7):308–16.doi: 10.1016/j.tips.2014.04.007. PubMed PMID: 24878326.

9. Kenakin T, Christopoulos A. Signalling bias in new drug discovery: detection, quantification and therapeutic impact. Nat Rev Drug Discov. 2013;12(3):205–16.doi: 10.1038/nrd3954. PubMed PMID: 23411724.
10. Luttrell LM, Maudsley S, Bohn LM. Fulfilling the Promise of "Biased" G Protein-Coupled Receptor Agonism. Molecular pharmacology. 2015;88(3):579–88.doi: 10.1124/mol.115.099630. PubMed PMID: 26134495.
11. Zhao X, Jones A, Olson KR, Peng K, Wehrman T, Park A, et al. A homogeneous enzyme fragment complementation-based beta-arrestin translocation assay for high-throughput screening of G-protein-coupled receptors. J Biomol Screen. 2008;13(8):737–47.doi: 10.1177/1087057108321531. PubMed PMID: 18660457.
12. Barnea G, Strapps W, Herrada G, Berman Y, Ong J, Kloss B, et al. The genetic design of signaling cascades to record receptor activation. Proc Natl Acad Sci U S A. 2008;105(1):64–9.doi: 10.1073/pnas.0710487105. PubMed PMID: 18165312.
13. Eishingdrelo H, Cai J, Weissensee P, Sharma P, Tocci MJ, Wright PS. A cell-based protein-protein interaction method using a permuted luciferase reporter. Curr Chem Genomics. 2011;5:122–8.doi: 10.2174/1875397301105010122. PubMed PMID: 22207892.
14. Oakley RH, Hudson CC, Cruickshank RD, Meyers DM, Payne RE Jr, Rhem SM, et al. The cellular distribution of fluorescently labeled arrestins provides a robust, sensitive, and universal assay for screening G protein-coupled receptors. Assay Drug Dev Technol. 2002;1(1 Pt 1):21–30.doi: 10.1089/154065802761001275. PubMed PMID: 15090153.
15. Iversen PW, Beck B, Chen YF, Dere W, Devanarayan V, Eastwood BJ, et al. HTS Assay Validation. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance Manual. Bethesda (MD)2004.
16. Beck B, Chen YF, Dere W, Devanarayan V, Eastwood BJ, Farmen MW, et al. Assay Operations for SAR Support. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance Manual. Bethesda (MD)2004.
17. Zaman GJ, de Roos JA, Blomenrohr M, van Koppen CJ, Oosterom J. Cryopreserved cells facilitate cell-based drug discovery. Drug Discov Today. 2007;12(13-14):521–6.doi: 10.1016/j.drudis.2007.05.008. PubMed PMID: 17631245.
18. Pierce KL, Lefkowitz RJ. Classical and new roles of beta-arrestins in the regulation of G-protein-coupled receptors. Nat Rev Neurosci. 2001;2(10):727–33.doi: 10.1038/35094577. PubMed PMID: 11584310.
19. Colquhoun D. Why the Schild method is better than Schild realised. Trends in pharmacological sciences. 2007;28(12):608–14.doi: 10.1016/j.tips.2007.09.011. PubMed PMID: 18023486.

Genome Editing Using Engineered Nucleases and Their Use in Genomic Screening

Joana R. Costa, Ph.D.,[1] Bruce E. Bejcek, Ph.D.,[2] James E. McGee, Ph.D.,[3] Adam I. Fogel, Ph.D.,[4] Kyle R. Brimacombe, M.S.,[5] and Robin Ketteler, Ph.D.[1]

Created: November 20, 2017.

Abstract

Understanding gene function is critical for developing therapeutic strategies to target disease. Common approaches to understanding gene function in a systematic and unbiased way include loss-of-function and gain-of-function genomic screening. Some of these rely on artificially increasing the copy number of gene transcripts using cDNA expression libraries. Others interrogate endogenous protein expression through genetic loss-of-function approaches such as siRNA screening. Over the past two decades, targeted approaches that reduce the endogenous expression level of genes or proteins have been developed that facilitate a much better understanding of genes in the context of the living cell. Technologies have been developed that enable precise modification of the genome rather than reduction at the transcript level. These include zinc finger nucleases, TALENs and the CRISPR/Cas9 technology. The use of these systems in large-scale high-throughput screening is an emerging field and we herein highlight recommendations for such applications.

Introduction

The ability to modify the expression of single genes and proteins has become one of the most important tools in molecular and cellular biology. Several methodologies have been developed to allow for specific gene manipulation in tissue culture cells, which have become colloquially known as "genome-editing". These rely on nucleases that are engineered to cut specific genomic target sequences, including *Zinc Finger Nucleases* (ZFN), *Transcription Activator-like Effector Nucleases* (TALEN) and *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR) nucleases (1). Homing meganucleases have also been used for these purposes but because they have not achieved widespread use they will not be discussed further.

The ZFNs, TALENs, and CRISPR/Cas9 (CRISPR associated) enzymes create double stranded breaks in the target DNA sequence which the cell will then repair using one of two pathways (Figure 1, Figure 2). The first of these processes is non-homologous end joining (NHEJ), which occurs without the use of a repair template. NHEJ results in a deletion or insertion (indel), the resulting sequence of which is essentially random and impossible to control. Practically, this pathway is useful in generating loss-of-function (knockout) of the gene of interest. The second pathway is homology-directed repair (HDR), which utilizes a repair template that contains homology to the sequences proximal to the DNA break (the endogenous template in cells is the sister chromatid). This pathway allows for precise control of the resulting insertion, but requires additional DNA sequences to be transfected, and is often less efficient than the NHEJ pathway.

In addition to these methods, *recombinant Adeno-Associated Viral vectors* (rAAVs) have been used to modify specific genomic DNA targets as well (2) (reviewed in (3) and (4)). The rAAV methodology for genomic modification does not rely on the activity of an exogenously derived nuclease but relies on the replacement of

Author Affiliations: 1 MRC Laboratory for Molecular Cell Biology, University College London; Email: j.costa@ucl.ac.uk; Email: r.ketteler@ucl.ac.uk. 2 Western Michigan University; Email: Bruce.bejcek@wmich.edu. 3 Eli Lilly; Email: Mcgee_james_e@lilly.com. 4 Biogen Inc.; Email: adam.fogel@biogen.com. 5 NIH/NCATS; Email: kyle.brimacombe@nih.gov.

Corresponding author.

sequences carried on a rAAV for the endogenous cellular homologue. rAAV methods will not be discussed further in this chapter.

All of these methods can be used in a wide variety of cell types to modify specific DNA sequences. In this chapter, we discuss the general application of different genome editing techniques for cell line generation and the specific use of CRISPR as a scalable platform for genetic screening.

A. Zinc Finger Nucleases (ZFNs)

Typically, genome editing platforms require two key features: an endonuclease that can cut a target location in the genome; and a sequence-specific adaptor that targets the endonuclease to a specified region on the DNA.

Zinc finger nucleases are chimeric proteins comprised of a DNA binding domain composed of zinc fingers (based on zinc finger transcription factors) and an endonuclease (5) (Figure 1, reviewed in (6)). Zinc fingers are ~30 amino acids that can bind to a limited combination of ~3 nucleotides. By using a combination of different zinc fingers, a unique DNA sequence within the genome can be targeted. Similarly, it was discovered that the *Fok I* type IIS restriction endonuclease had distinct and separable DNA binding and DNA cleavage domains, and the cleavage domain only has activity when there is dimerization (7). Making a chimera of the zinc finger binding domain with the *Fok I* digestion domain resulted in an artificial nuclease with a specificity that can theoretically be tailored to any sequence. When used in whole live cells it has become a powerful tool for the creation of cells that have single, insertion/deletion (indel) mutations in a region of interest in the genome.

The concept of use for zinc finger nucleases is relatively straightforward. Two chimeras are produced. Each chimera recognizes a specific and unique DNA sequence where the two sequences are relatively close to each other. When these chimeras are bound to the DNA they cut the DNA to create a fragment that is released, resulting in the activation of the cellular repair machinery in the form of NHEJ. NHEJ results in indels of various sizes and thus can disrupt open reading frames and result in a nonfunctional protein or a lack of an exon. Alternatively, a fragment of DNA can be also supplied to the cell, resulting in replacement of the sequence between the two cut sites by homologous recombination. The replacement fragment can be a normal or mutant form of the gene that is to be replaced.

There are some drawbacks with this technique:

Cost and time to engineer. ZFNs were the original platform used for genome editing and date back to the 1980s (8). Their use has always been hampered by the need to computationally predict and then iteratively engineer and test for sequence specificity. In comparison with the modular nature of TALE repeat binding, or the direct Watson-Crick affinity pairing of the CRISPR/Cas system, this makes the ZFNs much slower to develop and much costlier to pursue for a given target. This is the major reason why this technique has fallen out of favor in recent years.

Targeting specificity. Although the sequence that is targeted for digestion may be unique within the genome and the ZFN may have a high affinity for this sequence, off-target binding and cleavage of other DNA sequences can occur and result in undesired genome modification at other sites. For this interaction to result in digestion of DNA that will be repaired by NHEJ both halves of the ZFN must recognize sequences around the target site. Although this probability is small there are known zinc fingers that have off-site digestion patterns that will result in disruption of more than one site. Such off-target activities can be determined by sequencing the entire genome, but this is time consuming and costly. The best way to determine whether off-target activities might be influencing the biology of interest is to create more than one ZFN to digest the DNA at different sites (and presumably different off-target sites). When assessing knockout function, it is also imperative that the phenotype is rescued with a reintroduction of a functional gene in order to demonstrate the phenotype achieved is due to disruption of the gene of interest.

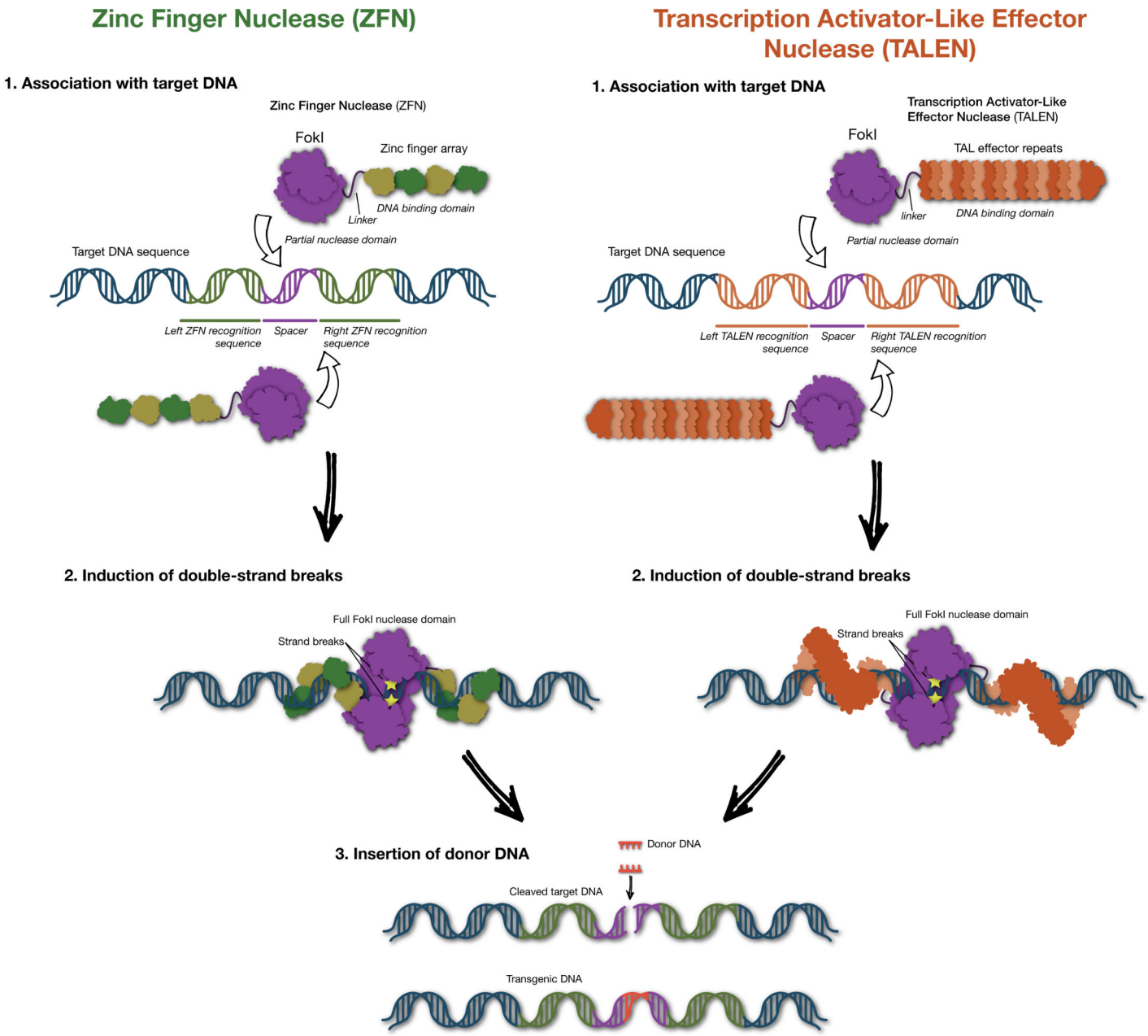


Figure 1: Overview of genome editing by zinc finger nucleases (ZFNs) and Transcription Activator-Like Effector Nucleases (TALENs). Zinc finger nucleases are composed of the Fok I endonuclease and an array of zinc finger binding domains that recognize the target DNA sequence. The action of two ZFNs on both strands of the DNA will result in a double strand break. A donor DNA provided in trans can then be integrated at the site of the break, resulting in a transgenic DNA sequence. In the absence of a donor template, the double strand break will be repaired by the host machinery, often resulting in insertions/deletions (indels) that disrupt the open reading frame. For TALENs, the zinc finger array is replaced by TAL effector repeats that guide targeting to the DNA.

B. TALENs

TALE-Nucleases (TALENs) are based on the transcription activator-like effector (TALE) proteins from *Xanthomonas* bacterial species (reviewed in (9,10)) that have a DNA binding domain and an effector domain (Figure 1). The DNA binding domain consists of several subdomains of slightly variable length (~34 amino acids) but have a nearly identical amino acid sequence. There is a region at the 12th to 13th amino acids that is highly variable (the repeat variable dinucleotide, RVD) and known to mediate nucleotide specificity (see Table 1). The combination of these ~34 amino acid regions and the RVD dictate which DNA sequence is bound.

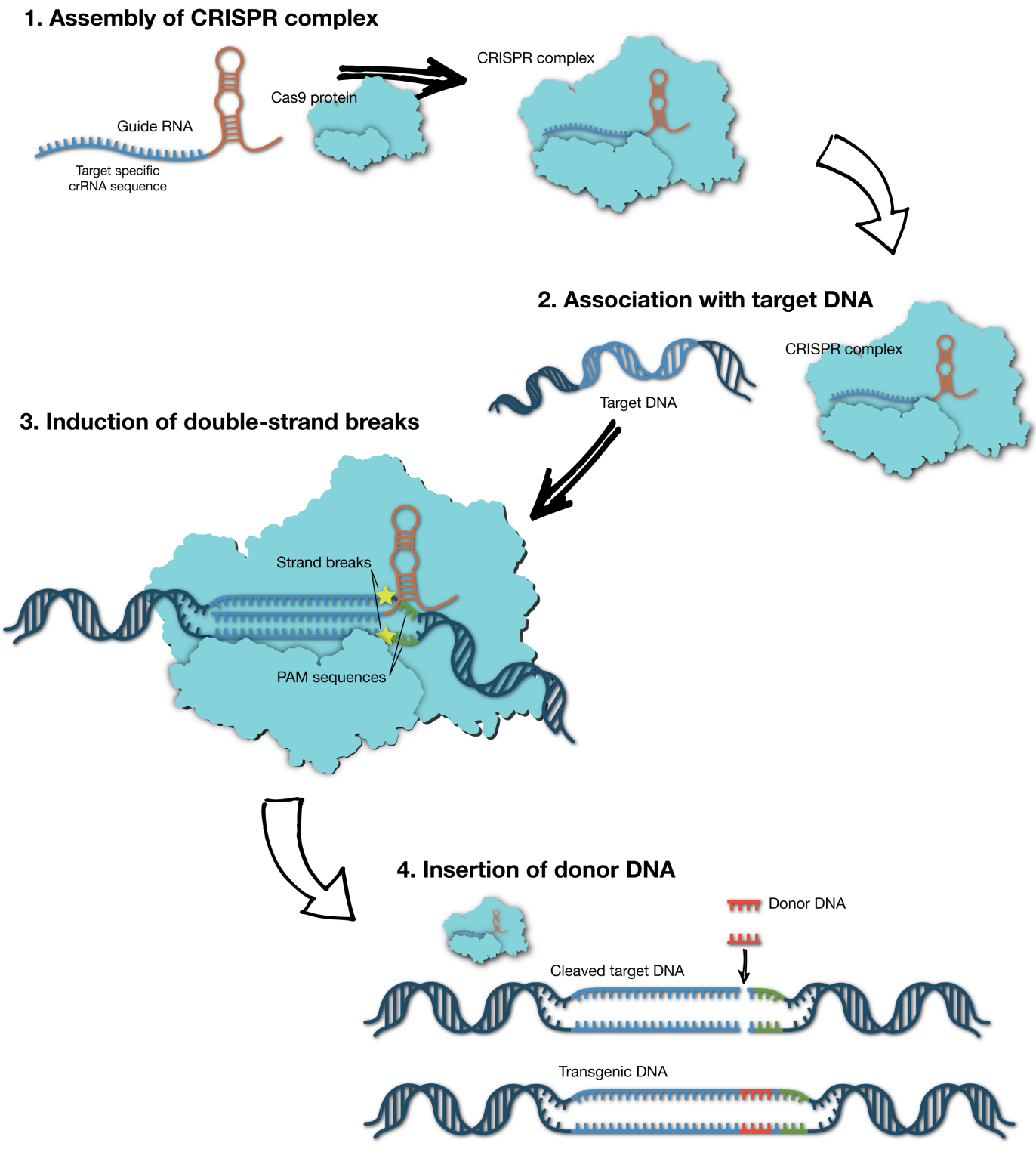


Figure 2: Mechanism of CRISPR/Cas9 genome editing. CRISPR/Cas9 genome editing requires a single guide (sg) RNA that directs the Cas9 endonuclease to a specific region of the genomic DNA, resulting in a double strand break. By providing a donor DNA in trans, a transgenic DNA can be created, whereas in the absence of a donor DNA, the double strand break will be repaired by the host cell, resulting in an insertion or deletion, thus potentially disrupting the open reading frame of a gene.

Therefore, the engineering of a sequence specific binding domain is more easily accomplished than with ZFNs as this is a more modular approach. It requires only varying the RVD nucleotides and not the rest of the repeat backbone to achieve specificity with reasonable avidity. As with ZFNs, this DNA recognition domain can then be fused with a variety of proteins including the *Fok I* catalytic domain.

Like ZFNs, TALENs can be used to make a site-specific cut in genomic DNA, which results in either indel mutations or allows for replacement specific pieces of DNA if a suitable HDR template is delivered. Since the TALE domains allow for a more highly predictable, modular creation of binding domains, they gained immediate popularity when compared with ZFNs because they obviated the need for expensive and time-consuming engineering. Cloning TALEN constructs could be done in 1-2 weeks and the highly modular nature of the binding domains made binding fidelity much more likely. They have been used in cells from a variety of species both in tissue culture as well as *in vivo*. The advantages of TALENs over ZFNs are increased predictability of binding of the TALE domain to the desired target sequence and cost.

Like ZFNs, TALEN construction requires verification *in vitro* to demonstrate an adequate level of cleavage efficiency before they can be employed in experiments. Due to the modular nature of the TALE repeats, the probability of rationally designing a TALEN that will be specific for a particular sequence appears to be greater than that for ZFNs, however it is not possible to rule out off-target cleavage in a given cell line without extensive testing.

Table 1. Nucleotide binding by amino acids of the RVD.

Amino acids in the repeat variable dinucleotide (RVD) at aa 12 and 13	Nucleotide(s) recognized
HG	T, A, C
NG	T
N*	C, T
HD	C
NN	G, A
NS	A, C, G, T
NI	A

C. CRISPR

The identification of clustered regularly interspaced short palindromic repeats (CRISPR) along with the CRISPR associated (Cas) protein is the most recent development and is even more rapid and modular than the TALEN platform. The CRISPR/Cas9 technology is a three-component system consisting of an endonuclease (Cas9), a sequence-specific targeting element (the crRNA), and another RNA that links Cas9 with the crRNA, the tracrRNA (Figure 3) (11). Both crRNA and tracrRNA can be combined in a single molecule termed single guide (sg) RNA (12,13). Since an RNA molecule acts as the guide in this system, and the technology for cloning DNA oligonucleotides has been well developed, this system results in a methodology that is simpler and more rapid to develop than TALENs and ZFNs. CRISPR/Cas9 is now widely used in multiple organisms, both *in vitro* and *in vivo*.

The mechanism of action for the CRISPR/Cas system relies on the expression of at least two components: an effector enzyme and a guide RNA (Figure 2). The effector enzyme most commonly used to date is Cas9 from the type II CRISPR/Cas system from *Streptococcus pyogenes*. Although there are three different types of CRISPR/Cas systems currently known, the type II enzymes are preferable as they are single polypeptides. To be used in this system the CAS9 protein must be expressed and translocated to the nucleus. Several laboratories have created codon-optimized versions with nuclear translocation signal sequences (NLS).

The crRNA recognizes and pairs with a sequence of 20 nucleotides in length within the targeted genome (Figure 3). Targeting relies on sequences at the 3' end of the desired site for digestion. These sequences, known as protospacer adjacent motifs (PAMs), are critical and unique to the CAS9 proteins isolated from different species. Alternative endonucleases such as the type V CPF1 exist that utilize different PAM motifs for targeting (14), thus expanding the repertoire of potential target sites (see Table 2). While the requirement for a PAM motif in the genome is a major restriction in this technique, it is estimated that every gene harbors multiple sites, allowing disruption of virtually any gene.

One advantage of the CRISPR/Cas9 technology is its great versatility. In effect, the sgRNA directs the endonuclease to a specific site in the genome. By attaching functional domains to Cas9, it is possible to target any functionality to a specific genomic location, including promoter sites and introns. For instance, using a translational fusion of catalytically inactive Cas9 to transcriptional activators, such as multimers of the VP16 peptide, and targeting a promoter region, it is possible to activate gene expression (15). Similarly, systems for gene repression, histone modification and epigenetic alterations have been created (16,17,18). In addition, the system can be used to specifically paint genomic sequences or tag sequence regions with a fluorescent protein or protein tag (19,20). It is also possible to generate an inducible CRISPR/Cas9 system either using classical repressor systems or small molecule inhibitors of Cas9 (21). The inventiveness of researchers to create novel applications of this technology seems to be infinite.

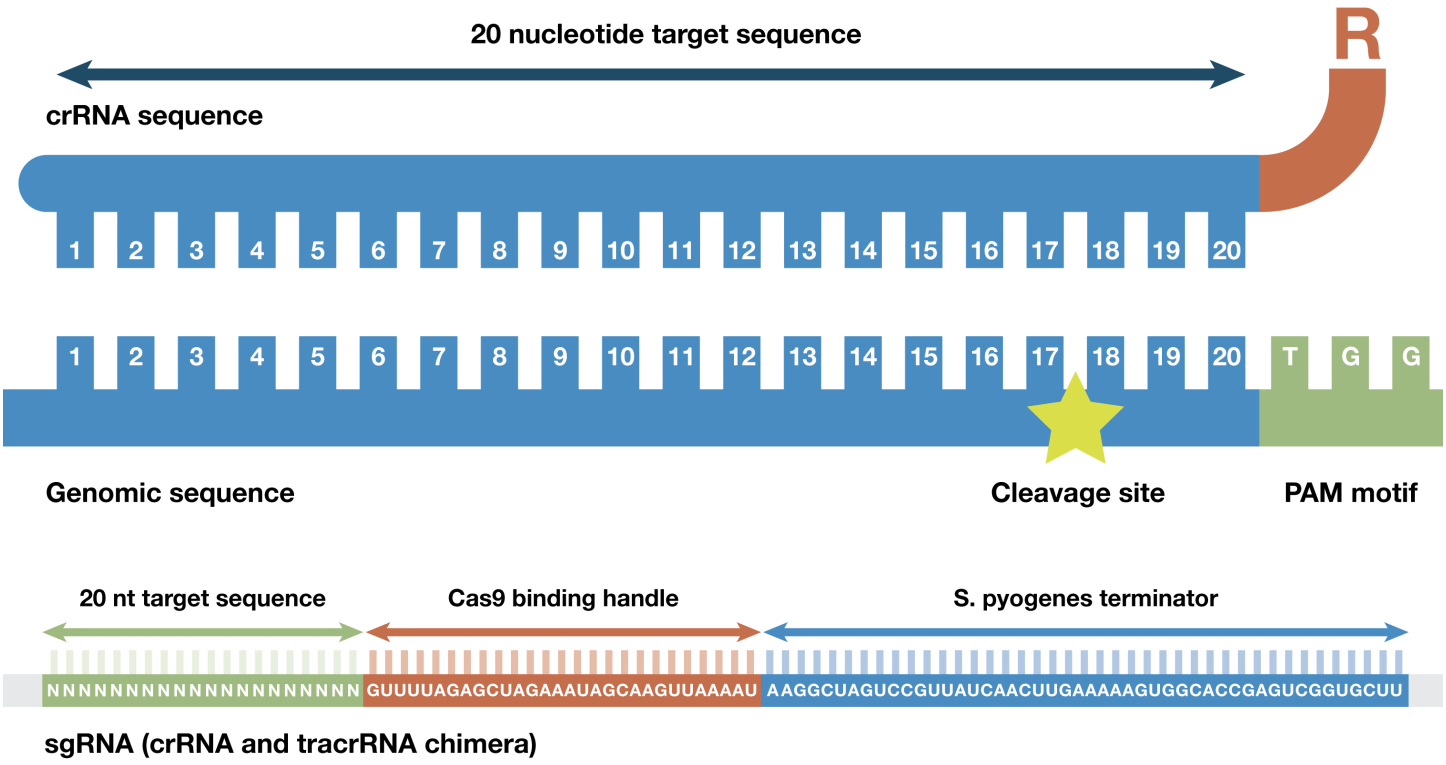


Figure 3: Anatomy of a single guide (sg) RNA. The sgRNA is composed of a crRNA sequence that is fused to the tracrRNA. The crRNA contains a 20 nt sequence that is identical to the genomic target DNA sequence. A pre-requisite for binding of the sgRNA to the target site is the presence of a PAM motif following the 20 nt recognition sequence. "R" denotes the remaining part of the sgRNA, consisting of a Cas9 binding handle and the *S. pyogenes* terminator sequence. Cleavage occurs at the -3 position upstream of the PAM motif.

Table 2: Endonucleases with different PAM specificity.

Name	Organism	PAM	Type	Ref
SpCas9	*Streptococcus pyogenes*	5'-NGG-3'	II	(11)
SpCas9 VQR	*Streptococcus pyogenes*	5'-NGAN-3'	II	(105,106)
SpCas9 EQR	*Streptococcus pyogenes*	5'-NGNG-3'	II	(105,106)
SpCas9 VRER	*Streptococcus pyogenes*	5'-NGCG-3'	II	(105,106)
SaCas9	*Staphylococcus aureus*	5'-NNGRRT-3'	II	(105,107)
NmCas9	*Neisseria meningitides*	5'-NNNNGATT-3'	II	(108)
St1Cas9	*Streptococcus thermophiles*	5'-NNAGAAW-3'	II	(109)
St3Cas9	*Streptococcus thermophiles*	5'-NNGGNG-3'	II	(109)
BlCas9	*Brevibacillus laterosporus*	5'-NNNCND-3'	II	(110)
AsCpf1 LbCpf1	*Acidaminococcus* *Lachnospiraceae*	5'-TTTC-3'	V	(14)
C2c1	*A. acidoterrestris*	5'-TTC-3'	V	(111)

Comparison of the Various Genome Editing Technologies

While commercial availability and technical aspects of design, assembly, and delivery are significant criteria when choosing a nuclease, performance is of equal importance. Factors that govern performance are influenced by the model system of choice (cell line, model species, etc.), the efficiency of nucleic acid delivery, and the presence of polymorphisms in the target region. Indeed, there are examples for each of the major genome editing technologies where very high frequency of modification is observed. Targeting efficiencies of 1%-50% have been reported for TALENS and ZFNs (22,23,24) and efficiencies of up to 70% have been reported for CRISPRs (13,25,26).

Despite the growing amount of data using these systems, there are very few examples of side-by-side comparisons, making it difficult to evaluate which system will work most effectively for an unvalidated target gene or sequence. Recently, the efficiency of TALENs was directly compared to that of CRISPR/Cas9. Site-specific editing of an EGFP transgene in mammalian cells showed that CRISPR/Cas9 is more efficient and precise than TALENs in the absence of a homology-repair DNA template. Contrarily, when supplied with a repair template, TALENs performed more efficiently (27).

Another study has focused on the signature of TALENs and ZFNs (insertions vs deletions). TALENs introduced mostly deletions whereas ZFNs introduced both insertions and deletions (28). However, a more systematic comparison of these techniques is required to draw any conclusions.

Very recently, a direct comparison of CRISPR, CRISPR-interference (CRISPRi) and shRNA on gene modulation was established by using a lethality screen (essential genes). Targeting of 46 essential and 47 nonessential genes showed that CRISPR performed better than the shRNA and CRISPRi methodologies as it allowed a better distinction of the two groups of genes. In this experiment, shRNA-mediated knockdown resulted in more noise (29).

A key aspect of genome editing technologies is their potential use in high-throughput screening applications for understanding gene function and drug target identification. While in principle all mentioned genome editing techniques are suitable for production of gene targeting libraries, there are some practicalities that make the use of the CRISPR/Cas9 the primary choice for such applications. In the next section, we will discuss the advantages of CRISPR/Cas9 in this context and its use in screening.

Screening

Cell-based high-throughput screening in the pharmaceutical industry serves two main purposes: a) the identification of genes underlying phenotypes that correlate with disease (identification of drug targets); and b) the identification of chemical or biological agents that bind to or modulate the activity of such drug targets in cell-based assays (phenotypic screening). Of particular interest in recent years has been high-content screening, due to the ease of use, the potential for multiplexing several readouts, and the high amount of information that can be recorded from relatively simple experiments (30). Nonetheless, enzymatic reporter assays and general viability assays are also very powerful tools even in the age of HCS (31).

Traditionally, modulation of gene expression using genomic libraries has been achieved with siRNA/shRNA libraries (see book-part://[cbrnai]). Although successful in many applications, the use of these libraries presents some challenges. Notably, off-target activities of siRNAs or shRNAs and incomplete knockdown represent major sources of false positives and false negatives, respectively. Nonetheless, the technology has been widely used and been successful for the identification of gene function in multiple organisms. Further, siRNA-based approaches (but also CRISPR/Cas9 activation and repression screens; see below) are preferable when identifying the gene function of essential genes, since in this case complete knockout will obscure any functional readout by disruption of cellular function. The two key reasons why genome editing technologies may overtake the

widespread use of siRNA technology in the near future are: *1) the possibility to create a "clean" knockout of a gene, where residual activity is not present and can be neglected; and 2) the reduced confusion created by off-target activities.*

The use of genome editing technologies in high-throughput approaches has not yet reached the same ease of automation and throughput as other technologies, but protocols are being developed by industry and academic laboratories that will address the current limitations (32).

There is general consensus that the applicability of both TALENs and ZFNs methodologies in screening approaches is more reduced than that of CRISPR. To date, there has been very little use of TALENs, ZFNs and AAV-related techniques for functional genomic screening. This is due to the design complexity and time-consuming generation of TALENs and ZFNs, relative to CRISPR. Nevertheless, an innovative cloning system has recently allowed the generation of a TALEN library comprising TALEN plasmids for 18,740 protein-coding genes (33,34). In a pilot study targeting 126 of those genes, the efficiency was around 90%. Moreover, TALEN-mediated knockout of genes involved in the NF-κB pathway performed better than siRNA knockdown (35). However, no large-scale screens or using TALENs have been published to date. Therefore, we will focus below mainly on the use of CRISPR/Cas9 in screening applications.

The main advantage of the CRISPR/Cas9 technology for screening approaches is that the DNA-targeting molecule is a short RNA. Single guide RNA and crRNA libraries can be produced easily and cost-effectively, thus putting this technology at the forefront of functional genomic screening.

When using genomic libraries to modulate genome sequences such as those mentioned above, two types of screening approaches can be designed: 1. Pooled screening, and 2. Arrayed screening (32). Both have very different requirements in terms of library design, assay design, screen optimization, quality control, hit selection and post-screen validation that is summarized in Table 3 and discussed below.

Table 3: Advantages and disadvantages of pooled vs arrayed screening.

	Pooled screening	Arrayed screening
Advantages	• simple setup • requires no automation and specialized equipment • requires less experimental manipulation • suitable for cell viability/proliferation assays identification of drug resistance genes • comprehensive profiling of genome-wide library • cost-effective	• easy genotype-phenotype correlation • suitable for high-content screening assays • multi-parametric readouts possible • suitable for primary cells and neurons • custom-made libraries are available
Disadvantages	• requires high cell number • not suitable for primary cells and neurons • not suitable for high-content screening • biosafety considerations for virus libraries • requires a selection step • analysis depends on the availability of next-generation sequencing (NGS) • rare gene transcripts may be more difficult to identify • limited number of readouts	• requires specialized equipment (lab automation) • costly and time-consuming • quality control difficult (internal controls and statistical adjustments may be needed) • genome editing efficiency and transfection efficiency may be low • plate and well variability high

Pooled Library Screening

CRISPR/Cas9 screens can be performed in two different formats (arrayed vs. pooled libraries). Most CRISPR-based screens performed to date utilize pooled libraries. In short, a pooled library is a single preparation of many different sgRNA plasmids. The use of pooled libraries requires a positive or negative selection step, and thus has limited use for high-content screening approaches. However, there is great potential for pooled library screening

as the setup is relatively simple and does not require extensive automation and high-throughput screening capabilities. Usual readouts include cell viability and/or proliferation measurements, and the identification of drug resistance genes (36). Several studies reported in the literature comprise investigations in several fields, such as for the identification of cancer tumour suppressor genes (37), viral infection (38), cell cycle regulation and DNA replication. CRISPR/Cas9 screens have been successfully performed not only in mammalian cells but also in Drosophila cells (39) and Zebrafish (40). Recently, a resource summarizing all high-throughput CRISPR/Cas9 screens has been published and is frequently updated (41).

In general, pooled lentiviral screening with CRISPR/sgRNA libraries uses several of the same principles as pooled shRNA screening. This is likely the main reason why pooled screening approaches have been developed very quickly, with essentially all the technology for library production, screening workflow and hit identification already established in key laboratories. Many of these principles can be found in book-part://[cbrnai] and will not be discussed in detail here. However, there are key differences in terms of library design, assay optimization, hit identification and validation that will be outlined below.

Library Design

Pooled Libraries consist of a pool of sgRNAs in a lentiviral vector backbone that is used to transfect or transduce in bulk the cell line of interest. Several tools exist for identification of guide RNA target sequences in the genome. The two key features in producing a high-quality library is a) identification of unique target sites in the genome as to limit potential off-target activities; and b) to predict highest efficiency for genome editing. The identification of unique target sites is usually done by a search for PAM motifs combined with a BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to remove similar regions. Deleterious consequences of off-target activity can be mitigated by strategies described above, and improvements to Cas9-based platforms have been occurring at a rapid pace in the last couple of years resulting in increased fidelity (42,43,44).

The rules for predicting the highest efficiency genome editing sequence targets are not well understood and are lagging behind our understanding of how to select good siRNA molecules. Nonetheless, recent progress has made it clear that there are key determinants that improve design of sgRNA sequence. A recent study has investigated the cleavage activity of sgRNAs on single nucleotide mismatched targets (45). The study found a 4 nucleotide sequence located at +4 to -7 upstream of the PAM motif that can help in the design of gene specific sgRNAs. Another study found nucleotides both at PAM distal and proximal sites to be important for on-target activity (46). Another feature that has been recognized to interfere with genome editing efficiency is PAM density, reasoning that sites harboring multiple PAMs are less amenable to editing (47). Further determinants for improved design have been made (48), suggesting that G is preferred at the -1 and -2 positions proximal to the PAM sequence, a preference for C at the -3 position, a preference for A in positions from -5 to -12 and for G in positions from -14 to -17. However, this prediction is based on a limited set of data and the overall design rules for efficient guide RNAs need to be further investigated.

Several online CRISPR design tools now exist, and these enable the prediction of off-target sites, along with the probability of those events based on the individual sgRNA designs. Some of these are summarized in recent reviews (49,50). A list of online tools and websites for sgRNA design can be found in the appendix.

A very interesting experimental approach to generate sgRNA libraries from any gene without the need for bioinformatics prediction of target sequences has recently been described as an alternative (51). This method relies on the use of a semi-random primer containing a PAM complementary sequence and standard molecular biology tools to generate a sgRNA library. This method is particularly useful for generating sgRNA libraries for species where limited sequence information is available.

Several companies have developed their own sgRNA design tools, which make it easy to design and order the reagents with a very simple two-click system. Often, the problem is that too many sequences are identified and

selection becomes a matter of being spoiled for choice. Given that most sgRNA sequences work reasonably well, it might be best to opt for location in the gene sequence rather than predicted efficiency. For instance, sequences close to the 5'-end of a gene might be favorable, but one should keep in mind that alternative start sites might result in expression of a downstream gene fragment. Also, ribosomal frameshifting (52) and "illegitimate translation" (53) can result in expression of a gene from out-of-frame alleles. An alternative is to target within a functional or catalytic domain. Often, even an in-frame insertion can result in disruption of protein function, as we have observed for the ATG4B protease (Robin Ketteler, unpublished observation).

Currently available libraries are listed in the appendix.

Assay Optimization

A pooled CRISPR screen typically requires three components: the sgRNA library, Cas9 endonuclease, a cell model of interest and a sequence analyzer for hit identification.

Cells: When choosing the right cell line, the first choice should always be that the most relevant cell type is used to address the specific biological question. However, one might want to consider the ploidy of cells. For instance, in diploid cells, one would need to edit two alleles in order to get a complete loss of function phenotype. In polyploid cells, this might be more complicated. To date there is no evidence that the number of alleles has an effect on genome editing efficiency, although it has been suggested that the efficiency may be lower with increasing number of alleles (54). Another consideration is that 1 in 3 modifications can result in an in-frame indel, thus potentially not leading to a loss of gene function. Identified phenotypes can therefore be a consequence of mosaicism in gene function in a certain cell. Furthermore, interfering with the endogenous DNA repair mechanism may interfere with genome editing efficiencies. All of these caveats will require addressing when validating individual hits from pooled library screening.

Delivery of library: Often, the pooled library is supplied as a glycerol stock where each plasmid needs to be amplified before use. This is a critical step, as during amplification one has to ensure that the representation of each sgRNA is maintained. In a next step, the plasmids will be delivered to producer cells that produce the lentivirus particles for transduction of the target cells. This step is critical and assessment of virus titer is necessary. During transduction of target cells, the key requirement is that each cell receives only one virus on average. This can be best achieved by transducing cells with a mixture of lentiviral particles at a low multiplicity of infection that allows a single integration per cell. On the other hand, one has to ensure that each plasmid is represented in the infected cell population. Therefore, an excess of cells is usually used, at least 100-fold higher than the number of plasmids. This also means that a large portion of cells will not be transduced. Another consideration in pooled lentiviral screening that affects the number of cells required for hit identification is that each transcript is represented at high enough numbers so that even rare transcripts can be identified. For this reason, the number of cells in the beginning of the screen should be sufficiently high to enable such representation. Typically, 100-200 million cells are transduced and screened.

Delivery of Cas9 endonuclease: The CAS9 endouclease can be delivered in conjunction with the sgRNA (using the same vector) or provided in trans, either in the form of a stable Cas9 expressing cell line or through transient delivery of the Cas9 gene or protein into cells. Some studies have suggested that the transient delivery of Cas9 protein can be advantageous to reduce off-target effects when compared to a stable Cas9 expressing cell line. However, for practical reasons and to reduce variability between batches, a stable Cas9 cell line is often used.

Selection: In the next step, a positive or negative selection will be applied to allow the identification of sgRNA-mediated editing events that lead to a desired phenotype. In a positive selection screen, a positive selective pressure will eliminate all cells that do not survive the selection (e.g. treatment with puromycin). The few cells that survive can then be sequenced and the underlying sgRNA identified. Positive selection screens are typically very robust, if the selection pressure is very strong. Obviously, the right amount of selective pressure has to be

optimized in a first experiment on a control population. In a negative selection screen, most cells will survive. Therefore, one has to perform next-generation sequencing on the starting population, apply selection, and perform another round of next-generation sequencing on the surviving cells. By comparing the results from the initial population with the selected population, one can identify the list of gRNAs that are underrepresented during negative selection. For both, positive and negative selection screens, access to next-generation sequencing platforms is essential. An alternative to positive or negative selection is the use of a fluorescent marker protein and cell sorting.

Pooled library screening is not feasible for cell types that have limited proliferative potential or a limited passage number such as neurons or primary cells. Also, pooled screening is limited for identification of morphological phenotypes, such as those typically observed in high-content screening assays.

Hit Identification

Pooled screens require a selection step in order to enrich for cells displaying the desired phenotype. Ultimately, this will result in the differential representation of transcripts in the selected cell population that can be identified by next-generation sequencing. The identity of each sgRNA in the final population can be determined by one of two approaches: the sgRNA sequence can be determined by designing primers for PCR amplification that recognize the vector backbone and are flanking the sgRNA, thus effectively directly sequencing the sgRNA. An alternative is to use sgRNA-specific barcodes in the lentiviral vector backbone that are sequenced and then infer the corresponding sgRNA sequence. The advantage of a barcode approach is that probe sequences can be designed so that optimal hybridization settings can be obtained for each plasmid, resulting in equal PCR amplification and representation of transcripts. Identification of sequences can be done by microarrays (55) or deep sequencing. Recently, a multiplexed barcoded library approach has been developed that allows hit identification in a high-throughput manner (56). In order to capture underrepresented sgRNAs, it is important to reach a high level of transcript representation, and – as mentioned above – to start with a sufficiently high cell number. There are a few things worth noting:

Sequencing Depth: A key consideration is sequencing depth. It has been estimated that more than $1x10^7$ reads are required for a complex library. While normalization is usually not required, it is important to assess the representation of each sgRNA in the initial library before selection. Therefore, a cumulative distribution normalized to the total number of reads may help in this. This can also be important for quality control, as deviations from this curve may indicate a loss of diversity in the starting population.

Representation/Reads: After determination of the representation of each sgRNA, it is important to apply statistical measures to determine the significance of the results. Several user-friendly programs have been designed to aid in statistical analysis, including RIGER (57), RSA (58), HitSelect (59) and MaGeCK (60).

Recently, an R package for data analysis and documentation of pooled CRISPR library screens has been developed (caRpools (61)).

Validation

Post-screen validation of hits from pooled and arrayed library screening use common principles and will be discussed below.

Arrayed Library Screening

The use of arrayed libraries is still in its early days and only recently have they become available. Arrayed libraries present a broader applicability, allowing direct analysis of image-based phenotypes as well as biochemical readouts based on intensity (colorimetric/fluorescence/luminescence). In this sense, the use of arrayed libraries is more direct. In some cases a selection step for the cells that incurred a genetic modification

may be required, whereas for robust biochemical readouts (such as luminescence), this step might be unnecessary. The same is true when the targeting efficiency is very high.

Library Design

Guide RNA design follows very similar rules for arrayed libraries as for pooled libraries. Tools for identifying target sequences are described above and listed in the appendix. Here, we will only focus on particular requirements for arrayed libraries.

An arrayed library usually consists of guide or crRNAs targeting a single gene within each well in a multi-well microplate format. In some cases, there is more than one sgRNA/crRNA per well, but the key point is that each well contains agents that target the same single gene. The main advantage is that the gene-function correlation can be much more easily established than by selection of over- or underrepresented sequences in a mixed pool.

Unlike pooled libraries, arrayed libraries can occur in many different formats.

Lentiviral arrayed libraries are available, and these often contain the same vectors and sequences as the corresponding pooled libraries. The use of lentiviral libraries provide an efficient delivery to almost any cell type, but may require a selection step and culture of the cells for extended time (>7 days) before analysis.

Synthetic crRNA libraries: Synthetic libraries of small RNA molecules have been successfully delivered to cells at high efficiency and with similar protocols developed for siRNA oligonucleotides. In this case, high transfection efficiency is essential as there is no ability to select cells. Experiments are on a shorter time frame (within a few days). When using crRNAs, the tracrRNA has to be provided in trans alongside the Cas9 endonuclease.

In vitro transcribed sgRNA libraries: Another format that is being offered, although usually at lower scale due to the cost in manufacturing, is an *in vitro* transcribed sgRNA library. These are being offered as custom libraries and it is unlikely that they will be provided as an off-the-shelf solution for large libraries.

It is worth noting that chemically synthesized crRNAs have been suggested to be more versatile due to the possibility of having incorporated chemical modifications that increase gene editing efficiency (see review (62)). It can be expected that further improvements in this area may result in alternative library design in the future.

The key advantages of arrayed libraries are that the gene-function correlation is much easier to establish, the handling can be automated using standard liquid handlers, and an arrayed format is easily adaptable to high-content screening technologies. Also, it is possible to deliver the Cas9-sgRNA as a ribonucleoprotein complex, which has been suggested to reduce off-target activities (63). A key disadvantage is the cost of arrayed libraries, plus uncertainties associated with the genome editing efficiency, which may be below the detection limit of conventional high-content screening assays (see Table 3).

The libraries currently available are designed to target the whole genome, individual classes of genes (e.g. tumour suppressors, hormone receptors) and specific pathways (e.g. angiogenesis, apoptosis, cell cycle). Additionally, custom-made libraries that match the need of the research community are also available. Resources for pooled and arrayed libraries are listed in the appendix.

Assay Optimization

An optimal assay is one that delivers the maximum number of reproducible positive hits and a minimum of false negatives. In arrayed CRISPR library screening, very similar measures can be applied as in siRNA or small molecule library based approaches. Thus, the optimization of an assay starts with evaluation of a positive (i.e. exerting the maximum known phenotypic effect) and a negative (i.e. neutral) control. A good signal-to-noise window is necessary and statistical measures for effect size such as a Z'-factor can be used to evaluate the reagents (64). One complication in CRISPR based approaches is that penetrance of a phenotype is typically

much lower than for siRNA or small molecule screening approaches. For instance, if genome editing efficiency is below 20%, it becomes hard to identify a phenotype in image-based approaches (unless the phenotype is an off-on switch of a fluorescently labeled reporter gene). Assays with a low background noise and a high signal-to-noise window obviously perform better than assays that record morphological changes in organelle structure or number.

The main issue is to ensure that the observed phenotype is indeed linked to the desired mutation or edited locus.

The parameters that will affect this are:

- transfection efficiency (the phenotype should only be observed in cells that have been transfected)
- genome editing efficiency (the phenotype should only be observed in cells that have been successfully edited)
- off-target activity (the phenotype should only be observed in cells that have been edited at the correct location)

Accordingly, controls should be deployed to test each of these steps.

- For transfection/transduction efficiency, one could include a control such as a plasmid encoding for a fluorescent protein that can be easily quantified. Alternatively, a reagent that results in cell death can be used, such as sgRNAs targeting the Plk1 gene, which is often used as readout for the setup of RNAi screens. As a negative control, one could use a scrambled sgRNA or a sgRNA targeting an exogenous gene such as firefly luciferase or GFP.
- Assessing genome editing efficiency during a screen is often difficult. Ideally, a positive control could be used here, targeting a gene that is known to exert a strong phenotypic effect. Genes often used are HPRT1, TK or DHFR, which can be used as selection markers.
- Off-target activity is most often assessed post-screening, when secondary assays are used to validate hits. However, off-target effects can be potentially reduced by selecting the right library in the beginning, as discussed above.

Depending on the method of delivery, the optimal conditions of transfection/transduction should be determined beforehand. For lentiviral particles, the multiplicity of infection (MOI) that gives the best transduction and genome editing efficiency must be empirically determined. Unlike pooled screens where it is crucial to have one integrant per cell, in arrayed screening it is less important to restrict the number of integrations per cell as all viral particles within one well will target the same gene. Therefore much higher MOIs are typically used. We recommend the use of an MOI of 2-5 for the sgRNA particles.

Assessing genome editing efficiency is usually done for identified hits post-screening. There are possibilities for investigating genome editing in a high-throughput manner, though, and if costs for such methods decrease in the coming years, it is in principle feasible to do this important quality control step while screening. The technologies for assessment of genome editing efficiency will be discussed below in the "validation" section.

Hit Identification

Hits in arrayed CRISPR library screening can be identified by the commonly used Z score measures that incorporate sample size and deviations from normal distribution to identify genes that enhance or reduce a phenotype. When edge effects might interfere with the assay, a B score can be applied post-screening to normalize for these effects. Often, a simple signal to background calculation is sufficient to identify the strongest phenotypes.

Validation

For high-throughput screening using genome editing technologies, it is very important to attribute the phenotypic effect to a successfully edited genome modification. There are multiple methods available, from a very simple setup requiring only a PCR instrument to very complex technologies that rely on next-generation sequencing (Table 4). Accordingly, the cost, but most importantly the potential of each technique is very different and requires careful assessment before choosing the appropriate method.

Table 4: Overview of validation methods.

Method	Cost	Resolution	High-throughput potential	Equipment	Comments
Sanger sequencing	low	high	yes	Sequencer	needs sub-cloning, time-consuming; can use deconvolution software
Mismatch Cleavage Detection Assay	low	low	no	Standard lab equipment	not sensitive, not quantitative
Poly-acrylamid gel electrophoresis	low	low	no	Standard lab equipment	simple, cheap method, but non-quantitative
Indel Detection by Amplicon Analysis (IDAA)	moderate	moderate	yes	PCR machine	simple method with good resolution
High Resolution Melt analysis (HRM)	low	high	yes	RT-PCR machine	simple, cheap method with high resolution
Competitive PCR	low	moderate	yes	PCR machine	simple and cheap method, but non-quantitative
Digital Droplet PCR	high	high	yes	Digital Droplet PCR machine	simple method with high resolution, amenable to 96-well format
Next generation sequencing	high	high	yes	Next generation sequencer	most comprehensive analysis of genome editing; has potential to detect off-target effects

PCR based assays

Mismatch Cleavage Detection Assay: Technologies with relatively low throughput include mismatch cleavage assay detection (65) such as the Surveyor assay (Invitrogen) and GeneArt Genomic Cleavage Detection Kit (ThermoFisher) that utilize the properties of an error-prone polymerase for amplification of fragments that harbor mismatch mutations.

PCR-PAGE: An alternative is the use of poly-acrylamide gel electrophoresis, where PCR fragments are generated from the genomic region, subjected to denaturation and re-annealing, and analysis of differential patterns of migration by homo- or heteroduplexes (66).

High Resolution Melting Curve Analysis: An alternative with higher throughput is high resolution melting (HRM) curve analysis (67). HRM makes use of different melting curve patterns for two individual PCR products and is efficient to detect indels with a resolution close to 1 nucleotide.

Indel Detection by Amplicon Analysis (IDAA): Another PCR-based method termed Indel Detection by Amplicon Analysis (IDAA) relies on tri-primer labeling of a PCR product and detection by gel electrophoresis (68).

Competitive PCR: A simple and cheap semiquantitative method to identify mutated cells based on competition-based PCR has been developed, where a mixture of three primers with one primer overlapping the Cas9 cleavage site can be used (69).

Digital Droplet PCR: Recently, Digital Droplet PCR has been used to detect genome editing events in a high-throughput manner (70). This method is based on partitioning of PCR reactions on single targets into droplets and using a fluorescence-based readout for each individual droplet PCR to visualize changes in the DNA sequence. Quantification of edited and wild-type alleles can be done in the same sample. More recent versions of this approach have enabled the simultaneous detection of NHEJ and HDR events in a single reaction mixture (71).

Sequencing based assays

Sanger Sequencing: Traditionally, genome editing has been assessed for single genes by sub-cloning the affected locus followed by Sanger sequencing. However, this method is time-consuming and not amenable to large-scale interrogation of genome editing events.

Next-Generation Sequencing: One caveat of the PCR based assays is that they only detect that a mutation has occurred but they do not detect the specific sequence alterations exerted by genome editing. Next-generation sequencing technologies can overcome this limitation. For instance, a variant of next generation sequencing, CRISPR Genome Analyzer (CRISPR-GA), can be used to map the specific alterations detected (72). CRISPR-GA provides information on size and location of indels, as well as the efficiency of NHEJ and HDR events. This is a huge advantage in terms of validation, as one could use this to determine if the phenotype corresponds to an edited site. An alternative is BATCH-GE (73), which provides batch analysis features, thus reducing the time required for analysis.

Additional validation steps

After a cell is identified as carrying the appropriate mutation additional steps are necessary for validation that the phenotype is directly the result of the known mutation and not due to some other effect. This can be done using secondary phenotypic assays which can include one or more of the following:

1. Rescue of the phenotype by expression of an exogenous normal copy of the mutated gene. This can be achieved by transfection or transduction of a plasmid or retrovirus particle, respectively. One caveat is that overexpression may result in a higher copy number, thus not representing the original endogenous levels.
2. Reversion of the genome modification back to wild-type by genome editing. In this approach it is important to include controls to demonstrate that the mutated cell line has been rescued and that the initial unmutated cell line has not been "reisolated" due to contamination. This can be easily accomplished by including insertion of silent mutations into the reversion step.

Conclusion

In summary, genome editing technologies have greatly advanced over the last decade and provided new insights into the function of genes and proteins. Recently, these technologies have been applied to high-throughput screening applications, with the CRISPR/Cas9 system at the forefront of this development. It can be expected that further improvements to this technology will be made in this fast-moving field and some of the challenges such as delivery and efficiency will be overcome in the not too distant future.

Acknowledgements

This work was supported by the Medical Research Council, grant number MC_U12266B.

Appendix

Commercial availability of ZFNs, TALENs, CRISPR/Cas9

ZFNs,TALENs and rAAVs may be custom synthesized or created. They may be ordered either by user design or by having the company design and synthesize and clone the relevant protein. There is a consortium for the development of ZFNs.

As of 2017, the following companies are known to provide ZFN synthesis services and/or reagents to create ZFNs:

- Sigma-Aldrich (www.sigma-aldrich.com)
- Available through Addgene (www.addgene.com)
  - Oligomerized Pool Engineering
  - Context dependent assembly
  - Modular assembly (Barbas kit, Joung kit, Wolfe system)

As of 2017, the following companies offer Tal synthesis services and or reagents to create TALENs:

- Cellectis Bioresearch (http://www.cellectis.com/)
- Life Technologies (www.lifetechnologies.com)
- Available through Addgene (www.addgene.com)
  - FLASH assembly
  - Modular assembly (Voytas kit, Joung kit, Zhang kit)

As of 2017, the following companies offer CRISPR/Cas9 development services:

- Agilent
- Dharmacon/GE
- Genecopeia
- Horizon Genomics
- IDT
- MolDiag Solutions
- Sigma/Merck
- Synthego
- ThermoFisher

Online Resources for Guide and crRNA Design

When choosing the best design algorithm for sgRNA, there are multiple choices (see below). All have in common a basic search algorithm for a PAM motif sequence and some blast search against potential off-target effects, i.e. similar sequences. The most useful websites are the ones that provide clear information and visualization of results, for example when displaying the location of the cut site, providing a suggestion of potential sequencing primers and scores about similar sequences that are easy to understand.

Design tools	Considerations	Reference
sgRNA Scorer 2.0	Identify putative guide sequences and assign a predicted activity. Considers off-target effects. Uses CasFinder algorithm.	(74)
CRISPR-DO (CRISPR Design and Optimization)	Predicts sgRNA efficiency and off-target scoring; applied to several genomes	(75)

Table continued from previous page.

Design tools	Considerations	Reference
Drosophila RNAi Screening Center (DRSC) Find CRISPRs	Very good graphical visualization in JBrowse. Considers off-target effects.	(76)
ge-CRISPR	Evaluates potential off-target sequences; very user-friendly tool	(77)
CT-Finder	Graphical visualization on JBrowse. Considers off-target effects. Applied to several genomes	(78)
CHOPCHOP	Considers off-target effects. Very informative table. Cas9 nuclease, Cas9 nickase, Cpf1 and TALEN	(79)
Breaking-Cas	Interactive design	(80)
Cas-Database	Graphical visualization in JBrowse. Both Cas9 and Cpf1	(81)
CLD (CRISPR Library Designer)	Design of CRISPR/Cas9 libraries. Also suitable for targeting non coding regions.	(82)
WU-CRISPR	Human or mouse genomes. No graphical visualization.	(83)
CRISPR-GA (Genome Analyzer)		(72)
CRISPRscan	Predictions are available as tracks that can be uploaded in the UCSC genome browser. Applies to several species	(84)
CRISPR-ERA	Editing, repression and activation; Cas9 nuclease and Cas9 nickase. Predictions are available as tracks that can be uploaded in the UCSC genome browser.	(85)
Protospacer Workbench		(86)
WGE | Sanger Institute	Along with Gibson assembly PCR oligo designer. Genoverse browser	(87)
E-CRISP		(88)
Phyto-CRISP-Ex	Protist genomes, 2 sequential filters are applied.	(89)
CRISPR-P	Applicable to Plants	(90)
Off-Spotter	Interactive way of selecting the seed sequence	(91)
CRISPR Direct		(92)
CRISPRseek	Bioconductor package	(93)
sgRNAcas9	Software package	(94)
GPP Web portal sgRNA designer for CRISPRa and CRISPRi	Gene activation and gene repression. No graphical visualization	(95)

Available Libraries for Screening

As of this date (2017), the following libraries for screening using CRISPR/Cas9 are available:

Commercial	Sanger Whole Genome & Sigma CRISPR arrayed libraries Thermo Fisher (LentiArray CRISPR Libraries, TrueGuide CRISPR Libraries) Dharmacon (synthetic guide RNA and lentiviral guide RNA) Genecopeia GenScript Horizon Adgene Agilent

Table continued from previous page.

Academic	Pooled lentiviral libraries: Mouse genome-wide lentiviral CRISPR-gRNA library (96) Human and mouse genome-wide lentivrial CRISPR/-gRNA library (GECKO) (97,98) Human genome-wide lentiviral CRISPR/gRNA library (99) Drosophila genome-wide CRISPR library (39) Human genome-wide lentiviral library (Toronto KnockOut TKO) (100) Human and mouse CRISPRi and CRISPRa libraries (101) Human lncRNA targeting library (102) Arrayed libraries: Human genome-wide, sequence-verified, arrayed lentiviral CRISPR library (103) Human and mouse lentiviral arrayed library (104)

References

1. Gaj T, Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology. 2013;31:397–405. PubMed PMID: 23664777.
2. Hirata R, Chamberlain J, Dong R, Russell DW. Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. Nature biotechnology. 2002;20:735–738. PubMed PMID: 12089561.
3. Daya S, Berns KI. Gene therapy using adeno-associated virus vectors. Clinical microbiology reviews. 2008;21:583–593. PubMed PMID: 18854481.
4. Russell DW. AAV vectors, insertional mutagenesis, and cancer. *Molecular therapy : the journal of the American Society of Gene Therapy* **15**, 1740-1743 (2007).
5. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proceedings of the National Academy of Sciences of the United States of America. 1996;93:1156–1160. PubMed PMID: 8577732.
6. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic acids research. 2005;33:5978–5990. PubMed PMID: 16251401.
7. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. Proceedings of the National Academy of Sciences of the United States of America. 1992;89:4275–4279. PubMed PMID: 1584761.
8. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. The EMBO journal. 1985;4:1609–1614. PubMed PMID: 4040853.
9. Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. Current opinion in biotechnology. 2012;23:644–650. PubMed PMID: 22342754.
10. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nature reviews Molecular cell biology. 2013;14:49–55. PubMed PMID: 23169466.
11. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013;2:e00471. PubMed PMID: 23386978.
12. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819–823. PubMed PMID: 23287718.
13. Mali P, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339:823–826. PubMed PMID: 23287722.
14. Zetsche B, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163:759–771. PubMed PMID: 26422227.
15. Gilbert LA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013;154:442–451. PubMed PMID: 23849981.
16. Qi LS, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152:1173–1183. PubMed PMID: 23452860.

17. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature protocols. 2013;8:2180–2196. PubMed PMID: 24136345.
18. Hilton IB, Gersbach CA. Enabling functional genomics with genome engineering. Genome research. 2015;25:1442–1455. PubMed PMID: 26430154.
19. Chen B, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013;155:1479–1491. PubMed PMID: 24360272.
20. Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S, Pederson T. Multicolor CRISPR labeling of chromosomal loci in human cells. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:3002–3007. PubMed PMID: 25713381.
21. Rauch BJ, *et al.* Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150-158 e110 (2017).
22. Miller JC, et al. A TALE nuclease architecture for efficient genome editing. Nature biotechnology. 2011;29:143–148. PubMed PMID: 21179091.
23. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic acids research. 2011;39:9283–9293. PubMed PMID: 21813459.
24. Maeder ML, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Molecular cell. 2008;31:294–301. PubMed PMID: 18657511.
25. Hwang WY, et al. Heritable and precise zebrafish genome editing using a CRISPR-Cas system. PloS one. 2013;8:e68708. PubMed PMID: 23874735.
26. Feng Z, et al. Efficient genome editing in plants using a CRISPR/Cas system. Cell research. 2013;23:1229–1232. PubMed PMID: 23958582.
27. He Z, Proudfoot C, Whitelaw CB, Lillico SG. Comparison of CRISPR/Cas9 and TALENs on editing an integrated EGFP gene in the genome of HEK293FT cells. SpringerPlus. 2016;5:814. PubMed PMID: 27390654.
28. Kim Y, Kweon J, Kim JS. TALENs and ZFNs are associated with different mutation signatures. Nature methods. 2013;10:185. PubMed PMID: 23396284.
29. Evers B, Jastrzebski K, Heijmans JP, Grernrum W, Beijersbergen RL, Bernards R. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nature biotechnology. 2016;34:631–633. PubMed PMID: 27111720.
30. Ketteler R, Kriston-Vizi J. High-Content Screening in Cell Biology. *Waltham, MA: Academic Press* **4**, 234-244 (2016).
31. Ketteler R. The Feynman trajectories: determining the path of a protein using fixed-endpoint assays. Journal of biomolecular screening. 2010;15:321–326. PubMed PMID: 20130209.
32. Agrotis A, Ketteler R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. Frontiers in genetics. 2015;6:300. PubMed PMID: 26442115.
33. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nature biotechnology. 2012;30:460–465. PubMed PMID: 22484455.
34. Reyon D, *et al.* Engineering customized TALE nucleases (TALENs) and TALE transcription factors by fast ligation-based automatable solid-phase high-throughput (FLASH) assembly. *Current protocols in molecular biology / edited by Frederick M Ausubel [et al]* **Chapter 12**, Unit 12 16 (2013).
35. Kim Y, et al. A library of TAL effector nucleases spanning the human genome. Nature biotechnology. 2013;31:251–258. PubMed PMID: 23417094.
36. Kurata M, et al. Using genome-wide CRISPR library screening with library resistant DCK to find new sources of Ara-C drug resistance in AML. Scientific reports. 2016;6:36199. PubMed PMID: 27808171.
37. Katigbak A, Cencic R, Robert F, Senecha P, Scuoppo C, Pelletier J. A. CRISPR/Cas9 Functional Screen Identifies Rare Tumor Suppressors. Scientific reports. 2016;6:38968. PubMed PMID: 27982060.
38. Park RJ, et al. A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. Nature genetics. 2017;49:193–203. PubMed PMID: 27992415.

39. Bassett AR, Kong L, Liu JL. A Genome-Wide CRISPR Library for High-Throughput Genetic Screening in Drosophila Cells. *Journal of genetics and genomics = Yi chuan xue bao* **42**, 301-309 (2015).
40. Shah AN, Davey CF, Whitebirch AC, Miller AC, Moens CB. Rapid reverse genetic screening using CRISPR in zebrafish. Nature methods. 2015;12:535–540. PubMed PMID: 25867848.
41. Rauscher B, Heigwer F, Breinig M, Winter J, Boutros M. GenomeCRISPR - a database for high-throughput CRISPR/Cas9 screens. Nucleic acids research. 2017;45:D679–D686. PubMed PMID: 27789686.
42. Cho SW, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome research. 2014;24:132–141. PubMed PMID: 24253446.
43. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nature biotechnology. 2014;32:279–284. PubMed PMID: 24463574.
44. Ran FA, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154:1380–1389. PubMed PMID: 23992846.
45. Zheng T, et al. Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. Scientific reports. 2017;7:40638. PubMed PMID: 28098181.
46. Liu X, Homma A, Sayadi J, Yang S, Ohashi J, Takumi T. Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. Scientific reports. 2016;6:19675. PubMed PMID: 26813419.
47. Malina A, Cameron CJ, Robert F, Blanchette M, Dostie J, Pelletier J. PAM multiplicity marks genomic target sites as inhibitory to CRISPR-Cas9 editing. Nature communications. 2015;6:10124. PubMed PMID: 26644285.
48. Xu H, et al. Sequence determinants of improved CRISPR sgRNA design. Genome research. 2015;25:1147–1157. PubMed PMID: 26063738.
49. Chuai GH, Wang QL, Liu Q. In Silico Meets In Vivo: Towards Computational CRISPR-Based sgRNA Design. Trends in biotechnology. 2017;35:12–21. PubMed PMID: 27418421.
50. Mohr SE, Hu Y, Ewen-Campen B, Housden BE, Viswanatha R, Perrimon N. CRISPR guide RNA design for research applications. The FEBS journal. 2016;283:3232–3238. PubMed PMID: 27276584.
51. Arakawa H. A method to convert mRNA into a gRNA library for CRISPR/Cas9 editing of any organism. Science advances. 2016;2:e1600699. PubMed PMID: 27574704.
52. Ketteler R. On programmed ribosomal frameshifting: the alternative proteomes. Frontiers in genetics. 2012;3:242. PubMed PMID: 23181069.
53. Makino S, Fukumura R, Gondo Y. Illegitimate translation causes unexpected gene expression from on-target out-of-frame alleles created by CRISPR-Cas9. Scientific reports. 2016;6:39608. PubMed PMID: 28000783.
54. Miles LA, Garippa RJ, Poirier JT. Design, execution, and analysis of pooled in vitro CRISPR/Cas9 screens. The FEBS journal. 2016;283:3170–3180. PubMed PMID: 27250066.
55. Brummelkamp TR, et al. An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. Nature chemical biology. 2006;2:202–206. PubMed PMID: 16474381.
56. Wong AS, et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. Proceedings of the National Academy of Sciences of the United States of America. 2016;113:2544–2549. PubMed PMID: 26864203.
57. Luo B, et al. Highly parallel identification of essential genes in cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:20380–20385. PubMed PMID: 19091943.
58. Konig R, et al. A probability-based approach for the analysis of large-scale RNAi screens. Nature methods. 2007;4:847–849. PubMed PMID: 17828270.
59. Diaz AA, Qin H, Ramalho-Santos M, Song JS. HiTSelect: a comprehensive tool for high-complexity-pooled screen analysis. Nucleic acids research. 2015;43:e16. PubMed PMID: 25428347.
60. Li W, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome biology. 2014;15:554. PubMed PMID: 25476604.
61. Winter J, et al. caRpools: an R package for exploratory data analysis and documentation of pooled CRISPR/Cas9 screens. Bioinformatics. 2016;32:632–634. PubMed PMID: 26508755.

62. Kelley ML, Strezoska Z, He K, Vermeulen A, Smith A. Versatility of chemically synthesized guide RNAs for CRISPR-Cas9 genome editing. Journal of biotechnology. 2016;233:74–83. PubMed PMID: 27374403.
63. Liang X, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. Journal of biotechnology. 2015. PubMed PMID: 26003884.
64. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. Journal of biomolecular screening. 1999;4:67–73. PubMed PMID: 10838414.
65. Langhans MT, Palladino MJ. Cleavage of mispaired heteroduplex DNA substrates by numerous restriction enzymes. Current issues in molecular biology. 2009;11:1–12. PubMed PMID: 18541926.
66. Zhu X, et al. An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. Scientific reports. 2014;4:6420. PubMed PMID: 25236476.
67. Thomas DK, et al. Comparative variation within the genome of Campylobacter jejuni NCTC 11168 in human and murine hosts. PloS one. 2014;9:e88229. PubMed PMID: 24516617.
68. Yang Z, et al. Fast and sensitive detection of indels induced by precise gene targeting. Nucleic acids research. 2015;43:e59. PubMed PMID: 25753669.
69. Harayama T, Riezman H. Detection of genome-edited mutant clones by a simple competition-based PCR method. PloS one. 2017;12:e0179165. PubMed PMID: 28586390.
70. Nelson CE, et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science. 2016;351:403–407. PubMed PMID: 26721684.
71. Miyaoka Y, et al. Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. Scientific reports. 2016;6:23549. PubMed PMID: 27030102.
72. Guell M, Yang L, Church GM. Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). Bioinformatics. 2014;30:2968–2970. PubMed PMID: 24990609.
73. Boel A, et al. BATCH-GE: Batch analysis of Next-Generation Sequencing data for genome editing assessment. Scientific reports. 2016;6:30330. PubMed PMID: 27461955.
74. Chari R, Mali P, Moosburner M, Church GM. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nature methods. 2015. PubMed PMID: 26167643.
75. Ma J, et al. CRISPR-DO for genome-wide CRISPR design and optimization. Bioinformatics. 2016;32:3336–3338. PubMed PMID: 27402906.
76. Housden BE, Lin S, Perrimon N. Cas9-based genome editing in Drosophila. Methods in enzymology. 2014;546:415–439. PubMed PMID: 25398351.
77. Kaur K, Gupta AK, Rajput A, Kumar M. ge-CRISPR - An integrated pipeline for the prediction and analysis of sgRNAs genome editing efficiency for CRISPR/Cas system. Scientific reports. 2016;6:30870. PubMed PMID: 27581337.
78. Zhu H, Misel L, Graham M, Robinson ML, Liang C. CT-Finder: A Web Service for CRISPR Optimal Target Prediction and Visualization. Scientific reports. 2016;6:25516. PubMed PMID: 27210050.
79. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic acids research. 2016;44:W272-276. PubMed PMID: 27185894.
80. Oliveros JC, et al. Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic acids research. 2016;44:W267-271. PubMed PMID: 27166368.
81. Park J, Kim JS, Bae S. Cas-Database: web-based genome-wide guide RNA library design for gene knockout screens using CRISPR-Cas9. Bioinformatics. 2016;32:2017–2023. PubMed PMID: 27153724.
82. Heigwer F, et al. CRISPR library designer (CLD): software for multispecies design of single guide RNA libraries. Genome biology. 2016;17:55. PubMed PMID: 27013184.
83. Wong N, Liu W, Wang X. WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome biology. 2015;16:218. PubMed PMID: 26521937.
84. Moreno-Mateos MA, et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nature methods. 2015;12:982–988. PubMed PMID: 26322839.

85. Liu H, Wei Z, Dominguez A, Li Y, Wang X, Qi LS. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. Bioinformatics. 2015;31:3676–3678. PubMed PMID: 26209430.
86. MacPherson CR, Scherf A. Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. Nature biotechnology. 2015;33:805–806. PubMed PMID: 26121414.
87. Hodgkins A, et al. WGE: A CRISPR database for genome engineering. Bioinformatics. 2015. PubMed PMID: 25979474.
88. Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. Nature methods. 2014;11:122–123. PubMed PMID: 24481216.
89. Rastogi A, Murik O, Bowler C, Tirichine L. PhytoCRISP-Ex: a web-based and stand-alone application to find specific target sequences for CRISPR/CAS editing. BMC bioinformatics. 2016;17:261. PubMed PMID: 27363443.
90. Lei Y, Lu L, Liu HY, Li S, Xing F, Chen LL. CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. Molecular plant. 2014;7:1494–1496. PubMed PMID: 24719468.
91. Pliatsika V, Rigoutsos I. "Off-Spotter": very fast and exhaustive enumeration of genomic lookalikes for designing CRISPR/Cas guide RNAs. Biology direct. 2015;10:4. PubMed PMID: 25630343.
92. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics. 2015;31:1120–1123. PubMed PMID: 25414360.
93. Zhu LJ, Holmes BR, Aronin N, Brodsky MH. CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. PloS one. 2014;9:e108424. PubMed PMID: 25247697.
94. Xie S, Shen B, Zhang C, Huang X, Zhang Y. sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PloS one. 2014;9:e100448. PubMed PMID: 24956386.
95. Doench JG, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nature biotechnology. 2014;32:1262–1267. PubMed PMID: 25184501.
96. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nature biotechnology. 2014;32:267–273. PubMed PMID: 24535568.
97. Shalem O, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343:84–87. PubMed PMID: 24336571.
98. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nature methods. 2014;11:783–784. PubMed PMID: 25075903.
99. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014;343:80–84. PubMed PMID: 24336569.
100. Hart T, et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell. 2015;163:1515–1526. PubMed PMID: 26627737.
101. Horlbeck MA, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. eLife. 2016;5 PubMed PMID: 27661255.
102. Liu SJ, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science. 2017;355 PubMed PMID: 27980086.
103. Erard N, Knott SRV, Hannon GJ. A CRISPR Resource for Individual, Combinatorial, or Multiplexed Gene Knockout. *Molecular cell* **67**, 348-354 e343 (2017).
104. Metzakopian E, et al. Enhancing the genome editing toolbox: genome wide CRISPR arrayed libraries. Scientific reports. 2017;7:2244. PubMed PMID: 28533524.
105. Kleinstiver BP, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015. PubMed PMID: 26098369.
106. Anders C, Bargsten K, Jinek M. Structural Plasticity of PAM Recognition by Engineered Variants of the RNA-Guided Endonuclease Cas9. Molecular cell. 2016;61:895–902. PubMed PMID: 26990992.
107. Ran FA, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015;520:186–191. PubMed PMID: 25830891.

108. Hou Z, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences of the United States of America. 2013;110:15644–15649. PubMed PMID: 23940360.
109. Muller M, *et al.* Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Molecular therapy : the journal of the American Society of Gene Therapy* **24**, 636-644 (2016).
110. Karvelis T, et al. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. Genome biology. 2015;16:253. PubMed PMID: 26585795.
111. Yang H, Gao P, Rajashankar KR, Patel DJ. PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease. *Cell* **167**, 1814-1828 e1812 (2016).

Suggested Reading

Iiizumi et al, Simple one-week method to construct gene-targeting vectors: application to production of human knockout cell lines, Biotechniques 41:311-316 2006

Jantz and Berg, Probing the DNA-binding affinity and specificity of designed zinc finger proteins Biophys J 98:852-860, 2010

Klug, Towards therapeutic applications of engineered zinc finger proteins FEBS Letters 579:892-894 2005

Perez-Pinera et al, Advances in targeted genome editing, Curr Opin Chem Biol 16:268-277, 2012

Suggested Websites

www.addgene.com

www.agilent.com

www.cellectis.com

www.genecopoeia.com

www.horizondiscovery.com

www.idtdna.com

www.lifetechnologies.com

www.moldiag.in

www.sigma-aldrich.com

www.synthego.com

Inhibition of Protein-Protein Interactions: Cell-Based Assays

Mark Wade, Ph.D.,[1] Jacqui Méndez, M.S.,[2] Nathan P. Coussens, Ph.D.,[3] Michelle R. Arkin, Ph.D,[4] and Marcie A. Glicksman, Ph.D.[5]

Created: November 20, 2017.

Abstract

There is a strong interest in discovering compounds that inhibit protein-protein interactions. High-throughput screening (HTS) approaches include formats using purified proteins (see AGM chapter Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats) (1) and those using whole cells. This chapter describes two types of cell-based HTS assays, energy transfer (Förster resonance energy transfer and bioluminescence resonance energy transfer) and protein complementation (fluorescence or enzymatic, e.g. luciferase).

Introduction

Protein-protein interactions (PPI) are a diverse and challenging group of targets for small-molecule discovery efforts (2). PPI interfaces come in a wide variety of sizes (from 4 amino acids to thousands of $Å^2$) and vary greatly in their binding affinities, dynamics, and complexity (from two proteins to tens of proteins in a complex). In order to capture the complexity of the PPI environment and maintain the integrity of the PPI complex, several groups have developed assays to monitor protein complexes in cells.

Cellular PPI assays offer different advantages and disadvantages when compared to biochemical approaches using purified proteins *in vitro*. Biochemical formats offer the advantage that the targets are studied in isolation, and hits from such screens are likely to interact directly with one of the proteins (see AGM chapter Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats) (1). However, the cellular context is lost in a protein-based assay. In cells, proteins are in their native environment, including discrete subcellular locations and formation of multi-protein complexes. This complexity provides additional mechanisms of interaction for a small molecule, and ensures that active compounds are cell permeable and able to reach the PPI target.

This chapter describes the development of energy transfer and protein complementation screening assays for identifying small molecules that modulate PPI in cells. General introductions to assay development for HTS, cell culture for HTS, and high content imaging can be found in dedicated chapters in this *Assay Guidance Manual*.

Background

Energy transfer assays include Förster resonance energy transfer (FRET) and bioluminescence resonance energy transfer (BRET). In a FRET assay, the fluorescence energy from the donor protein is transferred to the acceptor protein, which then emits light; in BRET, luciferase-induced chemiluminescence provides the donor energy that is then transferred to a fluorescent protein or dye. For energy transfer assays, donor and acceptor proteins are fused to each partner of a PPI. Formation of the PPI complex brings the two proteins close enough to undergo energy transfer.

Bimolecular protein complementation assays utilize a reporter protein that has been genetically split into two chains that do not fold into active protein on their own. Each chain is fused to one of the target protein partners; when the protein partners come into contact, the split reporter protein is reconstituted and its fluorescence or

Author Affiliations: 1 Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia; Email: Mark.Wade@iit.it. 2 Promega Corp.; Email: Jacqui.mendez@promega.com. 3 National Center for Advancing Translational Sciences, National Institutes of Health; Email: coussensn@mail.nih.gov. 4 University of California, San Francisco; Email: Michelle.arkin@ucsf.edu. 5 Orig3n Inc.; Email: Marcie@orig3n.com.

enzymatic activity is restored. Protein fragment complementation assays fall into 'indirect' and 'direct' reporter classes. For direct reporters, complementation of the two fragments leads to reconstitution of reporter function – such as fluorescence (for example, GFP) or enzymatic activity (such as luciferase). Indirect reporters trigger additional events that lead to the assay readout, such as transcription via the 2-hybrid assay (3-6). The direct reporter system provides a more proximal measurement of a PPI, and is therefore generally preferred for the discovery of PPI inhibitors.

Table 1 contrasts the energy-transfer and protein-complementation formats. All methods rely on generating fusion proteins with the reporter pairs. The signals that derive from energy transfer and protein complementation require that the two proteins of interest be in molecular contact. These assay formats are distinct from co-localization assays that are used to demonstrate that two proteins are in the same subcellular region.

The selection of a cell-based energy transfer versus protein complementation assay depends on several factors. Due to its relatively low signal and narrow dynamic range, FRET is generally read through an imaging-based assay or flow cytometry; by contrast, enzyme complementation and BRET assays utilize a standard multi-modal plate reader. An important distinction among formats is whether they are reversible. Upon reading out a dynamic process, it is highly advantageous for the reporter signal to also be reversible. FRET, BRET and bimolecular luminescence protein complementation (BiLC) are reversible (7-12), whereas bimolecular fluorescent protein complementation (BiFC), β-galactosidase complementation, and β-lactamase complementation are generally not reversible (6, 13-23).

Table 1: Cell-based PPI assays described in this chapter

Format	Notable features
Förster Resonance Energy Transfer (FRET)	**Read out: high content imaging or flow cytometry** **Benefits: reversibility, suited to live-cell assay** **Limitations: low signal/dynamic range, photobleaching, need to correct spectral overlap**
Bioluminescence resonance energy transfer (BRET, NanoBRET)	**Read out: plate reader with luminescence detection** **Benefits: reversibility, high signal/low background, suited to live-cell assay, cell-permeable chemiluminescent substrates and fluorescent tags available** **Limitations: Older configurations of BRET have too much spectral overlap resulting in narrow dynamic range**
Bimolecular fluorescent protein complementation (BiFC)	**Generally not reversible**
Bimolecular luminescence protein complementation (BiLC)	**Read out: plate reader with luminescence detection** **Benefits: reversibility, high signal/low background, suited to live-cell assay, cell-permeable chemiluminescent substrates available** **Limitations: signals vary depending on fusion constructs and selection of luciferase species/ substrate; compounds can inhibit luciferase directly**
Other enzyme protein complementation (β-galactosidase, β -lactamase)	**Read out: plate reader with fluorescence detection** **Benefits: high signal/low background, suited to live-cell assay, cell-permeable substrates available** **Limitations: generally not reversible**

General Considerations for Developing Cell-Based PPI Assays

Yeast versus mammalian cells: Methods to measure PPI in mammalian cell-based assays are robust and in most cases provide sufficient throughput to serve as primary screening formats. We therefore recommend using cells of the same species as the target PPI. *Saccharomyces cerevisiae* (yeast) has been extensively used to discover protein-protein interactions and has also been used to discover the mechanisms of action and off-target effects of

bioactive compounds (24). Several of the assays described below can be used in yeast cells. However, there are some drawbacks to yeast expression of mammalian proteins. First, yeast lacks some proteins that might be required to form multi-protein complexes and might not produce the posttranslational modifications required for some mammalian PPIs. Second, accumulation of some test compounds in yeast is prevented by the expression of multidrug transporters, though yeast strains with increased drug sensitivity are available (25).

Cell type considerations: Any transfectable cell line can be used, but there is an increasing trend towards using cell lines and primary cells that best recapitulate the native biology. For instance, primary astrocytes have been stably transfected and used in a high-throughput screen to detect neuronal cells (26). An important factor for run-to-run reproducibility is to maintain a consistent passage number for cells. We highly recommend that a screen-sized batch of cells are grown and frozen in single-use aliquots prior to initiating the screen.

Transient versus stable transfection: If at all possible, stable transfection is highly preferable to transient transfection. First, stable transfection increases run-to-run reproducibility, which is critical for high-throughput screening (HTS). Second, expression levels can be selected based on resistance markers or FACS-based selection. Higher expression levels are often obtained through transient transfection, which can increase signal strength but might decrease physiological relevance.

Orientation of the protein fusions: In protein-complementation (BiFC and BiLC) assays, steric hindrance can sometimes prevent reconstitution/proper folding of the split protein. For example, GFP N- and C-terminal regions might not be oriented correctly upon protein-protein interaction, and thus fluorescence will not be observed. FRET and BRET efficiencies depend on distance and angle (see **Equations 2 and 3** in the next section), so orientation can also affect energy-transfer efficiency. Therefore, fusions to the PPI partners should be tested in all orientations and combinations to determine the optimal pair. In some cases, the biology of the PPI under study will limit the configuration options. For example, fusion of polypeptides to the C-terminus of the *r*eally *i*nteresting *n*ew *g*ene (RING) domain can inhibit biological function. Additionally, placing a linker (such as a serine/glycine sequence) between the proteins of interest and each half of the split protein is recommended to reduce steric hindrance. Including an epitope tag in this linker region can also be useful for quantifying the relative amount of each fusion partner present in the cell.

Full-length versus fragments of the proteins of interest: There are biological and practical considerations to consider. The use of full-length proteins is recommended when possible, because additional domains remote from the primary site of interaction can sometimes contribute to binding. However, steric effects or alterations in the stability, expression level, or localization of the proteins are sometimes prohibitive. In this case, fragments rather than full-length proteins might be required for interrogating some PPIs.

Making the system inducible: Consistent expression of the PPI is preferable if the protein pair is constitutively expressed in the native case. On the other hand, there are potential advantages for regulated expression of the fusion proteins. First, if the system is under a doxycycline-regulated promoter (Dox-on), graded expression of the protein can be achieved by titrating in the amount of inducer. This aids assay development, as the minimal amount of fusion protein required to generate sufficient signal can be determined. Second, an inducible system allows cells to be pre-incubated with compounds before expression of the interacting partners. This might increase the number of compounds identified in the primary screen if the PPI is essentially irreversible.

To lyse or not to lyse: Sometimes the term 'cell-based' PPI assay is used when the cells are lysed before the complex is measured. *An important issue with this assay format is that protein-protein and protein-small molecule interactions will re-equilibrate during lysis and dilution.* These assays might therefore be more precisely defined as 'lysate-based' PPI assays. There are situations in which a lysate-based assay is beneficial. For instance, when the PPI is difficult to express in a native form, one could express the interacting protein partners in the cell and prepare lysates before performing the screen. In this way, one can also perform large-scale preparations from cells grown in bulk and generate assay-ready lysate that can be stored. It is also noteworthy that compounds that

are not cell-permeable can still be identified through a lysate-based assay. However, if the goal is to screen for PPI inhibitors in whole cells, a non-lytic format should be used.

Demonstrate that the PPI under study drives the complementation and not vice-versa: For fragment complementation assays, it is important to demonstrate that the affinity of the PPI of interest is not significantly altered by the addition of the protein fragments. Control constructs containing point mutation(s) that destroy the PPI can be used for this purpose. If available, a known small molecule inhibitor should be used as a positive control. As with any type of screen, the controls for specificity and selectivity of 'hits' identified in protein fragment complementation screens are critical.

Statistical Considerations: The Z′-factor should be used to evaluate the quality or performance of the optimized assay prior to implementing HTS (27). The Z′-factor takes into account the measurement variability for the high-signal (c^+) and low-signal controls (c^-) in addition to the overall dynamic range of the assay signal as shown below where σ_{c+} μ_{c+} and σ_{c-} μ_{c-} are the standard deviation and mean values for the high-signal and low-signal controls, respectively (**Equation 1**):

$$Z'\ factor = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|}$$

An assay with a Z′-factor value between 0.5–1 is considered to be robust with low assay variability and suitable for HTS. Variability within the plate, between plates, and between days needs to be established. Hits should be defined by a threshold of three standard deviations from the baseline and the percent coefficient of variation (in the absence of a positive control) should be less than 10%. For additional information on assay validation see the AGM Chapter HTS Assay Validation.

Assay artifacts and interferences: Prior to selecting an assay for optimization, it is important to carefully consider all potential sources of artifacts and interferences. In comparing multiple assay approaches for high-throughput screening, the primary assay should be selected to minimize susceptibility to interferences, provided that the assay is robust, feasible and of suitable quality for screening. Counter screens can be developed to identify compounds that interfere with the detection method or have intractable mechanisms of action. Additionally, orthogonal assays can be optimized that are susceptible to different types of artifacts compared with the primary assay, and also increase the evidence for on-target activities of hit compounds (see AGM chapters Interferences with Luciferase Reporter Enzymes and Interference with Fluorescence and Absorbance) (28, 29). Depending on the design and readout, assays have different susceptibilities to artifacts and interferences. The methodologies described in this chapter rely on the measurement of fluorescence and luminescence (luciferase) light as a readout. Test compounds in chemical libraries can interfere with the detection measurement due to autofluorescence properties, scattering of light, or by absorbing light at the excitation or emission wavelengths (inner filter effect) (28, 29). Fluorescence spectroscopic profiling studies of large compound libraries indicated that interference is most prominent in the UV and blue range, whereas it is substantially reduced for red-shifted readouts (30, 31). Therefore one can reduce library compound interferences by simply selecting the most red-shifted detection reagents that are available and applicable. Luciferase assays are susceptible to interference by colored compounds that can quench luminescence as well as enzymatic inhibitors (see AGM chapters Interferences with Luciferase Reporter Enzymes and Interference with Fluorescence and Absorbance) (28-30, 32). It is critical to ensure that the signals being measured are in the linear range of the detection instrumentation and that the appropriate hardware (such as filters and mirrors) are in place. While some interference mechanisms are assay or readout dependent, other mechanisms such as chemical reactivity (see AGM chapter Assay Interference by Chemical Reactivity) (33) and aggregation (see AGM chapter Assay Interference by Aggregation) (34) can affect a wide range of methodologies and should be tested for by appropriate assays during the hit triage process.

Förster Resonance Energy Transfer (FRET)

Introduction

Förster Resonance Energy Transfer (FRET) occurs when energy from an excited fluorophore (donor) is absorbed by another molecule (acceptor). FRET is a powerful method for measuring distances between two fluorophores and is used in unimolecular systems (e.g., the donor and acceptor on the same protein) and biomolecular systems (e.g., the donor and acceptor on two interacting proteins; **Figure 1**).

The underlying principles of FRET were introduced over fifty years ago (reviewed in (35)). Cell-based FRET was first used in flow cytometry (11), but it is also an established format for high content assays and plate-reader formats.

For a FRET assay, the donor and acceptor molecules must be in close proximity to one another (typically 10-100 Å). In addition, there must be overlap between the fluorescence emission spectrum of the donor and the excitation spectrum (usually the same as the absorption spectrum) of the acceptor (**Figure 2**). The extent of overlap is referred to as the spectral overlap integral (*J*). Förster demonstrated that the efficiency (*E*) of the energy transfer depends on the inverse sixth-distance between donor and acceptor (36) (**Equation 2**):

$$E = \frac{Ro^6}{\left(Ro^6 + r^6\right)}$$

Where *Ro* is the Förster distance at which half the energy is transferred to the acceptor and *r* is the distance between donor and acceptor. Förster distance (*Ro*) is dependent on a number of factors, including the fluorescence quantum yield of the donor in the absence of acceptor ($_d$), the refractive index of the solution (η), the dipole angular orientation of each molecule (k^2), and the spectral overlap integral of the donor and acceptor (*J*) (**Equation 3**):

$$Ro = 9.78 \times 10^3 \times \left(\eta^{-4} \times f_d \times k^2 \times J\right)^{\frac{1}{6}} \times \mathring{A}$$

The *Ro* distance is between 30-50 Å for the FRET pairs in Table 2. While this range could be important for large complexes, FRET pairs are generally chosen based on spectral characteristics rather than Ro. Energy transfer leads to a reduction in donor fluorescence and an increase in acceptor fluorescence, and is usually reported as the ratio donor emission/acceptor emission. Cell-based FRET generally uses fluorescent proteins, which have broad excitation and emission peaks and small Stokes shift (the distance between the emission and excitation wavelengths). It is therefore important to carefully select the donor and acceptor fluorescent proteins to maximize the distance between fluorescence wavelengths of donor and acceptor, and use narrow bandpass filters to avoid counting the overlap between the donor and acceptor (called cross-talk). Crosstalk should be measured and corrected (see below).

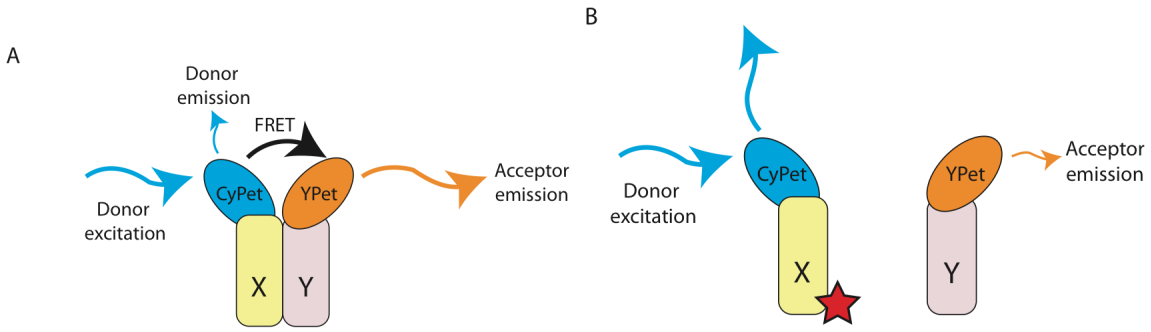


Figure 1. Schematic representation of FRET between two proteins (X and Y) fused to donor- and acceptor fluorescent proteins (CyPet and YPet, respectively). When excited, CyPet emits light, unless that energy is transferred to the YPet acceptor molecule. The YPet then emits light at its characteristic wavelength. If the PPI is inhibited (e.g., by a small molecule, red star) the CyPet donor and YPet acceptor are no longer in proximity, and FRET is reduced.

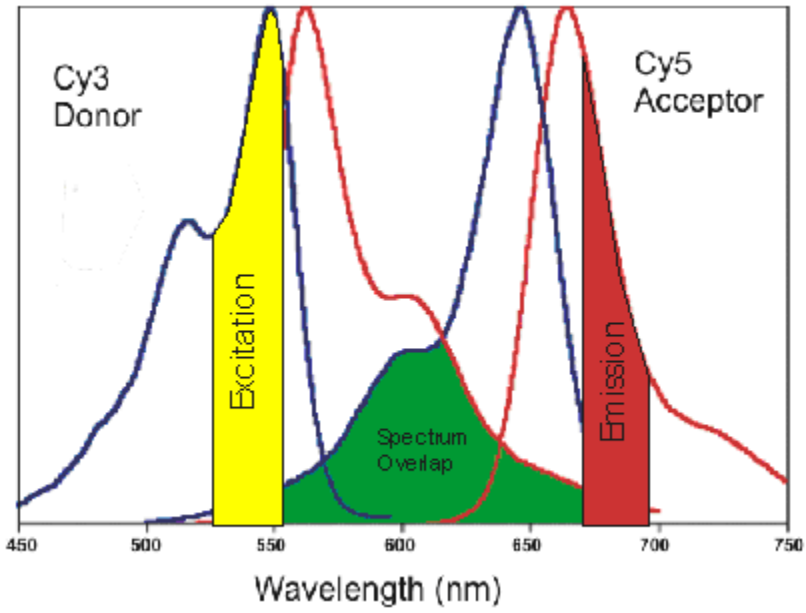


Figure 2. Spectral overlap and FRET. The donor (Cy3) is excited at a maximal wavelength of 550 nm, and emits with light at 650 nm. The acceptor (Cy5) absorbs light at 570 nm and emits at 670 nm. Energy transfer occurs due to the spectral overlap between the emission of the donor and excitation spectrum of the acceptor (shown in green). Crosstalk occurs when donor emission overlaps with acceptor emission, such that the donor emission is counted as FRET; similarly, the acceptor can be directly excited during donor excitation if their absorbance spectra overlap (as they do here).

Table 2. Common Donor/Acceptor pairs for intracellular FRET

Donor (ex/em)	Acceptor (ex/em)	Reference
Sirius (355/424)	mseCFP (434/474)	(62)
mTagBFP (402/457)	sfGFP (485/510)	(63)
CFP (433/475)	eYFP (514/527)	
CFP (433/475)	Venus (515/528)	
mTurquoise2 (434/474)	Venus (515/528)	(64)
CFP (433/475)	mCitrine (516/529)	

Table 2. continued from previous page.

Donor (ex/em)	Acceptor (ex/em)	Reference
mTFP1 (462/492)	mCitrine (516/529)	(65)
eCFP (439/476)	YPet (517/530)	(66)
CyPet (435/477)	YPet (517/530)	
Venus (515/528)	mKOκ (551/563)	(67)
mAmetrine (406/526)	tdTomato (554/581)	(65)
Sapphire (399/511)	DsRed (558/583)	(62)
CFP (433/475)	mRFP (584/607)	(68)
CFP (433/475)	mCherry (587/610)	(64)
eYFP (514/527)	mCherry (587/610)	(64)
Venus (515/528)	mCherry (587/610)	
mOrange (548/562)	mCherry (587/610)	(66)
mKOκ (551/563)	mLumin (587/621)	(67)

Assay Development

Instrumentation**:** For imaging, there are a number of instruments well suited for high-throughput imaging assays (e.g. Molecular Devices' Image Express, GE Healthcare InCell Analyzer, PerkinElmer Operetta). Many commercial plate readers can be used to detect the signal from FRET, and are appropriate for lysate-based assays. These include TECAN's Safire (now TECAN M1000), BMG's PherastarFS, BioTek's Synergy2 and -4, PerkinElmer's Envision.

Plates: Black solid or clear bottom plates are compatible with FRET. For imaging assays, there are plates optimized for microscopy with flat bottom, uniform, and thin plastic wells to minimize distortion (e.g., from Aurora Biotechnologies, Costar, and Greiner).

Cell types**:** FRET has been performed in many cell types including U2OS, HEK293, HeLa, COS-7, CHO, HCT-116. Selecting the most relevant cell type generally leads to identification of more relevant hits.

Buffers**:** For cell-based HTS in intact cells, phenol red-free medium should be used. The medium should be buffered against changes in pH by the addition of HEPES, for example.

Choice of FRET pair: Select the appropriate fluorescent protein pair to use in FRET (see Table 2). Ideal pairs have maximal overlap of donor emission and acceptor excitation, but minimal direct excitation of the acceptor at the excitation maximum of the donor. Historically, CFP/YFP has been the choice for FRET. However, with the continual expansion of the fluorescent protein palette, a large number of additional pairs are now available. Commonly used pairs include modified GFP/Venus (37, 38) and CyPet/YPet (12). Many of the newer pairs contain red-shifted acceptors, as this provides the opportunity for multiplexing with green-shifted acceptors in order to study more than one interaction simultaneously. Red-shifted variants are also advantageous as they can reduce the noise generated by cellular autofluorescence, which is generally in the green range. However, caution should be taken when choosing earlier versions of red-shifted proteins (such as DsRed or mRFP), as they generally have lower quantum yields than green fluorescent proteins.

For measuring PPI on the cell membrane, one can use donor and acceptor labeled antibodies raised against the proteins of interest. This method has been successfully used to detect changes in the interaction between the ERBB1/ERBB2 heterodimer, for example. (39)

Assay Conditions**:** The ratio of the donor and acceptor is critical for accurate FRET. This can be tricky for bimolecular FRET, where two fusion proteins are expressed at different levels or have different stabilities. This might lead to an excess of emission donors that do participate in the PPI, which will in turn mask the true FRET signal. An excess of acceptor can also complicate interpretation of the FRET signal, particularly if the acceptor is susceptible to direct excitation. In addition to expression level, other factors such as pH and temperature (in the case of live cell FRET) can interfere with the assay readout because they alter the rate of acceptor or donor maturation, and thus affect their concentrations.

Data Analysis

The type of analysis performed is linked to the precise FRET methodology used in the screen. In the simplest version, the ratio between acceptor and donor emission following donor excitation is calculated before and after compound addition. In order to overcome the issue of spectral cross-talk (in which the acceptor is excited directly by the donor excitation wavelength, or in which donor emission is present in the acceptor channel), filter cubes with bandpasses very specific for each FRET pair should be used. Ideally, several additional control cell lines should also be used in order to obtain a more accurate measurement of FRET efficiency. These include untransfected cells, and cells that express either the donor or the acceptor FRET partner. Measuring the fluorescence emission in these cells following excitation at donor and acceptor wavelengths provides information that can be used to calibrate the system. FRET is a powerful method but it does require careful analysis of images including measurement of bleed-through from each signal and normalization if the expression levels are different.

Alternatively, where ratio is not 1:1, the acceptor photobleaching (donor dequenching) FRET method might be considered, as the FRET efficiency calculated using this approach is independent of the ratio (see (39) for an example of FRET calculated using this method). In terms of HTS, however, this technique is relatively low throughput, as it generally requires laser scanning microscopy and high magnification, and bleaching times can be in the order of minutes for some photostable proteins.

Fluorescence lifetime microscopy (FLIM) is a third alternative to the measurement of FRET. This technique is independent of the ratio between donor and acceptor, as it only measures changes in donor fluorescence lifetime following illumination. Specifically, participation of the donor in FRET reactions with the acceptor alters the donor fluorescence lifetime, which can then be used to estimate changes in FRET efficiency in the presence and absence of compounds. However, conventional FLIM is unsuitable for HTS as its measurement is time-consuming and requires sophisticated instrumentation. Recently, FLIM microplate readers compatible with SBS standard labware have been introduced on the market (for example, see http://www.fluorescenceinnovations.com/cells.html). Furthermore, the throughput of FLIM has been improved, to give 96 well plate reads of tens of minutes, rather than hours (40). Because spectral bleed-through is not an issue with this technique, additional control cell lines expressing individual FRET proteins are not required. However, quantification of FLIM does require a series of calculations following image segmentation (for examples see (40) and (41)).

Assay Validation Steps

1. Test which orientation of the donor and acceptor fluorescent protein fusions produces the most robust FRET signal (depending on the protein of interest, validation of correct subcellular localization or biochemical activity might also be prudent). Also determine whether full-length proteins or only the interaction domains should be used for the FRET constructs. Inclusion of an epitope tag in the linker between the donor and acceptor fluorescent proteins and the proteins of interest can help determine relative ratio. If one member of the pair is more highly expressed, this protein should be selected as the acceptor because it reduces the potential of free donors contributing to spectral cross talk.

2. In the absence of small molecule positive controls, create point mutants or deletion constructs that disrupt the interaction under investigation.
3. Generate cells in which neither, or only one of the FRET pair is expressed in order to establish the degree of spectral cross-talk and bleed-through from each fluorophore.
4. Select specific filters based on the FRET pair chosen (i.e., it is not sufficient to use a general 'YFP' filter for all of the new yellow acceptor fluorescent protein variants). Use of dual band excitation/emission filters to allow simultaneous read of donor and acceptor emission can reduce variability during plate reads.
5. Optimize the illumination conditions for FRET: low illumination might result in poor signal to noise, whereas strong illumination can lead to photobleaching, both of which will complicate ratiometric analysis.
6. For automated image acquisition, minimize the exposure time for auto-focusing in order to avoid photobleaching effects.

Statistical Considerations for FRET

Enough cells should be prepared in bulk to run the full screen to avoid passage-number variation. Batch-to-batch (day-to-day) variations can also be significant with FRET. Inclusion of the same control on all plates within and between batches is recommended. The same statistics that have been described for other HTS methods apply to FRET-based assays, See the AGM chapter (see the AGM Chapter HTS Assay Validation) (42). See the overall statistical considerations section at the beginning of this chapter.

Sample Validation Data

Example 1. Development of FRET Assay into Quantitative and High-throughput Screening Technology Platforms for Protein–Protein Interactions by Song and colleagues (12). This manuscript describes development of a FRET-based assay for measuring the signaling cascade of SUMOylation using the FRET pair CyPet and YPet. Assay development and a pilot high-throughput screen in 384-well plates is described. Compounds are incubated with the stably transfected HEK293 cells for two hours, before reading the signal in a Molecular Devices FlexStation reader.

Example 2. Förster resonance energy transfer (FRET)-based subcellular visualization of pathogen-induced host receptor signaling by Buntru and colleagues (43). This study measures a protein-protein interaction in intact live cells using flow cytometry measurements. The authors further validated their assay with confocal microscopy and acceptor photobleaching to localize the sites of bacteria-host cell contact. This assay could be adopted for high-throughput screening.

Troubleshooting Guide

No or low FRET: The absence of energy transfer might be related to the actual FRET donor and acceptor constructs. Steric issues might prevent optimal FRET, while some combinations of fluorescent proteins and interacting proteins might be inherently unstable. Check expression of all orientations of N- and C-terminal fusions of fluorescent proteins for both donor and acceptor during assay development. An excess of free donor can reduce apparent FRET efficiency, and in this case one should titrate the FRET pairs during transfection optimization to favor an excess of acceptor. Suboptimal filter sets can significantly compromise FRET; thus the bandpass of the filters should be tailored to precisely match the absorption/emission spectra for the fluorescent proteins. If FRET is being estimated ratiometrically, care must be taken to avoid photobleaching. Reduce illumination intensity using neutral density filters, and illuminate only donor when scanning plates to identify cells to analyze. In order to maximize FRET efficiency, choose partners with minimal excitation of acceptor at maximal excitation of donor, and maximal overlap of donor emission and acceptor excitation.

Low signal-to-noise: Low signal-to-noise can be due to a low FRET signal (see above) or due to fluctuations in the light source. It is critical that the light source is stable for this assay. Another trivial explanation for low signal-to-noise is the mislabeling of donor and acceptor images during analysis.

Ratio is very high or low: In microscopy-based FRET, outliers with very high or low FRET ratios can be due to analysis of out-of-focus images or differences in expression in transient tranfections or pooled transfectants. In this case, different ROIs or cells must be selected and re-analyzed.

High FRET with negative controls: The fluorescent proteins should incorporate mutations to prevent oligomerization (44). If the genetic controls using binding-deficient mutants give a high FRET signal, this might indicate that the fluorescent proteins are oligomerizing and driving a non-specific interaction. Alternatively the expression level of the FRET pair might be too high and in this case re-optimization of expression is required. Also one should examine the signal overlap between donors and acceptors. Another possibility is that the interaction under investigation is more complex than originally thought. The investigator should validate the sequences of plasmid DNAs to confirm the presence of disruptive mutations.

Bioluminescence Resonance Energy Transfer (BRET)

Introduction

Bioluminescence Resonance Energy Transfer (BRET) is also a proximity-based assay that measures the energy transfer from a donor to an acceptor, in this instance luciferase is used as the donor, and a fluorescent protein as the acceptor (10). The NanoBRET method employs a NanoLuc fusion protein as the bioluminescent donor and a fluorescently labeled HaloTag fusion protein as the acceptor (**Figure 3**). The optimized blue-shifted NanoLuc donor paired with the red-shifted HaloTag acceptor minimizes spectral overlap, resulting in low bleed-through and improved signal-to-background and dynamic range. The extremely bright and stable NanoLuc® also allows for expression of the donor at or near physiological levels further reducing background and improving sensitivity.

Assay Development

Instrumentation: To measure BRET assays, an instrument capable of sequentially measuring dual-filtered luminescence values equipped with appropriate filters is required. For the NanoBRET configuration the ideal filter setup will include a band pass (BP) filter centered around 460nm to measure the donor signal (Ex. Emission 450nm/BP 80nm) and a long pass (LP) filter starting at around 600–610nm to measure the acceptor signal (Ex. Emission 610nm/LP). Filters outside of these ranges will miss critical measurements and compromise data quality. Commercial plate readers pre-equipped with the proper filter set up include Promega's GloMax Discover and BMG's CLARIOstar. Instruments that can be equipped with optional filters include Thermo's Varioskan (Edmunds Optics filters: donor 450nm CWL, 25mm diameter, 80nm FWHM, Interference Filter and acceptor 1 inch diameter, RG-610 Long Pass Filter) and PerkinElmer's EnVision (Chroma filters: Emission Filter (for EmSlot4) Cat. # AT600LP and Second Emission Filter (for EmSlot1) Cat. #AT460/50m).

For instruments using mirrors, select the luminescence mirror. An integration time of 0.2–1 second is typically sufficient. Ensure that the gain on the photomultiplier tube is optimized to capture the highest donor signal without reaching instrument saturation.

Plates: White tissue culture grade plates are preferred. Black plates can quench luminescence. Formats include 96 and 384 wells. 1,536 well formats might require optimization and a specialized instrument.

Cell types: NanoBRET can be performed in any cell line that can be transfected including HEK293, HeLa, HCT-116, NIH3T3, CHO, and Jurkats. For difficult to transfect cells, optimization might be required.

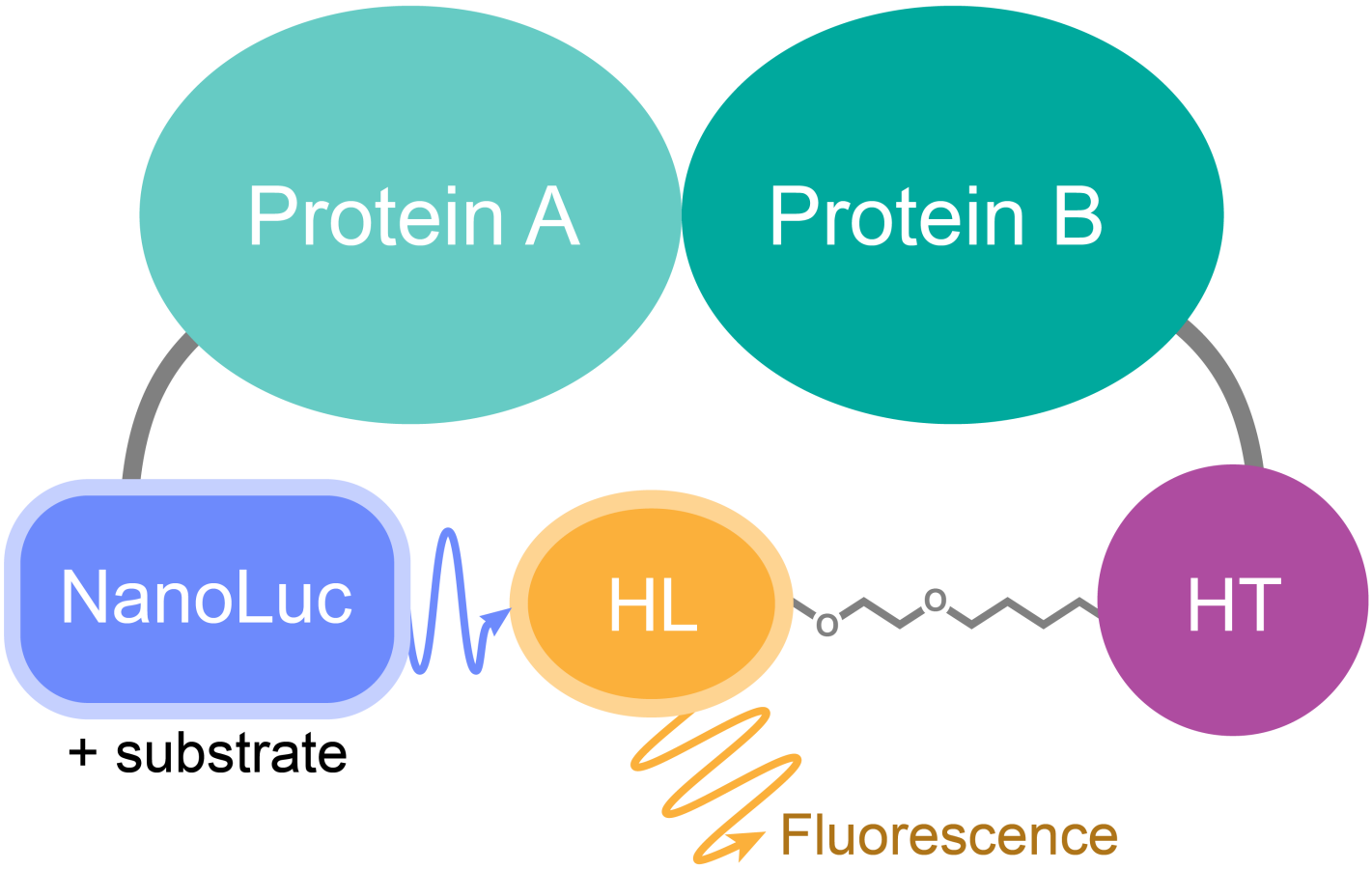


Figure 3. Overview of the NanoBRET assay principle. Energy is transfered from a NanoLuc-Protein A fusion (energy donor) to a fluorescently labeled HaloTag-Protein B fusion (energy acceptor) upon interaction of Protein A and Protein B.

Generation of stable cell lines expressing one or both protein partners is highly preferred, for the reasons described above.

Media: To avoid interference with the acceptor signal, phenol red-free medium should be used. Reduced serum media such as Opti-MEM + 4% FBS is recommended. The use of higher amounts of FBS can increase variability in 384 and 1,536 well formats.

Assay conditions: Find the optimal ratio of donor and acceptor expression levels to minimize unbound donor, which reduces background and maximizes dynamic range (see Assay Validation Steps below). Cells are transfected and grown under typical tissue culture conditions. Because the HT protein is not intrinsically fluorescent, a set of samples without ligand can be plated and used as a background control. The non-lytic assay is completed by the addition of the NanoLuc substrate furimazine and donor and acceptor signal measurements.

Kits and reagents: Kits for generating NanoLuc and HaloTag fusion clones and kits containing HaloTag ligand and NanoLuc substrate reagents are available from Promega.

Data Analysis

It is critical to use the appropriate filters to capture the specific donor and acceptor signals for NanoBRET, generic filters for other BRET configurations might not be applicable.

The donor emission occurs at 460nm, a band pass (BP) filter that covers close to 460nm with a band pass range of 8–80nm is preferred. A short pass (SP) filter that covers the 460nm area also can be used but it might result in an artificially large value for the donor signal measuring the bleed-through into the acceptor peak. This artificially large value could compress the ratio calculation and reduce the assay window. The acceptor emission occurs at 618nm. To measure the acceptor signal, a long pass filter starting at 600–610nm is preferred.

To calculate the BRET ratio the acceptor signal is divided by the donor signal (**Equation 4**):

$$\frac{618\ nm_{Em}}{460\ nm_{Em}} = Raw\ NanoBRET\ ratio = BU$$

To convert raw BRET units (typically decimal values) to milliBRET units (mBU; whole numbers), multiply each raw BRET value by 1,000 (**Equation 5**):

$$\frac{618\ nm_{Em}}{460\ nm_{Em}} = BU \times 1000 = mBU$$

To account for any background contribution, the mean ratio from the no ligand control samples is subtracted from the ligand containing samples to yield a corrected ratio (**Equation 6**):

$$Mean\ mBU(experimental) - Mean\ mBU(no\ ligand\ control) = Mean\ corrected\ mBU$$

Assay Validation Steps

1. Determine the donor and acceptor configuration with the optimal distance and geometry for efficient energy transfer between the protein partners. There might be instances when there is prior knowledge of the protein biology that would prevent a certain terminus from being tagged. If no such constrains exist, both proteins of interest can be tagged with either NanoLuc or HaloTag on either the N- or C-terminus resulting in up to 8 possible constructs and 8 possible combinations. Full-length proteins or protein domains might be used. Test the various combinations in the assay to see which one(s) give the most robust ratio(s).
2. Optimize transfection conditions for the top pair configuration(s). This is done by finding the ratio of donor to acceptor DNA that minimizes unbound donor, reducing the background and maximizing dynamic range. Generally, the amount of HaloTag DNA is kept at a higher concentration while the NanoLuc DNA is reduced to $1/10^{th}$ to $1/1{,}000^{th}$ relative to the amount of HaloTag DNA.
3. If a known modulator of the interaction is available, either an inhibitor or an activator, confirm the proper biological response. In some cases the highest possible ratio pair might not always equate the most robust response to the modulator. It is advisable to test more than one combination to find the most responsive to the expected biology. In addition, some small molecules might have been developed against a particular isolated domain or using *in vitro* assays and might not elicit the same response when using full-length proteins or inside the cell.
4. If no known modulator is available, donor saturations assays (DSA) might be performed to show assay specificity (8). In a DSA, the amount of NanoLuc donor DNA is kept constant while steadily increasing the amount of acceptor DNA. As the acceptor-to-donor (A/D) ratio increases, a specific BRET assay will show ratios that increase in a hyperbolic manner and reach a plateau representing complete saturation of all donors with acceptor molecules. Non-specific interaction, as it can be tested with a negative control protein, will generate much weaker ratios that plot in a linear manner known as bystander BRET (**Figure 4**).

Statistical Considerations for BRET

The ratiometric nature of the BRET assay intrinsically results in low variability. Use the corrected mBU and standard deviation values to calculate the Z'-factor (see **Equation 1**).

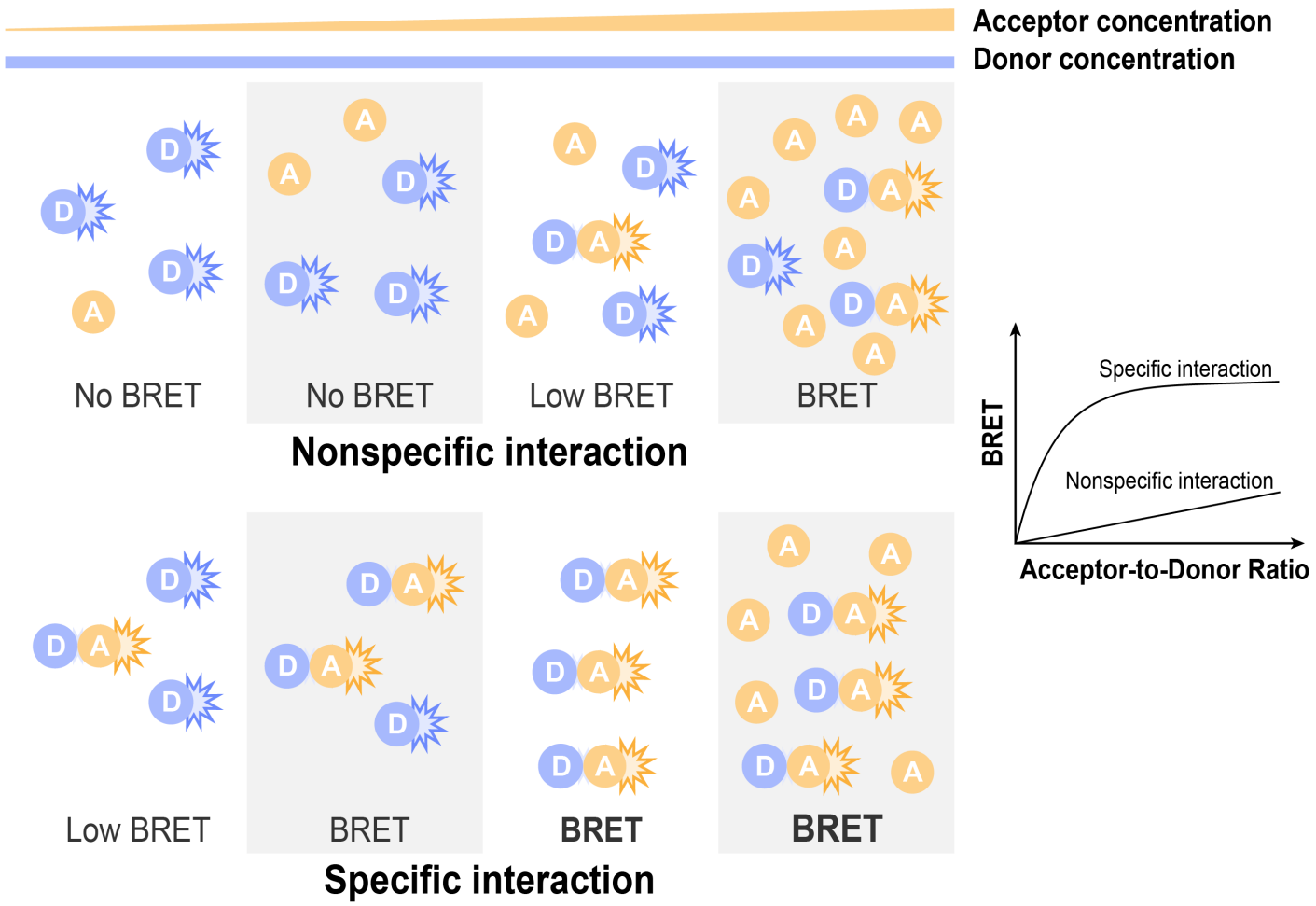


Figure 4. Principle of the Donor Saturation Assay (DSA) to validate BRET assay specificity.

Sample Validation Data

Example 1. NanoBRET—A Novel BRET Platform for the Analysis of Protein–Protein Interactions by Machleidt and colleagues (9). In this example, the authors describe the development of the NanoBRET configuration of BRET and its application to a novel assay developed for analyzing the interactions of bromodomain proteins with chromatin in living cells.

Example 2. Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy by Picaud and colleagues (45). The authors report the development and preclinical evaluation of a novel, potent inhibitor targeting CBP/p300 bromodomains that impairs aberrant self-renewal of leukemic cells. This is a good example of a small molecule that when studied in live cells, exhibits differential inhibitory profiles in full-length proteins vs. individual protein domains

Troubleshooting Guide

No or low BRET:

a. Ensure filters are specific for NanoBRET: 460nm (with 8–80nm band pass) for donor signal and 600–610nm long pass for acceptor signal. Make sure the gain on the photomultiplier tube is set to detect donor

signal without instrument saturation. Confirm ratio calculation is acceptor divided by donor values (618nm/460nm).
b. Test all possible donor/acceptor combinations to find best pair.
c. Ensure protein expression by checking luminescence for NanoLuc constructs. Expression for HaloTag constructs can be checked by fluorescently labeling with a HaloTag ligand such as TMR which can then be run on SDS-PAGE and analyzed using a fluorescence scanner.
d. Optimize transfection conditions and relative amounts of donor and acceptor DNA. Excess donor will compress assay window.
e. Some PPI might be dependent on specific biological events or stimuli to activate specific pathways. Ensure proper pathway activators are added. If possible, check proper phenotypical responses by other means.
f. Confirm cell health by a viability assay (see AGM chapter Cell Viability Assays).
g. Absolute raw values and ratios will vary among PPI systems. Absolute BRET values depend on the proximity of the protein partners, the affinity of the interaction, the relative occupancy with other interacting proteins and the instrument setup. When possible, check specificity with a known modulator such as an inhibitor or by the Donor Saturation Assay (DSA).

High variability: Optimize all parameters of the assay aiming for Z'-factors between 0.5 and 1. If the assay parameters have been optimized, consider automated dispensing to reduce variability.

Unable to detect proper expected biology in the presence of small molecule: Increase the concentration of small molecule, treat overnight or both to see maximum effect. If the compound was developed *in vitro* against the domain or region alone, it might not disrupt the interaction of the full-length proteins. Test region or domain alone versus full-length protein. Always consider the possibility that the compound might affect the luciferase activity without actually disrupting the PPI (see next section).

Fluorescent Protein Fragment Complementation

Introduction

Direct reporter complementation assays have split reporter proteins fused to the protein partners of interest; formation of the PPI causes reconstitution of the reporter protein and induction of fluorescence or enzymatic activity (**Figure 5**).

The use of split GFP as a PPI reporter was first described by Regan and colleagues (46). In this method, a GFP variant was separated between residues 157 and 158 into two polypeptides, each of which were fused to interacting leucine zippers. Reconstitution of fluorescence was not observed when the GFP fragments were co-expressed without the leucine zippers, indicating that PPI could be reliably detected using this method. GFP and additional fluorescent proteins were rapidly adopted for the study of PPIs. Hu and Kerppola extended the use of split GFP to measure protein-protein interactions in mammalian cells (16). This so-called '*Bi*molecular *F*luorescence *C*omplementation' (BiFC) has been extended to many other fluorescent proteins, greatly expanding the potential combinations that could be applied *in vivo*. BiFC offers some advantages over other fluorescence-based methods for PPI detection. For example, FRET requires careful control of the ratio between interacting partners, and is also less sensitive than BiFC due to the presence of background fluorescence emitted by the acceptor fluorophore (**Section 2**) (47). Many BiFC pairs have been demonstrated (**Table 3**).

Because the primary amino acid sequence of different fluorescent proteins varies at only a few residues, it is possible to 'multiplex' BiFC. For example, the C-terminal amino acid sequences of GFP, YFP, and BFP are identical after amino acid 155, whereas their N-terminal sequences differ at several residues. In this way, the interaction of one protein with 2 potential partners can be interrogated simultaneously. For example, interaction of the bJun transcription factor with its binding partners bFos and bATF2 has been demonstrated (16).

Despite its versatility, BiFC does suffer from some limitations. First, refolding of the fluorescent chromophore fragments is essentially irreversible (48, 49). While this might be advantageous for detecting weak or transient PPIs, the irreversible nature of BiFC in its present form greatly complicates its application as a discovery platform for inhibitors of PPIs. To get around this issue, one can engineer an inducible BiFC reporter (using a doxycycline-responsive promoter, for example); the reporter cell lines can then be incubated with compounds prior to expression of the BiFC constructs, providing a time window in which their activity can be detected.

A second issue with BiFC (and other complementation assays in general) is that the signal intensity, compared to that achieved with the full-length protein, is often orders of magnitude lower. Although the background of BiFC is low, the reduced signal can sometimes reduce the robustness of the HTS assay. Variants of Venus (a YFP derivative) have been reported to significantly improve the signal/background for complementation (13).

Finally, a fluorescence-maturation step follows re-association of the split chromophore. In some cases, this maturation is temperature-dependent; the BiFC signal from some YFP fragments and from mCherry is greatly enhanced if the temperature is lowered to 25-30 °C prior to readout (13, 17). This change in temperature presents an additional logistical consideration and potential source of variability for BiFC-based HTS campaigns.

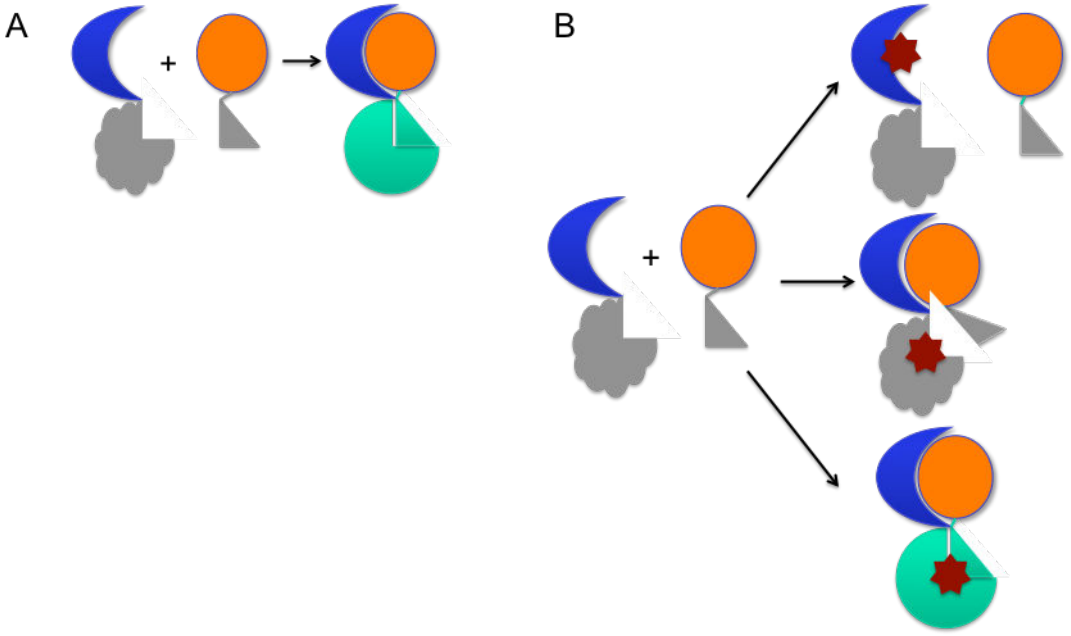


Figure 5: Principle of protein fragment complementation. (A) Two interacting proteins (blue) are fused to N- and C-terminal fragments of a reporter protein such as GFP or luciferase (gray). Following interaction, the reporter is reconstituted (green). (B) A small molecule (red star) can affect the reporter signal in several ways. The molecule might bind to and specifically inhibit the PPI under interrogation, a 'true positive' (top). Alternatively, false positive compounds could interfere with reconstitution of the reporter protein (middle) or block the catalytic activity of the reporter (bottom). Secondary screens are required to identify such false positive events.

Table 3. Examples of BiFC partners

Fluorescent Protein	N-terminus	C-terminus	Comments	Reference
GFP	1-157	158-238	Utilized a brighter variant of wtGFP	(69)
Venus	1-154	155-238	A brighter variant of YFP	(13)
Venus	1-210 (V150A)	210-238	Additional mutations added to Venus to improve S/N ratio	(14)
YFP	1-154	155-238	Increasingly replaced by other 'yellow' variants (Citrine and Venus) due to sensitivity to high temperatures	(15)
CFP	1-154	155-238		(16)
mCitrine	1-172	173-238		(70)
mCherry	1-159	160-237	Sensitive to high temperatures, requires longer maturation times	(17)
mLumin	1-151	152-233	Brighter mKate variant, less sensitive to high temperature compared to other red variants	(71)

Assay Development

Chromophore selection**:** **Table 3** lists common fluorescent proteins used in split constructions. Each reference contains additional variations in terms of the location of the split, and point mutations to improve fluorescence emission or stability in the BiFC application. Splitting chromoproteins can also slightly alter the maximum excitation and emission spectra compared to the full-length proteins, and wavelength scanning to identify the optimal excitation and emission is therefore recommended when using new BiFC pairs.

Instrumentation**:** BiFC can be measured on high content fluorescence microscopes (e.g., Molecular Devices Image Express and Image Express Ultra, GE Healthcare INCell Analysts 2200 or 6000, or the PerkinElmer

Operetta or Opera) or flow cytometry (e.g., BD Biosciences, Intellicyte). If the fluorescence intensity is sufficiently high, the assay might be readable on a fluorescent plate reader (e.g., TECAN Safire^2 or M1000, BMG Pherastar, BioTek Synergy series, Molecular Devices M5 or Flexstation, PerkinElmer Envision). Reading fluorescence from the bottom of the well is less sensitive to the effects of fluorescence interference from the medium, but top reading plates with a z-plane optimization function can also be used. If a new BiFC reporter is being evaluated, monochromator-based readers can be useful to determine the peak emission of the reporter during assay development.

Plates: Due to the fluorescent readout, black-walled plates are recommended. Clear bottom plates, such as μClear plates from Greiner, are suitable for imaging or for detecting fluorescence from the bottom of the plate.

Cell types: a variety of cells have been used for BiFC assays, including U2OS osteosarcoma, HEK293, COS-1, HeLa and NIH3T3. A primary consideration for high content imaging applications is that the cells should adhere well to the plate. For plate-reader based assays, it might be necessary to perform multiple reads per well (several plate readers allow reading at different spots within a well) to compensate for unequal cell distribution.

Buffers: Cells can be grown in the medium recommended by ATCC. However, for maximal signal (especially when using GFP) it is preferable to use phenol red-free and low-glucose medium in live cells, or switch to PBS if fixed cells are used. These solutions can overcome some of the issues associated with autofluorescence coming from medium components. Fixation should be optimized to reduce quenching of fluorescent proteins.

Assay Conditions**:** In general, lower temperatures (between 4 °C and RT) favor fluorescence complementation. However, fluorescence is reconstituted in many BiFC pairs (for example, mCherry) following incubation at higher temperatures (RT to 30 °C). With newer versions such as Venus, the temperature dependence might be less critical. The temperature and time of incubation for each BiFC pair must be determined empirically, and binding-deficient mutants should be included to control for non-specificity of the BiFC reporter. This is particularly important at lower temperatures, where complementation can occur independently of the proteins to which the fragments are fused. Following completion of the screen, cells might be visualized live, or can be washed in PBS and fixed in 4% paraformaldehyde and stored protected from light until examination. Include additional negative control wells where possible (medium in the absence of cells, for example) to determine the background auto-fluorescence.

Data Analysis

General considerations: Stably transfected cells are highly preferable due to the improved run-to-run stability and ability to subsequently separate cells with the desired expression profiles. If the BiFC screen is performed by transient transfection, however, it is critical to co-transfect a full-length fluorescent protein that is spectrally distinct from the BiFC pair as an internal control for transfection efficiency. In this way, the BiFC signal can be normalized to the full-length fluorescent protein signal on a per cell basis, and the results plotted as a histogram. Ideally, the control protein is expressed from the same plasmid as the BiFC pair; otherwise, not all cells will receive both the BiFC and the full-length fluorescent protein with the same ratio. During assay development, plates should include cells with genetic controls for non-interacting mutants (e.g. N- and C-terminal fragments of the fluorescent protein fused to point or deletion mutants of one of the interacting pair). These constructs can be used to validate that PPI formation drives BiFC and set background readings for the assay.

Imaging-based assays: a nuclear stain (e.g., Hoechst 33342) is added to live or fixed cells. Plates are imaged at the appropriate resolution (typically 10x) with the appropriate fluorescence filters for the nuclear stain and fluorescent protein. Cells are identified by nuclear stain and by the intensity and localization of the fluorescent protein. Fluorescence can be monitored in several ways, such as intensity/cell, number of cells above an appropriate threshold, or number or intensity of puncta/cell. The most reproducible and robust measure should be used to calculate the signal for high-signal and low-signal controls and to determine Z'-factor values (see

Equation 1). Normalization of fluorescence to cell number is important to account for inconsistencies in cell plating or cell-growth and to identify toxic compounds. Dead or dying cells can be excluded by morphological criteria (e.g., cell rounding, nuclear condensation) during image segmentation.

Plate-reader based assays: If a plate reader is used for the BiFC measurement, readings are best performed in cells that have been washed to remove dead and floating cells that can give confounding bright fluorescent signals. Normalization of the BiFC to cell number is important to exclude false positives that are toxic. With a genetically engineered system, this normalization can be achieved using an IRES driving expression of a second fluorescent protein or luciferase. Alternatively, a fluorescent indicator of cellular viability (such as resazurin or CellTiter-Fluor) can be used as an indirect measure of cell number. Care should be taken to ensure the spectral overlap between this marker and the BiFC pair is minimal.

Assay Validation

Note: many of the assay validation considerations are similar to those described for BRET.

Optimal protein pair configuration: Determine the optimal orientation of the fusion proteins for both partners testing both N- and C-terminal orientations (total of 8 combinations). Linker design is also a critical factor at this stage; Gly-Ser rich sequences of varying lengths, as well as RPACKIPNDLKQKVMNH and AAANSSIDLISVPVDSR have been used successfully. In order to validate expression by western blot, an epitope tag such as HA or FLAG can also form part of the linker.

Assay validation controls: Known small-molecule inhibitors are ideal, but if no such modulator is available, design a negative control construct where one of the interacting proteins is mutated to disrupt binding. Alternatively, donor saturation can be used (see Assay Validation Steps for NanoBRET).

Optimize expression: Optimize transfection conditions by varying the ratio of the donor and acceptor plasmids. Determine the optimal temperature and time following induction to yield maximal fluorescence. Ideally, fast maturing fluorescent proteins should be chosen, although due to the split, the maturation times for the BiFC pairs might differ from the full-length fluorescent proteins. If the assay is to be performed transiently, titrate the BiFC pair in order to give the most stable signal with the lowest expression of protein. Also establish the shortest time at which the signal is stable, to avoid issues associated with overexpression. Perform both microscopy and plate reader based analyses at this step. Note that the level of fluorescence from split fluorescent proteins is far less than the full-length counterparts, so adjust gain accordingly. Maximal emission might be slightly offset compared to the full-length protein, which must be taken into account when selecting filters. Determine whether the subcellular localization (and if possible biochemical activity) of the BiFC pair is consistent with the biology. For a multiwell format and especially if the reading will be high content, optimize cell seeding density. If individual cells need to be segmented, then aim for a final confluence at measurement of around 50-60%.

Determine the timing of compound addition: Because BiFC for most (if not all) fluorescent proteins is irreversible, it is recommended that compounds are added either before or concomitantly with introduction of the BiFC pair. Set high-signal (vehicle only) low-signal (known inhibitor, mutated PPI pair, etc.) controls.

Statistical Considerations for Fluorescent Protein Fragment Complementation

For an assessment of assay robustness, the Z'-factor can be used (see **Equation 1**). Ideally the low-signal control should be a known small molecule inhibitor of the PPI under investigation. Otherwise, the non-interacting genetic control can be substituted, although this might give an artificially low BiFC signal compared to actual small molecule inhibitors. For initial hit determination, each sample can be compared to the plate-based mean BiFC signal, with a cutoff defined as between 1 and 3 x σ from the mean, depending on the assay robustness. Secondary screens should be in place in order to test for selectivity (for instance, a different PPI using the same

BiFC fluorescent reporter) and for assay interference (for example, test the putative hits against the original PPI, but using split luciferase instead of BiFC).

Sample Validation Data

Example 1. Drug screening for autophagy inhibitors based on the dissociation of Beclin 1-Bcl2 complex using BiFC technique and mechanism of eugenol on anti-influena A virus activity by Dai and colleagues (50). The goal of this study was to identify compounds that disrupt the interaction between Beclin and Bcl2, thereby increasing autophagy. BiFC used mCherry (N-terminal mCherry aa1-169-Bcl2 and C-terminal mCherry aa160-262-Beclin1). For 96-well plate assays, A549 cells were transfected with BiFC constructs before compound addition for 8 h. The cell-based assay using a microplate reader gave a Z'-factor of 0.5. Known autophagy inhibitors (ERK inhibitor, MAPK inhibitor) increased the BiFC signal, whereas activators (anisomycin, H_2O_2) reduced the signal.

Troubleshooting Guide

No or low BiFC signal: Check all orientations of the fusion proteins and linker lengths to find maximal signal and check for expression by western analysis. Ensure that the filter is correct; this can also be optimized by performing a wavelength scan on monochromatic instruments. Determine the optimal temperature for the maturation of the fluorescent BiFC construct (a range is from 4 °C for 1 h through 37 °C). Due to the lower fluorescence intensity of split fluorescent proteins, reads of multiwell plates should be performed in optically clear medium, and might require significant optimization of the photomultiplier tube gain.

BiFC signal not affected by positive controls: The level of the BiFC constructs might be too high. In that case, switch to weaker promoters, shorter incubation times, or use a lower dose of the inducing agent if using an inducible system. Pre-dose cells with compound prior to expression of the BiFC constructs. Also, older fluorescent proteins used for BiFC can be more prone to self-assembly. Use newer variants such as Venus (V150A mutant, split at 1-210 and 210-238) (14).

Luciferase Enzyme Fragment Complementation

Introduction

*Bi*molecular *l*uciferase *c*omplementation (BiLC) is proving to be a powerful tool in many discovery projects (51). For example, PPIs involved in pathways associated with cancer (52, 53), Alzheimer's disease (54) and plant signaling (55) have all been interrogated with this method. Compared with BiFC, luciferase complementation has three distinct advantages. First, the interaction between the two fragments does not require covalent bond formation, and is therefore completely reversible (49). Second, background luminescence is extremely low because the overlapping luciferase fragments, have no activity on their own, and can reassemble into an active enzyme when the proteins they are fused to are brought into close proximity by the interacting proteins of interest. In addition, the affinity of the split luciferase fragments for each other is much weaker than that of the interacting proteins of interest to which they are fused. Finally, the enzymatic activity of luciferase combined with a very low background leads to robust assays when evaluated by common screening metrics such as the Z'-factor (27). Several groups have also reported the use of BiLC in xenograft studies in mice (7, 53, 56), opening the possibility of further characterization of lead compounds during preclinical studies. **Table 4** highlights the many flavors of BiLC. Some of these enzymes require ATP (e.g., firefly), while others do not (e.g. Renilla).

As discussed above for fluorescent protein fragment complementation assays, BiLC can also be multiplexed. Multiplexing is possible because the N-terminus within a particular luciferase family determines it spectral characteristics, whereas the C-terminus contributes the enzymatic function (57, 58). Ozawa and colleagues (52) have identified mutations in the C-terminus of click beetle red (CBR) that allow complementation with N-

terminal fragments of firefly, CBR and click beetle green (CBG) luciferases. Thus, the interaction of a 'bait' protein (fused to the C-terminus of CBR) with two competing proteins (each of which is fused to a different N-terminal luciferase fragment) can be interrogated. In a second study, complementing CBR and CBG fragments were rationally designed using a 'bait' fused to the C-terminus of CBG and prey fused to either the N-terminus of CBG or CBR (59).

The multiplexed BiLC setup allows the detection of multiple PPIs using one common substrate (D-Luciferin or commercial reagents), but also requires careful selection of imaging medium, luminescence emission filters, and software to perform spectral deconvolution of the data. In terms of HTS, the time taken to image each microwell plate will depend on the total photon output, which in turn will be determined by the expression level of the constructs and the strength of their interaction.

BiLC also has some limitations, such as sensitivity to false positives (inhibitors of luciferase enzymatic activity, for example) and false negatives (such as autofluorescent test compounds). Additionally, as with all protein fragment complementation assays, the choice and orientation of fusion partners might be limited to some extent by steric effects. Controls for normalization of cell number or transfection efficiency are also critical in this assay. Suitable controls include co-transfection or IRES-driven expression of a reporter such as Renilla luciferase or fluorescent protein. Finally, because the chemiluminescent reaction requires molecular oxygen, this assay might be affected in HTS screens performed at low oxygen concentrations or those performed in strictly anaerobic organisms.

Assay Development

Instrumentation**:** Many commercial multimodal plate readers can be used to detect the signal from BiLC. These include Promega's GloMax, TECAN's Safire^2 (now TECAN M1000), BMG's Pherastar, BioTek's Synergy series, PerkinElmer's Envision.

Choice of luciferase species: There are multiple choices for luciferase species. Click beetle luciferases show little color sensitivity to pH and temperature changes as opposed to firefly luciferase, in which the maximal emission color can vary by as much as 40 nm under different reaction conditions (59).

Plates: Solid white plates (for example Corning 3570) are recommended. If the assay is performed with live cells and is to be combined with a fluorescence readout, then black walled Greiner μClear plates can be used.

Cell types**:** BiLC has been performed in many cell types including U2OS, SaOS-2, HEK293, H4 neuroglioma, HeLa, COS-7, CHO.

Buffers**:** For kinetic BiLC studies in intact cells, phenol red-free medium containing 100 μM D-luciferin (made from a concentrated stock solution) should be used. The medium should be buffered by the addition of HEPES, for example. For endpoint assays in which the cells are lysed, phenol red-free medium is still recommended.

Kits: For large numbers of plates in HTS format, commercially available luciferase detection reagents can be beneficial due to the batch-to-batch consistency. Although these reagents are frequently added at a 1:1 (reagent:medium) ratio, lower ratios still produce robust signals and can be optimized during assay development.

Titration of reagents: In a 384-well format, a cell density between 2000 and 15000 per well is recommended, depending on the cell type. This is based on a stable cell line system; if the assay will be performed using transient transfection, the optimal cell density might differ.

Assay Conditions**:** If stable cell lines with constitutive expression are used, it is important to screen several clones and pick those that exhibit a robust signal with the minimum amount of overexpression**.** An alternative (but more labor-intensive) option is to develop a bidirectional, doxycycline-inducible reporter system, where each

pair of the fragment complementation module is expressed following addition of doxycycline to the medium. This can aid in control of the level of expression, permit compounds to be added before protein complexes are formed, and restrict the expression of potentially toxic fragment pairs.

Timing: the optimal time for detection of luminescence will vary according to the experimental setup. For example, if a doxycycline inducible system is used, robust signal can be detected 6-10 h following the addition of doxycycline. One advantage of a short incubation time is that secondary compound effects, such as cytotoxicity, are likely to be avoided.

As with conventional luciferase assays, it is important to equilibrate the plates to room temperature for up to 30 min before reading. This improves plate Z'-factors. In comparison with full-length luciferase, the window for detection of the BiLC signal might be significantly shorter when using reagents such as SteadyGlo. Therefore during assay development the time post-lysis at which optimal luciferase activity is detected must be determined. This will dictate the number of plates that can be processed per unit time while maintaining acceptable plate metrics.

Data Analysis

Using real-time analysis of click beetle green and click beetle red luciferase heteroprotein fragment complementation, the hybrid proteins rapidly and directly reconstitute an enzymatic reporter activity within the cell, offering the potential benefit of real-time analysis and enhanced sensitivity for detecting weakly interacting proteins.

Assay Validation Steps

To validate the assay, well-characterized binding partners would be useful. For example, the rapamycin-binding domain of human mTOR (residues 2024–2113) and human FK506-binding protein-12 can provide a well-characterized macromolecular interaction platform to validate protein fragment complementation pairs. In the case of Villalobos et al., the N-terminal fragments of all luciferase PCA constructs were each fused to a flexible G-S linker and the C-terminus of the rapamycin-binding domain of human mTor, while the C-terminal fragments of all constructs were each fused to a flexible G-S linker and the N-terminus of FK506 binding protein-12 (59). Subsequently, HEK293T cells were used for the validation and the protein interactions were analyzed by live cell bioluminescence imaging.

It is important to establish that the half-life of the protein of interest was unaffected by the addition of the fusion protein. To rule out that any change is related to a change in the half-life of the protein, monitor the stability of the protein complex in real time using a HeLa TetOn stable line expressing the two binding partners. Ilagan et al. used the luciferase complementation activity in live NΔE-NLuc/CLuc-RBPjκ-expressing cells that was monitored continuously for 6 h in the presence or absence of the protein translation inhibitor cycloheximide (60). The decay in bioluminescence indicated a Notch intracellular domain half-life of ~180 minutes, consistent with the half-life obtained by pulse-chase experiments in HeLa cells.

Another way to validate PPI assays is to use known inhibitors of the interaction if they are available. This allows the demonstration that the complementarity can be inhibited. Without the availability of a known inhibitor, a peptide can be selected that will block the two complementary fragments.

Statistical Considerations for Luciferase Enzyme Fragment Complementation

The investigator should calculate the Z´-factor (see **Equation 1**) to gauge assay quality and suitability for HTS (27). For the low-signal control, omission of the substrate, an inhibitor (small molecule or peptide), or the cell line expressing only one of the complementary fusion proteins can be used.

Sample Validation Data

Example 1: Dual-color click beetle luciferase heteroprotein fragment complementation assays by Villalobos and colleagues (59). This paper describes a set of reversible, multicolored heteroprotein complementation fragments based on the click beetle luciferases that utilize the same substrate as firefly, D-luciferin. Using luciferase fragment complementation a dual-color quantification was performed with two discreet pairs of interacting proteins simultaneously or two distinct proteins interacting with a third shared protein in live cells with real-time analysis. The proteins examined were β-TrCP, an E3-ligase that regulates both β-catenin and IκBα, and GSK3β as the kinase that regulates IκBα processing.

Example 2: Real time imaging of notch activation with a luciferase complementation reporter by Ilagan and colleagues (60). This paper measures ligand binding of Notch with a luciferase complementation assay. Notch receptors undergo ectodomain shedding followed by γ-secretase-mediated release of the Notch intracellular domain. The Notch intracellular domain translocates to the nucleus and associates with the DNA-binding protein CBF1/RBPjκ/Su(H)/Lag and then subsequently activates gene expression. The activation of specific Notch paralogs was monitored in live cells and in real time using luciferase complementation imaging. In order to validate the system, kinetic analyses of Notch signaling with agonist- and ligand-dependent activation were conducted in live cells.

Alternative Enzyme Complementation Assays

Two other enzyme complementation formats – β-lactamase and β-galactosidase – are also used for cell-based assays. The luciferase approach offers reversibility and the ability to multiplex PPI readouts, which are clear advantages. Because the assay development and troubleshooting will be analogous to the luciferase complementation described above, an extensive assay description is not included here. Nevertheless, a couple of important points are worth noting for each system.

Beta-galactosidase is a natural example of protein fragment complementation, because two polypeptides (lacα and Lacω) derived from the LacZ gene assemble to form the active enzyme. Mutagenesis studies led to the identification of weakly complementing Δα and Δω polypeptides, whose interaction is driven by the protein partners to which they are fused (6). One advantage of this system is that there are diverse methods for the detection of β-galactosidase activity, including immunohistochemical, fluorescent and chemiluminescent detection reagents. Thus, PPIs measured using split β-galactosidase can be detected on plate readers, flow cytometers and conventional microscopic systems. Furthermore, a choice of β-galactosidase reagents with different spectral characteristics facilitates multiplexing of this assay with other readouts. As with the BiLC approach, the enzymatic nature of split β-galactosidase assays can also lead to high signal-to-noise ratios, and a robust format for HTS (for examples, see (23)). Measuring PPI by β-galactosidase fragment complementation can also be used as a surrogate measure of other cellular processes, such as GPCR activation (DiscoveRx PathHunter), proteolysis, and cytoplasmic-nuclear translocation (61).

Beta-lactamase is also amenable to protein fragment complementation (21). Because β-lactamase is a smaller protein, its usage might be advantageous in situations where β-galactosidase complementation is perturbed by steric hindrance. Beta-lactamase has been used to monitor the interaction between components of the Bcl2 pathway (18), GPCRs and arrestin (22) and proteins involved in Toll-like receptor signaling, and it provides a robust assay for HTS (19, 20).

Suggested Websites

Engineering split GFP

http://www.stanford.edu/group/boxer/gfp.html

http://en.wikipedia.org/wiki/Green_fluorescent_protein

Protein folding from fragments

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2211703/?tool=pubmed

FRET considerations

http://www.olympusmicro.com/primer/techniques/fluorescence/fret/fretintro.html

https://www.biotek.com/resources/white-papers/fluorescent-proteins-filters-mirrors-and-wavelengths/

Luciferase

Harvard luciferase blackboard green and red (firefly) luciferase

http://2009.igem.org/Team:Harvard/Split

Spectra for different luciferases

http://www.targetingsystems.net/drug-discovery.php

Click Beetle differences

http://www.promega.com/products/reporter-assays-and-transfection/reporter-vectors-and-cell-lines/chroma_luc-vectors/

References

1. Arkin MR, Glicksman MA, Fu H, Havel JJ, Du Y. Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
2. Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. Chem Biol. 2014;21(9):1102–14. doi: 10.1016/j.chembiol.2014.09.001. Epub 2014/09/23. PubMed PMID: 25237857.
3. Bruckner A, Polge C, Lentze N, Auerbach D, Schlattner U. Yeast two-hybrid, a powerful tool for systems biology. Int J Mol Sci. 2009;10(6):2763–88.doi: 10.3390/ijms10062763. PubMed PMID: 19582228.
4. Karimova G, Pidoux J, Ullmann A, Ladant D. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci U S A. 1998;95(10):5752–6. PubMed PMID: 9576956.
5. Pelletier JN, Campbell-Valois FX, Michnick SW. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. Proc Natl Acad Sci U S A. 1998;95(21):12141–6. PubMed PMID: 9770453.
6. Rossi F, Charlton CA, Blau HM. Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. Proc Natl Acad Sci U S A. 1997;94(16):8405–10. PubMed PMID: 9237989.
7. Luker GD, Luker KE. Luciferase protein complementation assays for bioluminescence imaging of cells and mice. Methods Mol Biol. 2011;680:29–43.doi: 10.1007/978-1-60761-901-7_2. PubMed PMID: 21153371.
8. Bacart J, Corbel C, Jockers R, Bach S, Couturier C. The BRET technology and its application to screening assays. Biotechnol J. 2008;3(3):311–24.doi: 10.1002/biot.200700222. PubMed PMID: 18228541.
9. Machleidt T, Woodroofe CC, Schwinn MK, Mendez J, Robers MB, Zirnmerman K, Otto P, Daniels DL, Kirkland TA, Wood KV. NanoBRET-A Novel BRET Platform for the Analysis of Protein-Protein Interactions. Acs Chem Biol. 2015;10(8):1797–804.doi: 10.1021/acschembio.5b00143. PubMed PMID: 26006698.

10. Pfleger KD, Eidne KA. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). Nat Methods. 2006;3(3):165–74.doi: 10.1038/nmeth841. PubMed PMID: 16489332.
11. You X, Nguyen AW, Jabaiah A, Sheff MA, Thorn KS, Daugherty PS. Intracellular protein interaction mapping with FRET hybrids. Proc Natl Acad Sci U S A. 2006;103(49):18458–63.doi: 10.1073/pnas.0605422103. PubMed PMID: 17130455.
12. Song Y, Madahar V, Liao J. Development of FRET assay into quantitative and high-throughput screening technology platforms for protein-protein interactions. Ann Biomed Eng. 2011;39(4):1224–34.doi: 10.1007/s10439-010-0225-x. PubMed PMID: 21174150.
13. Kodama Y, Hu CD. An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. Biotechniques. 2010;49(5):793–805.doi: 10.2144/000113519. PubMed PMID: 21091444.
14. Ohashi K, Kiuchi T, Shoji K, Sampei K, Mizuno K. Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments. Biotechniques. 2012;52(1):45–50.doi: 10.2144/000113777. PubMed PMID: 22229727.
15. Hu CD, Chinenov Y, Kerppola TK. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell. 2002;9(4):789–98. PubMed PMID: 11983170.
16. Hu CD, Kerppola TK. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat Biotechnol. 2003;21(5):539–45.doi: 10.1038/nbt816. PubMed PMID: 12692560.
17. Fan JY, Cui ZQ, Wei HP, Zhang ZP, Zhou YF, Wang YP, Zhang XE. Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. Biochem Biophys Res Commun. 2008;367(1):47–53.doi: 10.1016/j.bbrc.2007.12.101. PubMed PMID: 18158915.
18. Galarneau A, Primeau M, Trudeau LE, Michnick SW. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. Nat Biotechnol. 2002;20(6):619–22.doi: 10.1038/nbt0602-619. PubMed PMID: 12042868.
19. Lee HK, Brown SJ, Rosen H, Tobias PS. Application of beta-lactamase enzyme complementation to the high-throughput screening of toll-like receptor signaling inhibitors. Mol Pharmacol. 2007;72(4):868–75.doi: 10.1124/mol.107.038349. PubMed PMID: 17615244.
20. Lee HK, Dunzendorfer S, Tobias PS. Cytoplasmic domain-mediated dimerizations of toll-like receptor 4 observed by beta-lactamase enzyme fragment complementation. J Biol Chem. 2004;279(11):10564–74.doi: 10.1074/jbc.M311564200. PubMed PMID: 14699116.
21. Wehrman T, Kleaveland B, Her JH, Balint RF, Blau HM. Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments. Proc Natl Acad Sci U S A. 2002;99(6):3469–74.doi: 10.1073/pnas.062043699. PubMed PMID: 11904411.
22. Zhao CK, Yin Q, Li SY. A high-throughput screening system for G-protein-coupled receptors using beta-lactamase enzyme complementation technology. Acta Pharmacol Sin. 2010;31(12):1618–24.doi: 10.1038/aps.2010.154. PubMed PMID: 21102483.
23. Graham DL, Bevan N, Lowe PN, Palmer M, Rees S. Application of beta-galactosidase enzyme complementation technology as a high throughput screening format for antagonists of the epidermal growth factor receptor. J Biomol Screen. 2001;6(6):401–11.doi: 10.1089/108705701753364896. PubMed PMID: 11788058.
24. Smith AM, Ammar R, Nislow C, Giaever G. A survey of yeast genomic assays for drug and target discovery. Pharmacol Ther. 2010;127(2):156–64.doi: 10.1016/j.pharmthera.2010.04.012. PubMed PMID: 20546776.
25. Rogers B, Decottignies A, Kolaczkowski M, Carvajal E, Balzi E, Goffeau A. The pleitropic drug ABC transporters from Saccharomyces cerevisiae. J Mol Microbiol Biotechnol. 2001;3(2):207–14. PubMed PMID: 11321575.
26. Colton CK, Kong Q, Lai L, Zhu MX, Seyb KI, Cuny GD, Xian J, Glicksman MA, Lin CL. Identification of translational activators of glial glutamate transporter EAAT2 through cell-based high-throughput screening:

an approach to prevent excitotoxicity. J Biomol Screen. 2010;15(6):653–62.doi: 10.1177/1087057110370998. PubMed PMID: 20508255.
27. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. 1999;4(2):67–73. PubMed PMID: 10838414.
28. Auld DS, Inglese J. Interferences with Luciferase Reporter Enzymes. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
29. Simeonov A, Davis MI. Interference with Fluorescence and Absorbance. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
30. Simeonov A, Jadhav A, Thomas CJ, Wang Y, Huang R, Southall NT, Shinn P, Smith J, Austin CP, Auld DS, Inglese J. Fluorescence spectroscopic profiling of compound libraries. J Med Chem. 2008;51(8):2363–71.doi: 10.1021/jm701301m. PubMed PMID: 18363325.
31. Jadhav A, Ferreira RS, Klumpp C, Mott BT, Austin CP, Inglese J, Thomas CJ, Maloney DJ, Shoichet BK, Simeonov A. Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. J Med Chem. 2010;53(1):37–51.doi: 10.1021/jm901070c. PubMed PMID: 19908840.
32. Auld DS, Southall NT, Jadhav A, Johnson RL, Diller DJ, Simeonov A, Austin CP, Inglese J. Characterization of chemical libraries for luciferase inhibitory activity. J Med Chem. 2008;51(8):2372–86.doi: 10.1021/jm701302v. PubMed PMID: 18363348.
33. Dahlin JL, Baell J, Walters MA. Assay Interference by Chemical Reactivity. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
34. Auld DS, Inglese J, Dahlin JL. Assay Interference by Aggregation. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
35. Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998;67:509–44.doi: 10.1146/annurev.biochem.67.1.509. PubMed PMID: 9759496.
36. Forster T. *Zwischenmolekulare Energiewanderung Und Fluoreszenz. Ann Phys-Berlin. 1948;2(1-2):55–75.
37. Zimmermann T, Rietdorf J, Girod A, Georget V, Pepperkok R. Spectral imaging and linear un-mixing enables improved FRET efficiency with a novel GFP2-YFP FRET pair. FEBS Lett. 2002;531(2):245–9. PubMed PMID: 12417320.
38. van der Krogt GN, Ogink J, Ponsioen B, Jalink K. A comparison of donor-acceptor pairs for genetically encoded FRET sensors: application to the Epac cAMP sensor as an example. PLoS One. 2008;3(4):e1916doi: 10.1371/journal.pone.0001916. PubMed PMID: 18382687.
39. Roszik J, Toth G, Szollosi J, Vereb G. Validating pharmacological disruption of protein-protein interactions by acceptor photobleaching FRET imaging. Methods Mol Biol. 2013;986:165–78.doi: 10.1007/978-1-62703-311-4_11. PubMed PMID: 23436412.
40. Alibhai D, Kelly DJ, Warren S, Kumar S, Margineau A, Serwa RA, Thinon E, Alexandrov Y, Murray EJ, Stuhmeier F, Tate EW, Neil MA, Dunsby C, French PM. Automated fluorescence lifetime imaging plate reader and its application to Forster resonant energy transfer readout of Gag protein aggregation. J Biophotonics. 2013;6(5):398–408.doi: 10.1002/jbio.201200185. PubMed PMID: 23184449.

41. Sun Y, Hays NM, Periasamy A, Davidson MW, Day RN. Monitoring protein interactions in living cells with fluorescence lifetime imaging microscopy. Methods Enzymol. 2012;504:371–91.doi: 10.1016/B978-0-12-391857-4.00019-7. PubMed PMID: 22264545.
42. Iversen PW, Beck B, Chen YF, Dere W, Devanarayan V, Eastwood BJ, Farmen MW, Iturria SJ, Montrose C, Moore RA, Weidner JR, Sittampalam GS. HTS Assay Validation. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, Austin C, Baell J, Bejcek B, Chung TDY, Dahlin JL, Devanaryan V, Foley TL, Glicksman M, Hall MD, Hass JV, Inglese J, Iversen PW, Kahl SD, Kales SC, Lal-Nag M, Li Z, McGee J, McManus O, Riss T, Trask OJ, Jr., Weidner JR, Xia M, Xu X, editors. Assay Guidance Manual. Bethesda (MD)2004.
43. Buntru A, Zimmermann T, Hauck CR. Fluorescence resonance energy transfer (FRET)-based subcellular visualization of pathogen-induced host receptor signaling. Bmc Biol. 2009;7. doi: Artn 8110.1186/1741-7007-7-81.
44. Zacharias DA, Violin JD, Newton AC, Tsien RY. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science. 2002;296(5569):913–6. 10.1126/science.1068539. doi: DOI. PubMed PMID: 11988576.
45. Picaud S, Fedorov O, Thanasopoulou A, Leonards K, Jones K, Meier J, Olzscha H, Monteiro O, Martin S, Philpott M, Tumber A, Filippakopoulos P, Yapp C, Wells C, Che KH, Bannister A, Robson S, Kumar U, Parr N, Lee K, Lugo D, Jeffrey P, Taylor S, Vecellio ML, Bountra C, Brennan PE, O'Mahony A, Velichko S, Muller S, Hay DC, Daniels DL, Urh M, La Thangue NB, Kouzarides T, Prinjha R, Schwaller J, Knapp S. Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. Cancer Res. 2015;75(23):5106–19.doi: 10.1158/0008-5472.Can-15-0236. PubMed PMID: 26552700.
46. Ghosh I, Hamilton AD, Regan L. Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. Journal of the American Chemical Society. 2000;122(23):5658–9. 10.1021/ja994421w. doi: DOI.
47. Kerppola TK. Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. Nat Protoc. 2006;1(3):1278–86.doi: 10.1038/nprot.2006.201. PubMed PMID: 17406412.
48. Magliery TJ, Wilson CG, Pan W, Mishler D, Ghosh I, Hamilton AD, Regan L. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J Am Chem Soc. 2005;127(1):146–57.doi: 10.1021/ja046699g. PubMed PMID: 15631464.
49. Stefan E, Aquin S, Berger N, Landry CR, Nyfeler B, Bouvier M, Michnick SW. Quantification of dynamic protein complexes using Renilla luciferase fragment complementation applied to protein kinase A activities in vivo. Proc Natl Acad Sci U S A. 2007;104(43):16916–21.doi: 10.1073/pnas.0704257104. PubMed PMID: 17942691.
50. Dai JP, Zhao XF, Zeng J, Wan QY, Yang JC, Li WZ, Chen XX, Wang GF, Li KS. Drug Screening for Autophagy Inhibitors Based on the Dissociation of Beclin 1-Bcl2 Complex Using BiFC Technique and Mechanism of Eugenol on Anti-Influenza A Virus Activity. Plos One. 2013;8(4). doi: ARTN e6102610.1371/journal.pone.0061026.
51. Verhoef L, Wade M. Visualization of Protein Interactions in Living Cells Using Bimolecular Luminescence Complementation (BiLC). Curr Protoc Protein Sci. 2017;90:30 5 1- 5 14. Epub 2017/11/02. doi: 10.1002/cpps.42.
52. Hida N, Awais M, Takeuchi M, Ueno N, Tashiro M, Takagi C, Singh T, Hayashi M, Ohmiya Y, Ozawa T. High-sensitivity real-time imaging of dual protein-protein interactions in living subjects using multicolor luciferases. PLoS One. 2009;4(6):e5868doi: 10.1371/journal.pone.0005868. PubMed PMID: 19536355.
53. Luker KE, Piwnica-Worms D. Optimizing luciferase protein fragment complementation for bioluminescent imaging of protein-protein interactions in live cells and animals. Methods Enzymol. 2004;385:349–60.doi: 10.1016/S0076-6879(04)85019-5. PubMed PMID: 15130748.
54. Hashimoto T, Adams KW, Fan Z, McLean PJ, Hyman BT. Characterization of oligomer formation of amyloid-beta peptide using a split-luciferase complementation assay. J Biol Chem. 2011;286(31):27081–91.doi: 10.1074/jbc.M111.257378. PubMed PMID: 21652708.

55. Gehl C, Kaufholdt D, Hamisch D, Bikker R, Kudla J, Mendel RR, Hansch R. Quantitative analysis of dynamic protein-protein interactions in planta by a floated-leaf luciferase complementation imaging (FLuCI) assay using binary Gateway vectors. Plant J. 2011;67(3):542–53.doi: 10.1111/j.1365-313X.2011.04607.x. PubMed PMID: 21481030.
56. Paulmurugan R, Umezawa Y, Gambhir SS. Noninvasive imaging of protein-protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. Proc Natl Acad Sci U S A. 2002;99(24):15608–13.doi: 10.1073/pnas.242594299. PubMed PMID: 12438689.
57. Branchini BR, Southworth TL, Murtiashaw MH, Wilkinson SR, Khattak NF, Rosenberg JC, Zimmer M. Mutagenesis evidence that the partial reactions of firefly bioluminescence are catalyzed by different conformations of the luciferase C-terminal domain. Biochemistry. 2005;44(5):1385–93.doi: 10.1021/bi047903f. PubMed PMID: 15683224.
58. Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, Kato H. Structural basis for the spectral difference in luciferase bioluminescence. Nature. 2006;440(7082):372–6.doi: 10.1038/nature04542. PubMed PMID: 16541080.
59. Villalobos V, Naik S, Bruinsma M, Dothager RS, Pan MH, Samrakandi M, Moss B, Elhammali A, Piwnica-Worms D. Dual-color click beetle luciferase heteroprotein fragment complementation assays. Chem Biol. 2010;17(9):1018–29.doi: 10.1016/j.chembiol.2010.06.018. PubMed PMID: 20851351.
60. Ilagan MX, Lim S, Fulbright M, Piwnica-Worms D, Kopan R. Real-time imaging of notch activation with a luciferase complementation-based reporter. Sci Signal. 2011;4(181):rs7doi: 10.1126/scisignal.2001656. PubMed PMID: 21775282.
61. Zhu PJ, Zheng W, Auld DS, Jadhav A, Macarthur R, Olson KR, Peng K, Dotimas H, Austin CP, Inglese J. A miniaturized glucocorticoid receptor translocation assay using enzymatic fragment complementation evaluated with qHTS. Comb Chem High Throughput Screen. 2008;11(7):545–59. PubMed PMID: 18694391.
62. Tomosugi W, Matsuda T, Tani T, Nemoto T, Kotera I, Saito K, Horikawa K, Nagai T. An ultramarine fluorescent protein with increased photostability and pH insensitivity. Nat Methods. 2009;6(5):351–3.doi: 10.1038/nmeth.1317. PubMed PMID: 19349978.
63. Su T, Pan S, Luo Q, Zhang Z. Monitoring of dual bio-molecular events using FRET biosensors based on mTagBFP/sfGFP and mVenus/mKOkappa fluorescent protein pairs. Biosens Bioelectron. 2013;46:97–101.doi: 10.1016/j.bios.2013.02.024. PubMed PMID: 23517824.
64. Woehler A. Simultaneous quantitative live cell imaging of multiple FRET-based biosensors. PLoS One. 2013;8(4):e61096doi: 10.1371/journal.pone.0061096. PubMed PMID: 23613792.
65. Ai HW, Hazelwood KL, Davidson MW, Campbell RE. Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. Nat Methods. 2008;5(5):401–3.doi: 10.1038/nmeth.1207. PubMed PMID: 18425137.
66. Ouyang M, Sun J, Chien S, Wang Y. Determination of hierarchical relationship of Src and Rac at subcellular locations with FRET biosensors. Proc Natl Acad Sci U S A. 2008;105(38):14353–8.doi: 10.1073/pnas.0807537105. PubMed PMID: 18799748.
67. Su T, Zhang Z, Luo Q. Ratiometric fluorescence imaging of dual bio-molecular events in single living cells using a new FRET pair mVenus/mKOkappa-based biosensor and a single fluorescent protein biosensor. Biosens Bioelectron. 2012;31(1):292–8.doi: 10.1016/j.bios.2011.10.034. PubMed PMID: 22088261.
68. Galperin E, Verkhusha VV, Sorkin A. Three-chromophore FRET microscopy to analyze multiprotein interactions in living cells. Nat Methods. 2004;1(3):209–17.doi: 10.1038/nmeth720. PubMed PMID: 15782196.
69. Ghosh PS, Hamilton AD. Noncovalent template-assisted mimicry of multiloop protein surfaces: assembling discontinuous and functional domains. J Am Chem Soc. 2012;134(32):13208–11.doi: 10.1021/ja305360q. PubMed PMID: 22839643.
70. Shyu YJ, Liu H, Deng X, Hu CD. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. Biotechniques. 2006;40(1):61–6. PubMed PMID: 16454041.

71. Chu J, Zhang Z, Zheng Y, Yang J, Qin L, Lu J, Huang ZL, Zeng S, Luo Q. A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein interactions under physiological conditions. Biosens Bioelectron. 2009;25(1):234–9.doi: 10.1016/j.bios.2009.06.008. PubMed PMID: 19596565.

Pharmacological Characterization of GPCR Agonists, Antagonists, Allosteric Modulators and Biased Ligands from HTS Hits to Lead Optimization

Chi Shing Sum, Ph.D.,[1] Brian J. Murphy, Ph.D.,[1] Zhuyin Li, Ph.D.,[1] Tao Wang, M.Sc.,[1] Litao Zhang, Ph.D.,[1] and Mary Ellen Cvijic, Ph.D.[1]

Created: November 1, 2019.

Abstract

G-Protein Coupled Receptors (GPCRs) have been an important target class due to their involvement in a plethora of physiological processes. Their high druggability has led to the discovery of many medicines used to treat various diseases and conditions by modulating their functions. Understanding of the molecular mechanisms of GPCR functions has been advancing through the years. This knowledge opens up new opportunities to modulate this family of drug targets beyond the traditional agonists and antagonists, to include allosteric modulators and biased ligands. Benefiting from this new knowledge necessitates updates in the methodologies used for the drug discovery process, especially during the lead discovery and optimization phases. This chapter, based on current understanding of GPCR functions, describes various important mechanistic pharmacology assays useful during lead discovery and optimization. These assays can be applied to better understand the underlying pharmacological properties of compounds and increase the chance of success.

Introduction to GPCRs as Drug Targets

G-Protein Coupled Receptors (GPCRs) have been an important target class in drug discovery. A number of important medicines have been developed that target GPCRs for the treatment of diseases from oncology, immunology, cardiovascular, neurological, and metabolic disorders. Based on a 2016 review, GPCRs represent the major family targeted by small molecule drugs, accounting for 33% of all approved small molecule drugs (1). Moreover, GPCR-based drugs have the highest first-in-class approval rate among four major families of target classes—GPCR, nuclear hormone receptor, ion channel, and kinase—in each of the five-year periods from before 1990 through 2015 (1). GPCR-based therapeutics are mostly agonists that activate GPCRs or antagonists that inhibit the action of natural agonists. These days, drugs targeting GPCRs as allosteric modulators have been gaining attention in the industry (2,3), and a number of drugs based on allosteric mechanism have been successfully developed (4,5). Furthermore, with a deeper understanding of GPCR function today, there is great enthusiasm for the development of biased ligands that selectively turn on the therapeutically beneficial cellular pathways while avoiding pathways that lead to unwanted side effects (6-9).

As with other target classes, GPCR drug discovery campaigns have benefited from the wide acceptance of automation technologies, which have greatly facilitated lead identification and optimization (10). In addition, the adoption of recombinant technologies has enabled better assay sensitivity and robustness, and fostered the development of new screening technologies. With the ability to assay a recombinantly expressed target in comparison to a null background, it greatly improves the ability to specifically detect on- and off-target activities. This is an important means to ensure the lead identification and optimization campaigns are on target and to assist potential liability assessment. HTS likely generates hits with different mechanisms of action, and attention should be paid to examine the underlying pharmacological modes of action of the candidate molecules.

Author Affiliation: 1 Bristol-Myers Squibb Company, NJ; Email: chi.sum@bms.com; Email: brian.j.murphy@bms.com; Email: zhuyin.li@bms.com; Email: tao.wang2@bms.com; Email: litao.zhang@bms.com; Email: maryellen.cvijic@bms.com.

Corresponding author.

Otherwise, one may encounter potential pitfalls that could confuse the structure-activity relationship, or widen the gap between *in vitro* and *in vivo* connectivity. In addition to routine screening assays, follow-up assays that address the mechanisms and those that are more physiologically relevant should be conducted to bridge the *in vitro-in vivo* connectivity in order to select compounds with the desirable pharmacological profiles for downstream testing (11,12).

With the renewed understanding of GPCR signaling mechanisms and the availability of novel assay technologies, GPCR-based drug discovery programs continue to identify new drugs with better efficacy and safety profiles. This chapter reviews the current state-of-the-art cell- and biochemical-based GPCR *in vitro* pharmacological assay methodology important for the discovery and progression of compounds during the lead discovery and optimization phases.

GPCR Signaling from a Drug Discovery Perspective

GPCRs have been known to couple to multiple subtypes of G proteins (Figure 1). Different GPCRs primarily couple to specific G protein subtypes, for example, $G\alpha_i$, $G\alpha_s$ and $G\alpha_q$; as such, they are referred to as $G\alpha_i$-coupled, $G\alpha_s$-coupled, and $G\alpha_q$-coupled receptors. In addition to the primary coupling, secondary coupling to other G proteins subtypes may also occur.

The activation of any given G proteins leads to the activation of downstream second messenger signaling. Common in GPCR drug discovery, compound activities are measured by their effects on specific second messenger signaling. Hence, the activation of a GPCR that couples to the $G\alpha_s$ protein is usually assessed by the production of cAMP, the formation of which is catalyzed adenylate cyclase upon activation by the $G\alpha_s$ protein. In contrast, the activation of a GPCR that couples to the $G\alpha_i$ protein causes suppression of adenylate cyclase activities, leading to reduced cAMP levels. The second messenger downstream of the activation of $G\alpha_q$ protein coupling is inositol-1,4,5-trisphosphate (IP_3), which is produced from phosphatidylinositol 4,5-bisphosphate through the action of phospholipase C. IP_3 is labile, but it mediates the increase in intracellular calcium through release from sarcoplasmic reticulum. A variety of technologies are available that measure intracellular cAMP levels and calcium levels. Although IP_3 is labile and degrades to IP_2 and subsequently to IP_1, technology is also available the measures IP_1 accumulation in the presence of Li^{2+}, which prevents its degradation. Table 1 lists some of the widely used commercially available assay technologies for assessment of GPCR activation of the G_s, G_i and G_q classes of G proteins. Furthermore, detailed guidance on these technologies is available in other chapters of this manual (13-16). Although no commercial signaling assays have been developed measuring the proximal signaling from other G protein classes such as the $G_{12/13}$ subtypes, G protein activation in general can also be assayed by various gene reporter assays or by biosensors (e.g.,(17, 18)).

Independent of the G protein coupling, the recruitment of β-arrestin also constitutes an important piece in GPCR signaling as well as desensitization response. Similarly, commercial technologies are available for the measurement of β-arrestin recruitment, and details of the assay design is covered in another chapter of this manual (14). There are also literature reviews of the variety of signaling assay technologies available commercially for compound screening (19).

Although GPCRs may couple to various G proteins and modulate a variety of second messenger pathways and downstream cellular effects, in the past it was thought that a receptor may activate all its downstream pathway to a similar extent (6). Thus, for practicality, primary or surrogate signaling pathways, including promiscuous G proteins, were chosen as a readout for the assessment of GPCR function (20,21), regardless of whether different secondary messenger pathways may lead to different biological consequences. While the use of surrogate readouts, including the utilization of chimeric G proteins and promiscuous G proteins has continued to play an important role, the use of surrogate readout alone is inadequate for GPCR drug discovery. This is due to the fact that GPCR ligands, either natural agonists, synthetic small molecules or receptor activating antibodies, can differentially activate one signaling pathway over another (signaling bias) with differing rank orders of potency

and efficacy. Failure to recognize the possibility of a ligand signaling bias when using a surrogate signaling assay to assess compound activity can potentially be misleading if the compound structure-activity relationship required for the therapeutic response is different than that used in the surrogate signaling assay.

With deeper understanding of GPCR function, it is now clear that GPCR signaling can be biased toward or against certain pathways, and different signaling pathways can be modulated differentially by their ligands with different potency and efficacy (i.e., biased ligands) (6). Therefore, it is essential in GPCR drug discovery to assess compound activities in the therapeutically relevant pathways, and ideally to profile compound activities across these different pathways. In addition to biased ligands, allosteric modulators belong to another important class of ligands in GPCR drug discovery. Allosteric modulators may display special pharmacological properties not shared by orthosteric ligands. For example, they may potentiate or reduce potency and/or efficacy of orthosteric ligands, without intrinsic activity of their own. They may achieve these by allosterically modulating the affinity and/or functional efficacy of another ligand.

While traditional GPCR drug discovery programs have very much focused on identifying agonists and antagonists, the discovery of biased and allosteric mechanisms has spurred pharmaceutical programs to look into developing these other classes of pharmacological agents (4,22,23). A number of preclinical studies *in vitro* and *in vivo* have demonstrated the potential benefits of a biased ligand when it can be harnessed to selectively activate a therapeutically relevant pathway while avoiding another pathway that could cause potential liabilities. Interestingly, clinically approved drugs that were developed as antagonists or inverse agonists were later found to be possess biased signaling properties. The β blockers are among one of the well-characterized examples of hidden bias (24), although the relation to clinical efficacy is still not clear. Additionally, allosteric modulators have been successfully developed into approved drugs (25,26). By targeting the allosteric site, the ligand might be more selective among closely related members, wherein the orthosteric ligand may more likely cross-react due to similarity of the orthosteric pockets across the family members. Moreover, unlike agonists and antagonists, the nature of allosteric mechanisms also means that their modulatory effect may also reach a ceiling. Under certain circumstances, this can be advantageous in that they would not cause over-stimulation or over-suppression of the targeted pathway. With respect to discovering allosteric ligands vs. orthosteric ligands in the different classes of GPCRs, class A GPCRs bind small molecules, lipids, peptides, and proteins in their orthosteric sites, class B GPCRs interact with peptides and proteins agonists, and class C recognize ions, amino acids, and small molecules agonists. It is therefore conceivable that these structural divergences in the orthosteric sites could influence the potential for discovering orthosteric vs. allosteric small molecule ligands among the different GPCR classes.

Along with agonists and antagonists, biased and allosteric ligands therefore offer additional tractability for modulating any given GPCR target. Importantly, whether a discovery program is aiming to specifically develop an agonist, antagonist, biased ligand or allosteric ligand, it is unavoidable that the screening programs will come across compounds of all kinds of pharmacological modes of action. Consequently, it is essential a GPCR drug discovery program be able to design assays that properly identify and characterize all these classes of compounds.

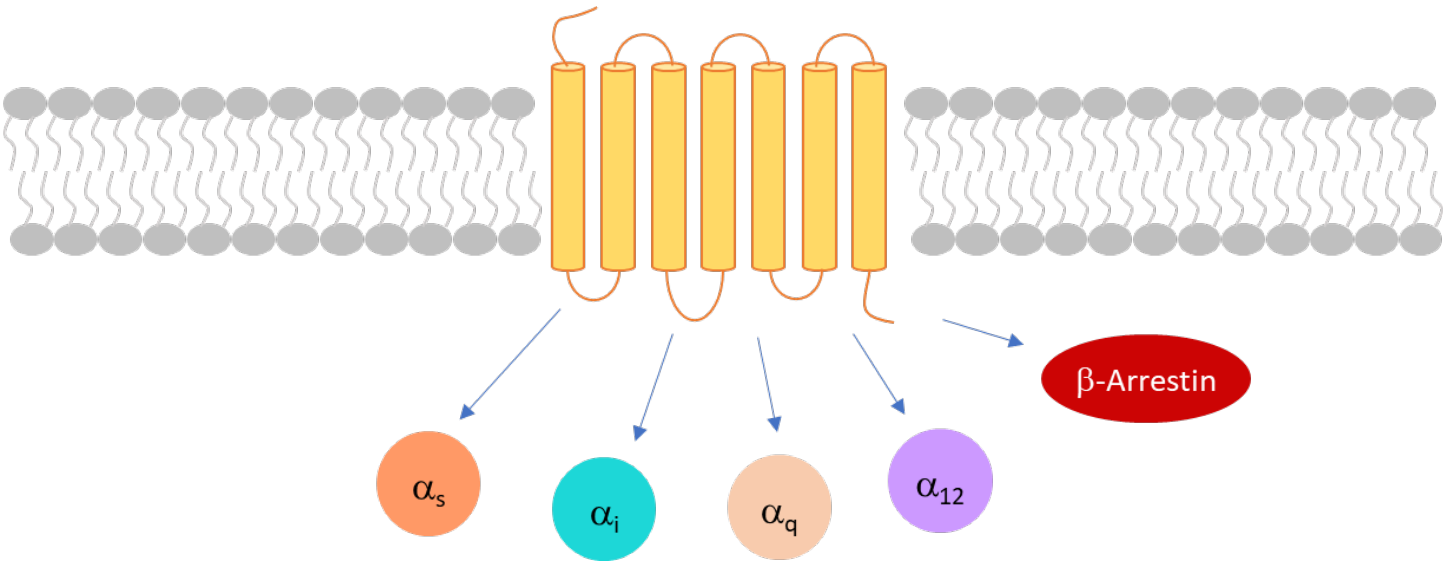


Figure 1. Key GPCR signaling pathways shown with representative G protein subtypes

Table 1. Commonly measured proximal signaling readouts for various G protein coupling in drug discovery

Classes of G protein Coupling	Commonly Measured Proximal Signaling Readouts
G_s	cAMP accumulation
G_i, G_o, G_z	cAMP reduction
G_q, G_{11}, G_{14}, G_{15}, G_{16}	IP_1 and calcium

High-Throughput Screening Assays for Identification and Optimization of GPCR Ligands

HTS Assay Considerations

While a high-throughput screening (HTS) campaign may sometimes employ similar assay technologies as in lead optimization, considerations for successful HTS and lead optimization campaigns can differ due to their different focuses (Table 2). An HTS campaign typically lasts one to two months, with the goal to identify leads for the target of interest. In contrast, lead optimization efforts in the drug discovery process may last many years, and involve many aspects, such as structure-activity relationship studies, optimization of pharmacokinetics properties for *in vivo* testing, *in vitro* to *in vivo* translation, as well as the assessment and mitigation of potential liabilities. Consequently, considerations and requirements for establishing HTS assays can be different from screening assays for lead optimization. Guidelines are provided in this manual for the setup, validation and statistical assessment of screening assay performance, ADME, and *in vivo* testing especially in the section Assay Validation, Operations and Quality Control and the chapter Pharmacokinetics and Drug Metabolism. These guidelines are applicable to GPCR drug discovery in general and this chapter will focus on those specific areas of GPCR drug discovery which are not covered elsewhere in this manual.

This manual contains several chapters that describe in detail the necessary considerations for an HTS screen in general (27-30). In essence, although quantitative HTS can be done, HTS campaigns are conducted primarily in one compound concentrations either as single compound or in a mixture. In addition to proper instrumentation, assay design that provides high precision and minimal background noise is important to allow detection of low activity hits. In order to maintain data precision, HTS assays would ideally be conducted with cells that are cultured from within one to two passages and from the same prepared batch. Biochemical reagents of the same

lots can be acquired in quantities sufficient for an entire HTS campaign. This is generally not difficult, as HTS campaigns are close-ended, therefore the anticipated resource requirement can be determined. Since an HTS run can go continuously for days to weeks, stability of the reagents (e.g., cell lines, agonists, tool compounds, detection reagents) is often an important factor such that stability testing should be built into the HTS assay design. Various pharmaceutical companies have conducted in-house studies suggesting that compounds in the repository are generally stable under various described storage conditions (e.g., (31,32)). If needed, hits that come out of an HTS campaign can be verified for the integrity of the compounds by means such as mass determination, structural analysis, fractionation studies, as well as new chemical synthesis.

With regard to GPCRs in particular, in an HTS screen it is important to design the screening assays in relationship to the known biology of the target, and to ensure that the identified hits are acting on the therapeutically relevant pathways, thus minimizing the potential of interrogating the wrong cellular functions. Endogenously expressed receptors from cells that originate from the target tissues would likely resemble the *in vivo* function more closely. In contrast, recombinant cell systems overexpressing the GPCR target of interest can offer increased sensitivity for detection of compound activity. The advancement of assay miniaturization technologies has helped drive down the cost of an HTS campaign, which may amount to over one million compounds. Along with assay miniaturization, there is an increased demand for sensitivity and precision. Experiences from multiple pharmaceutical screening departments have suggested that sensitivity and functional relevance can sometimes compete against each other in the design of a robust assay. As the main objective for an HTS campaign is to identify leads with all possible mechanisms, at times a balance or trade-off has to be made between sensitivity and functional relevance.

Figure 2 illustrates an example of a general HTS screening paradigm. Ever since the availability of homogeneous assays that measure signaling responses from GPCR activation, these assays have gained favor over the relatively more expensive binding assays, which are often non-homogenous. Upon selection of the primary screening assay, the primary HTS is usually conducted in a single high compound concentration due to cost. The primary hits are then followed up by concentration-response assays, although quantitative HTS using a concentration series has been described to allow one step assessment of compound concentration-response (33). During an HTS campaign, which may span weeks, certain screening plates may fail and need to be retested. In addition to the primary assays, a counter screen is usually conducted using an unrelated GPCR target. When the primary screen is a signaling assay, this counter screen will be a GPCR that couples to the same signaling pathway and can be assayed in the same readout as the primary HTS assay, in order to filter out false positive hits that could be acting elsewhere along measured signaling pathways. For example, compounds that activate adenylate cyclase would be detected as hits in both the targeted screen and the counter screen and should be eliminated as false positive hits. Foreseeably, the extent of false positive hits depends on the compound collection in the screening deck. Certain collections may contain compounds that generate more false positive hits in one signaling pathway versus another.

After the selection of hit candidates based on the primary and counter screening results, a concentration-response assay can be performed to allow rank ordering of compounds by potency and efficacy. The same counter screening assay can be performed in a concentration-response manner at this stage. Moreover, an orthogonal assay measuring the response using a different technology may be useful so as to further confirm the hit. Depending on the resource availability, subtype selectivity and species cross-reactivity can be assessed at this time or the hits can be progressed to lead optimization upon chemistry triage.

Figure 2

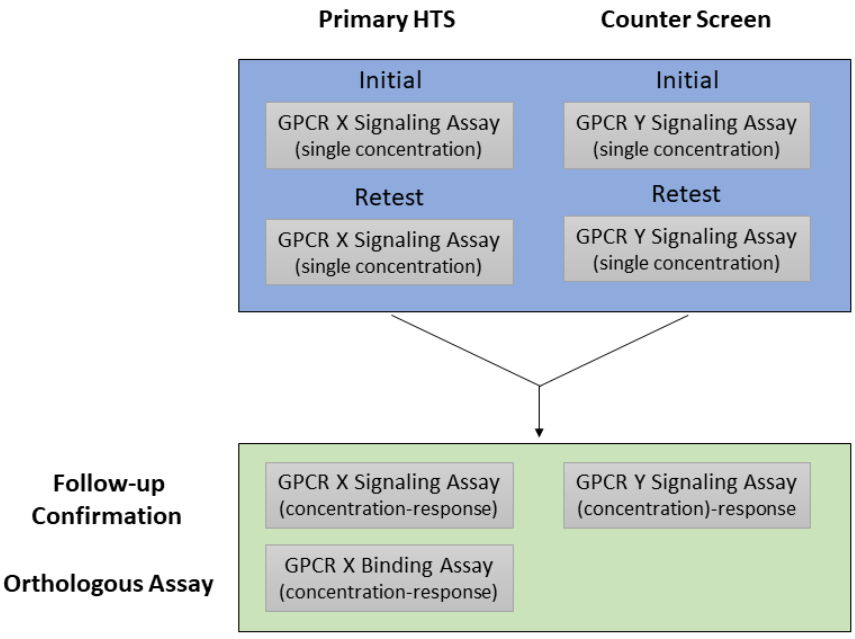


Figure 2. An example of a general HTS screening paradigm for GPCR X and an unrelated GPCR Y as counter screen.

Table 2. Important considerations that differ between high-throughput screening campaigns and lead optimization of efforts.

	HTS	Lead Optimization
Time	1 to 2 months	Over many years
Priority of performance	Assay sensitivity > functional relevance	Assay sensitivity ≈ functional relevance
Accuracy and precision	High precision required for single-point assay	Accuracy and precision required from multiple runs
Cellular model	Recombinant vs. endogenous targets	Recombinant, endogenous targets and primary cells within screening trees
Cell culture	Restricted to one or a few passages	Cell culture may span longer passages;
Reagents	Typically a large supply of single batches of reagents	Batch-to-batch variations of cryopreserved cells, primary cells and biochemical reagents
Compound integrity	Structural integrity of compounds in the deck collection	Compound stability in dry vs. solubilized forms.
Signal tolerance	False positive and negative rates	n/a

Lead Optimization Assay Considerations

Due to differences in objectives and time frames, lead optimization assays have additional considerations compared to HTS campaigns. In addition to sensitivity and reproducibility, lead optimization assays should provide the functional relevance for better prediction of efficacy and liability. As in HTS campaigns, overexpressing cell lines are useful for their generally higher sensitivity and robustness of performance. Also, endogenously expressing cell lines and primary cells are valuable, as they can potentially provide more physiologically relevant readouts. Accordingly, lead optimization assays are usually set up into tiers, in which

more sensitive assays may be positioned earlier in the screening tree, followed by assays that may be less sensitive or robust but offer more relevant end point measurements. With increasing technological advancements, it is also possible to develop lead optimization assays that are sensitive and functionally relevant.

Lead optimization assays are primarily conducted to examine the concentration-response, and/or time-dependent response for kinetics studies. The precision is assessed as confidence intervals of the measured pharmacological parameters across multiple experiments (e.g., EC_{50}, IC_{50}, K_d, k_{on}, k_{off}). The accuracy and precision required for lead optimization assays also comes from the fact that data needs to show reproducibility over a long period of time for guiding Structure-Activity-Relationship (SAR) or Structure-Liability-Relationship (SLR) studies, and selectivity profiling. Again, this demands assays of high statistical robustness.

Batch-to-batch differences in cellular and biochemical reagents may introduce assay variation into the assays and this is more likely to happen because of the longer time frame for lead optimization. Cell culture processes should be standardized, in terms of the passage schedules, passage limit and cryopreservation method. Serum should be batch tested if possible and a large enough lot acquired to minimize variations. Batch testing is especially important when primary cells are used, as it is not uncommon that the response from each batch may vary considerably. If possible, the ability to test multiple batches will be beneficial to assess the effect on the assays. For biochemical reagents, batch testing can be performed similarly if needed. Overall, it is important that a log sheet of reagent batches be maintained particularly for those that are prone to variations.

Unless compounds are replenished through an enrichment effort, HTS campaigns are carried out mostly using pre-existing solubilized compounds. In contrast, solutions from newly synthesized and freshly solubilized compounds are generally used in the initial testing in lead optimization assays. Subsequent testing may then utilize pre-solubilized material. Depending on the storage conditions, presence of contaminants, solubility, and stability of the compounds either intrinsic or in the solvent or vessel, assay reproducibility can be affected. Additionally, batch-to-batch difference in compound purification is possible. It is important to be able to maintain logs of the compound batch and the solubilization dates of compounds to assess the potential of the above issues and inform the validity of the results. If possible, the opportunity to test freshly solubilized solutions against pre-existing solutions, or testing multiple batches of the same compounds in parallel can be informative.

Figure 3 illustrates in general how screening assays could be organized during the lead optimization phase. The same assay from HTS or an orthogonal assay may be employed for routine testing of compound structure-activity relationship (SAR). Recombinant cell-based assays may still be useful as these offer opportunities to overexpress GPCRs from a variety of preclinical species for efficacy and toxicity models. The use of a null receptor host cell line from which the recombinantly expressed GPCR target derives is essential to ensuring on-target activity. If recombinant cells are chosen for routine SAR studies, it is important to realize that the cellular context (i.e. expression profile of various signaling components) may differ from the cellular context in the targeted organs/tissues *in vivo* either in animals or humans. To promote better *in vitro-in vivo* connection, including a cell line that expresses the targeted GPCR endogenously and that also displays the biological response relevant to the therapeutic indication can be tremendously valuable. If such an endogenous cell-based assay is sufficiently robust, it can be incorporated into the routine SAR. With CRISPR technology, it is now possible to generate a null target cell line, making it more attractive to use endogenous cell lines for routine SAR studies. Otherwise, the endogenous cell-based assay can be layered as a tier 2 assay. A binding assay is also valuable in providing affinity and kinetic assessments. From these screening assays, candidate compounds can be selected for downstream assays, including *in vitro* safety pharmacology assays and ADME assays. Ideal candidates can further be tested in *in vivo* pharmacokinetic models and subsequently efficacy models. The lead optimization phase of a discovery program can be dynamic, such that the screening assays can be positioned in any particular manner as required in accordance to resource availability and the SAR challenges faced by the set of compound series. For example, an identified off-target activity from the *in vitro* safety assay may be added to the top tier routine screening if such activity needs to be mitigated.

During lead optimization, while potency and efficacy values from the screening assays are important metrics to assess the pharmacological properties of the compounds, it is important to realize that such metrics need to be interpreted in the context of the cellular model. Even the use of an endogenous cell line does not guarantee the obtained metrics are translatable directly to *in vivo* context. Again, this is mainly because the cellular context may be different, such as levels of receptor expression or others. Moreover, the conditions employed in the screening assays may not necessarily resemble the *in vivo* context. For example, the level of agonist is usually applied as EC_{50} to EC_{80} concentrations in routine agonist assays, but this may not be the case in the disease state in animals or humans. If it turns out that the agonist is actually present in a very high concentration in the target environment, an expected fall off in potency or even a complete loss of efficacy if the compound turns out to be a weak negative allosteric modulator. Consequently, as outlined below, the potency and efficacy values should not be the only deliverable from lead optimization assays, but the screening assays have to be evaluated from time to time to reveal underlying pharmacological properties of the compounds from different aspects.

Pharmacological Deliverables of HTS and Lead Optimization

The main objective of an HTS campaign is to identify compounds from the screening deck that demonstrate the desired on-target activities. In addition to discovering novel structural scaffolds, HTS provides as the main pharmacological deliverables: potency and efficacy of the various hits. Assays for target selectivity can also be conducted to assess or eliminate cross-reactive compounds. Contrary to HTS, lead optimization screening assays serve to provide a comprehensive knowledge of the *in vitro* compound pharmacological properties during iterative rounds of structural activity relationship studies, the selectivity profile, and the best candidates for *in vivo* testing (Figure 4).

A discovery program may seek to discover a GPCR agonist, antagonist, inverse agonist, allosteric modulator, or biased ligand. An agonist is a ligand that can stimulate (agonize) the GPCR to activate intracellular signaling and trigger a biological response. In contrast, an antagonist is one such ligand that can inhibit (antagonize) the action of an agonist, either natural or synthetic, to suppress the signaling and biological response. An inverse agonist acts similarly to an antagonist in that it can also inhibit receptor response. The term inverse agonist is coined when it is realized that GPCRs may exhibit basal (constitutive) signaling in the absence of an agonist, as a result of conformational dynamics that at any given moment the receptor may adopt an active conformation to trigger biological response. The more favorable the conformational dynamics for the receptor to adopt the active state, or the more abundant the receptor due to high expression level, the higher is the observed basal signaling. A bound inverse agonist can shift the conformational equilibrium to the inactive state, thereby suppressing the receptor response. In fact, antagonists may exhibit inverse agonistic properties. On the other hand, an antagonist that merely occupies the receptor and block the agonist from binding is called a neutral antagonist.

By definition, an allosteric ligand in one that binds outside of the orthosteric site of a GPCR. Pharmacologically, an allosteric modulator acts to potentiate and suppress (modulate) the potency and/or efficacy of the orthosteric agonist. When it potentiates, it is called a positive allosteric modulator (PAM), and when it inhibits, it is called a negative allosteric modulator (NAM). A pure allosteric modulator does not exhibit any agonist activity on its own. If it does, it is sometimes referred as ago-allosteric modulator, such as ago-PAM, in order to describe its dual properties.

As described in the introduction, biased ligands can be regarded as those that can preferentially stimulate/suppress one signaling pathway over another pathway. The behavior of a biased ligand can be rather complex; hence, it may display an "agonistic" activity in one pathway, and if it does not activate another pathway, it may also "antagonize" the activity of an agonist in this other pathway. Due to this potential complex nature, it is important to characterize the compounds for potential bias in as many relevant pathways as possible.

Whether the therapeutic program aims to discover an agonist, antagonist, allosteric modulator, or biased ligand, a repertoire of assays should be established for lead optimization to enhance the predictability for the desirable

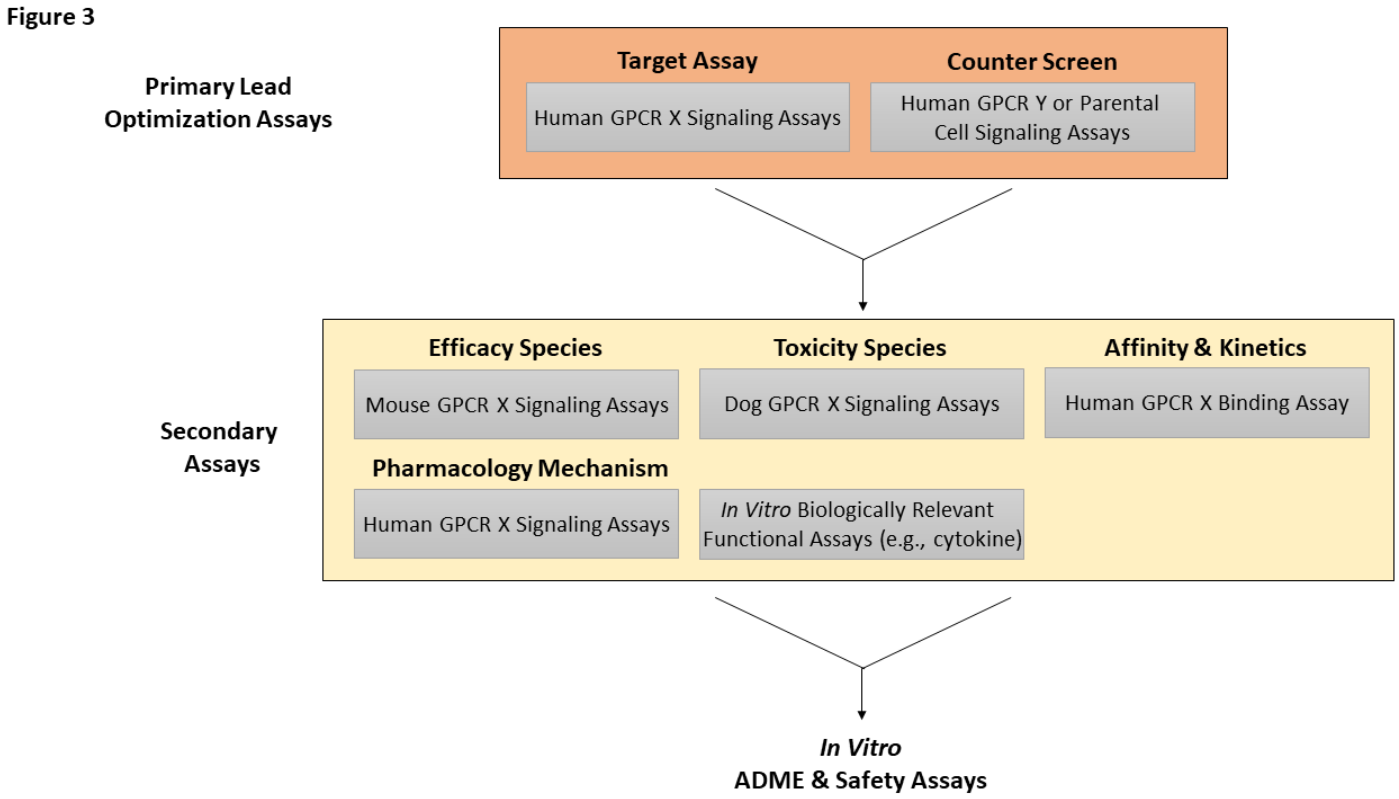


Figure 3. An example of a general lead optimization paradigm for GPCR X and an unrelated GPCR Y as counter screen.

biological outcomes. In addition to potency and efficacy, lead optimization assays should address receptor subtype selectivity and selectivity across other GPCRs, especially closely-related ones. Furthermore, understanding the mechanism of action, such as orthosteric/allosteric interactions, receptor reserve, kinetics, reversibility, and surmountability of competition is important to predict the *in vivo* efficacy. Assessment of compound activity and consistency across animal species is important for translation to downstream *in vivo* testing for pharmacokinetics, efficacy and liability. Functional assays based on the specific G protein coupling or β-arrestin recruitment assays as described above, as well as binding assays, provide potency, efficacy, and affinity determinations that are useful for informing the receptor selectivity and activities across species. In addition, recent literature has indicated the importance of addressing the kinetic aspects of ligand-receptor interaction for GPCR drug discovery (34). The subsequent sections focus on cell-based assay designs that are of particular importance in the lead optimization phase of GPCR drug discovery, while other specific assay technologies are covered in other chapters in this guidance manual.

Important Pharmacological Considerations During Lead Optimization

As lead optimization activities progress, attention should be paid to assess how the model systems and assay conditions may influence the observed pharmacological properties of test compounds, their structure-activity relationships, as well as the translation of their activity from *in vitro* to *in vivo*. These considerations are described and explained in the following sections.

Figure 4

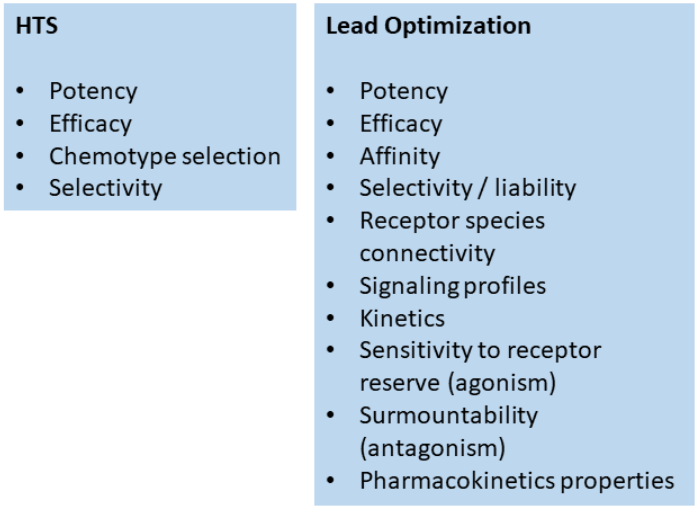


Figure 4. Main *in vitro* pharmacological deliverables of high-throughput screening campaigns and lead optimization of efforts in GPCR drug discovery processes in relationship to the compound activity.

How Receptor Density Affects Agonist Potency and Efficacy Measurements

In assessing agonism, it is very important to recognize how the expression level of the target GPCRs will affect the test agonist's observed potency and efficacy. This is related to the well-known phenomenon of receptor reserve, as described earlier by Black and Leff (35) and re-emphasized recently by others, and has long been recognized to play a significant role in modulating agonist response measurements (36,37). Such influence on potency and efficacy measurements occurs *in vitro* in overexpressing recombinant systems as well as may happen *in vivo* where expression level of receptor is altered between normal and diseased states. Hence, understanding the agonist behavior at various receptor levels can shed light into the efficacy of the agonist *in vivo*.

The potential presence of receptor reserve can be evaluated from a comparison between the ligand affinity measured from a binding assay and the potency values obtained from a functional assay. If the measured affinity value is better than the potency, it is possible that receptor reserve is causing a higher potency than the affinity of ligand binding. From a similar perspective, the receptor occupancy can be computed from the measured binding affinity. If the maximal response observed in a functional assay occurs at a ligand concentration less than 100% receptor occupancy, this would be a sign that receptor reserve is contributing to better than expected functional response. Furthermore, a comparison of the levels of receptor expression across cell lines and tissues that express that GPCR target can provide an idea of the extent of variation of receptor levels in case that receptor reserve may be at work.

When the presence of receptor reserve is suspected, recombinant cell systems can provide a useful model to examine any influence of receptor levels on compound potency and efficacy. There are several means to assess

the effect of receptor expression levels on agonist potency and efficacy parameters. Different clones of primary or stably expressing recombinant cells can be used to assess the relationship between receptor expression and compound agonist activity for any given receptor. A more sophisticated assessment can be done using recombinant cell systems with adjustable levels of receptor expression for better assessment as described below.

Method 1: Transient transfection studies

The plasmid encoding the GPCR can be transfected into the cells at different quantities. The transfection is supplemented with empty vector plasmid or plasmid encoding a different protein (e.g., β-galactosidase) such that the total amount of DNA is the same in the different transfections, so as to maintain the transfection efficiency.

The amount of receptor plasmid should be reduced to a level that the maximal efficacy of the compound is observably lowered, compared to that using the highest quantities of receptor plasmid (i.e., a condition where there is no observable receptor reserve). Any concomitant change in potency should be noted. In addition, receptor expression can be verified by flow cytometry or high-content analysis if a specific antibody for the GPCR or an epitope-tagged antibody is available.

Method 2: Use of cell lines with tunable expression

Host cell lines are commercially available that permit expression levels to be experimentally modulated, such as the T-REx System (Tetracycline-Regulated Expression System) by Invitrogen. In this system, the GPCR expression is constitutively repressed by a tetracycline-sensitive repressor protein binding to an operator sequence in the promoter region of the GPCR-expressing gene. In the presence of tetracycline, the binding of the repressor to the operator is inhibited, hence the GPCR expression is turned on.

Using the T-REx system, receptor expression can be modulated by 1) changing the concentration of tetracycline and harvesting the cells for assays at a fixed time, or 2) using a fixed concentration of tetracycline and harvesting the cells for assays after various times of induction. The expression levels of the receptor can also be measured using antibodies by flow cytometry or in a high-content screening platform. The effect of changing receptor expression levels on test compound potency and/or efficacy can then be assessed.

Figure 5 shows that increasing receptor densities can enhance the observed potency and/or efficacy of test compounds. If activating less than 100% of the tissue/cell receptors are required to achieve maximal efficacy, the tissue/cell system is said to have 'receptor reserve', or 'spare receptors'. In general, as receptor level increases, the functional response become stronger, vice versa. Under conditions of high receptor reserve, for example, in a highly expressing system, compounds with vastly different pharmacological properties can yield exactly the same potency and efficacy. Hence, Compound A with a high intrinsic efficacy but low affinity agonist may yield highly potent and maximal response as much as Compound B with a low efficacy and high affinity agonist (Figure 5A & 5B, red curves). However, when the receptor level drops, the high intrinsic efficacy of Compound A allows it to maintain strong potency and response even when the receptor level drops significantly (Figure 5A, blue curve). In contrast, when there is a significant drop in receptor density, Compound B show significantly declined activity (Figure 5B, blue curve). The potency and efficacy behavior under different levels of receptor expression can differ greatly depending on the underlying pharmacological properties of the compounds.

Further to this notion is that, in using a highly overexpressed system for lead optimization, it is difficult to unequivocally determine whether a compound structural modification is driving an improvement in affinity or intrinsic efficacy, or both, as changes in one of both properties can result in an increase in apparent compound potency. Situations may arise that an optimized high affinity but low efficacy agonist may show significant decrease in potency and efficacy in a non-overexpressed system (e.g. *in vivo*), in certain disease states when receptor expression is lower, or when the receptor become desensitized by continuous agonist exposure that render the agonist non-efficacious for therapeutic treatment.

Questions to consider:

- Is the observed potency in a functional assay better than the measured ligand affinity?
- Is the maximal response attainable with less than 100% receptor occupancy?
- Is the agonist activity driven by potency or efficacy?
- Is the agonist potency maintained under low receptor density?
- Is the agonist efficacy maintained under low receptor density?
- Are there primary cells available to assess agonist activity in relationship to receptor expression level?
- Are there different clones of a recombinant cell line to assess agonist activity in relationship to receptor expression level?
- Is there any receptor specific antibody or epitope tag available on the receptor?
- Is there a transfectable cell line that can be used to modulate expression of the receptor?

How the On- and Off-Rates of the Test Agonist (Compound Kinetics) Affect Compound Functional Activity Measurements and Thereby Impact Lead Optimization

The methodologies of setting up assays to measure the kinetics of compound association and dissociation to GPCRs are covered in the AGM chapter Receptor Binding Assays for HTS and Drug Discovery. Binding kinetics can play a significant role in the observed compound potency measured in functional assays, especially in transient kinetic-based signaling assays, such as in the measurement of calcium flux (38). Consequently, when evaluating the contribution of binding kinetics, it can be useful to assess the agonist response using assays that the readout sustains over the time course of the experiment, for example, cAMP accumulation assay for G_s and G_i-coupled receptors, or IP1 accumulation assay for G_q-coupled receptors. Transient signaling readouts such as calcium flux are captured over short time periods of time (<2 min) and therefore can be significantly affected by the test compound binding kinetics (38). Therefore there is a potential that the calcium transient being restored to resting level before the agonist can fully bind to the receptor at equilibrium.

When instrumentation and assay technologies are available that allow real-time measurement of receptor response, the functional consequence of binding kinetics can be assessed readily by continuous or time-lapse measurements. Examples of this type of real-time assays are the signaling biosensors (39-41). In many cases, such real-time assays may not be available, and the response assessment can be performed as described below.

Assessing the Influence of Binding Kinetics on Measured Agonist Response

Figure 6A–F illustrates how compound binding kinetics may affect the measured responses for three different agonists. The three agonists exemplify a scenario in an optimization effort that leads to improved affinity from one to another, due to the reduction in the dissociation rates. One of the agonists has lower affinity and potency (10 nM), and associates rapidly with the receptor (k_{off} = 0.04 min^{-1}) (Figure 6A). The second agonist has an intermediate affinity and potency (1 nM), and associates with the receptor at a moderate rate (k_{off} = 0.004 min^{-1}) (Figure 6B). The third agonist has the highest affinity and potency (0.1 nM), but associates slowly with the receptor (k_{off} = 0.0004 min^{-1}) (Figure 6C). The observed association rate (k_{obs}) is related to the ligand concentration (L) and both the k_{on} and k_{off} as described by the equation:

$$k_{obs} = k_{off} + k_{on}(L)$$

Re-expressing the equation in terms of the equilibrium dissociation constant (K_d) results in the following equation:

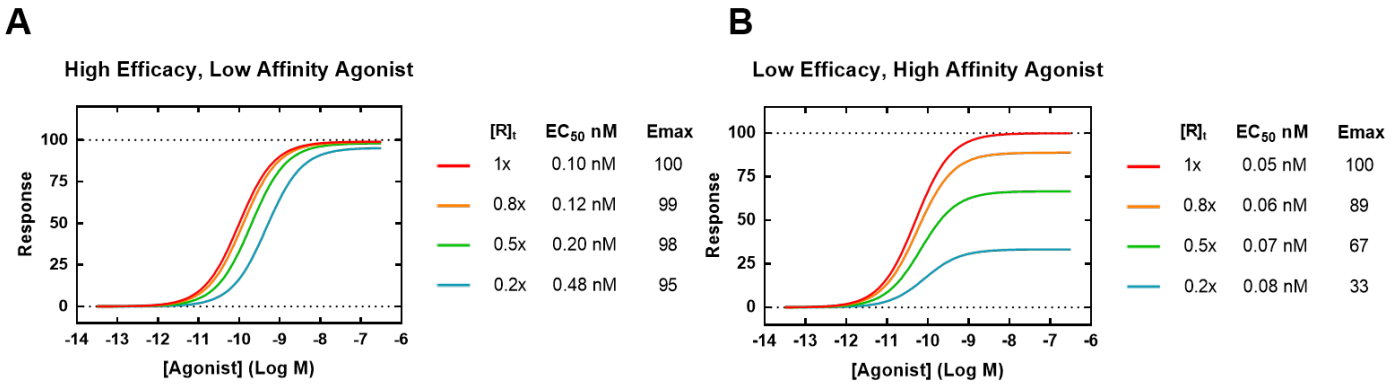


Figure 5. Effect of changing receptor density on response of (A) a high efficacy agonist with comparatively low affinity (K_A = 10 nM) and (B) a lower efficacy agonist with high affinity (K_A = 0.1 nM). $(R)_t$ % = relative receptor density. Simulations are carried out in accordance to the Operational Model (35).

$$k_{obs} = k_{off} + (k_{off} \,/\, K_d)(L) = k_{off}(1 + (L) \,/\, K_d)$$

It is worth noting from the above expression that the overall observed association rates of the agonists are highly dependent on the k_{off} when (L) ≤ K_d, that is when the agonist is present at a concentration less than the K_d value. In other words, as a result of a reduced dissociation rate through optimization, the agonist being tested will take longer to equilibrate with the receptor (compare Figures 6A to 6C).

During assay development, while an optimal period of incubation may be determined using any tool compounds available at the time, the tool compound can be one such as the first agonist above. Thus, it would have suggested that a 30 min incubation is adequate (Figure 6D). Nevertheless, as the affinity improves and the dissociation rate constant of the compound decreases, it is important to realize that the potency may be underestimated due to the much longer time needed to attain equilibrium (Figures 6E and 6F). The results could lead to difficulty in the interpretation of the true structure-activity relationship (SAR). It is therefore essential to regularly evaluate the influence of binding kinetics from time to time when SAR is unclear, kinetic contribution is suspected or as the potency of agonist compounds matures.

The contribution of binding kinetics can be periodically assessed by performing the agonist concentration-response assays for increasing periods of incubation and finding out if the compounds may give higher observed potency (Figure 6D–F). Figure 6G shows a plot of the observed potency (EC_{50}) as a function of time t for the same three agonists. The observed time-dependent increase in measured potency reveals the underlying influence of binding kinetics on observed potency and calls for the need to revise the assay procedure to accurately capture the potency values of slow binders.

Assessing the Influence of Binding Kinetics on Antagonist Response

When antagonists are tested, their activity is measured in the presence of an agonist that is used to trigger the cellular response. The kinetics of association of the test compounds will influence the observed antagonistic activity in the mutual competition with the agonist during the approach to equilibrium. Figure 7A–F explains how binding kinetics may affect the measured responses for three antagonists. They exemplify the scenario that through lead optimization, the affinity of antagonist test compounds increases due to a reduction the dissociation rates. The first antagonist has a lower affinity (K_d = 10 nM) and associates rapidly with the receptor (k_{off} = 0.04 min^{-1}) to a similar extent as the agonist trigger (Figure 7A). The second antagonist has an intermediate affinity (K_d = 1 nM) and associates with the receptor at a moderate rate (k_{off} = 0.004 min^{-1}) (Figure 7B). The third antagonist has the highest affinity (K_d = 0.1 nM) but associates slowly with the receptor

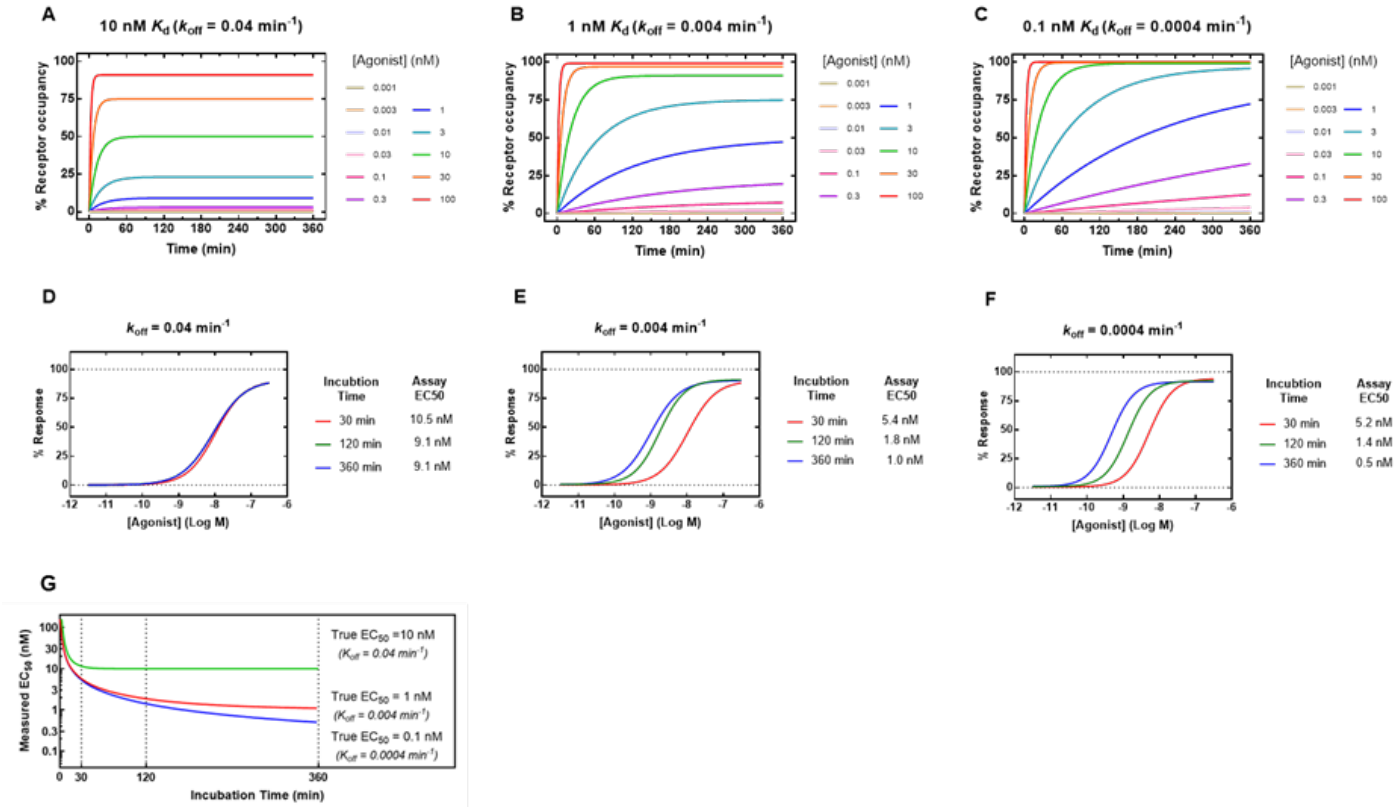


Figure 6. Influence of binding kinetics on the measured agonist potency and the assessment of time-dependent agonist response. Figures 6A–C. Time-dependent receptor occupancy of agonists with the EC_{50} value and equilibrium dissociation constant (K_d) of 10, 1 and 0.1 nM and corresponding dissociation rate constant (k_{off}) of 0.04, 0.004 and 0.0004 min^{-1} as indicated on the graphs. Figures 6D–F. Concentration-response relationship of the three agonists measured after 30 min, 120 min and 360 min incubation and the potency values determined at the corresponding time point showing in the associated legends. Figure 6G. A plot of observed potency (EC_{50}) as a function of time to assess the time-dependent agonist response and influence of binding kinetics on observed potency.

(k_{off} = 0.0004 min^{-1}). It is important to note that it takes longer to attain equilibrium when the off-rate is slower, based on the same notion as described in the above section on agonist kinetics. The antagonist assay is usually carried out using an agonist concentration that by itself triggers 50% of the maximal response (*i.e.*, EC_{50} concentration). In this example, the agonist has an EC_{50} of 10 nM and is used at 10 nM concentration. If the initial assay incubation condition was established using the first antagonist, which is a fast binder, it would have suggested that a 30-min assay incubation is sufficient. While this is true for a rapidly associating antagonist, as the affinity of the test compound improves, the off-rate may possibly decrease; as a result, the inhibitory potency of the potent antagonist might be under-estimated (Figure 7D and 7F). While these simulations use an agonist that equilibrates comparatively faster, it is important to recognize that the issue is largely governed by the kinetics of the antagonists themselves but is exacerbated by an agonist that equilibrates much faster than the antagonists.

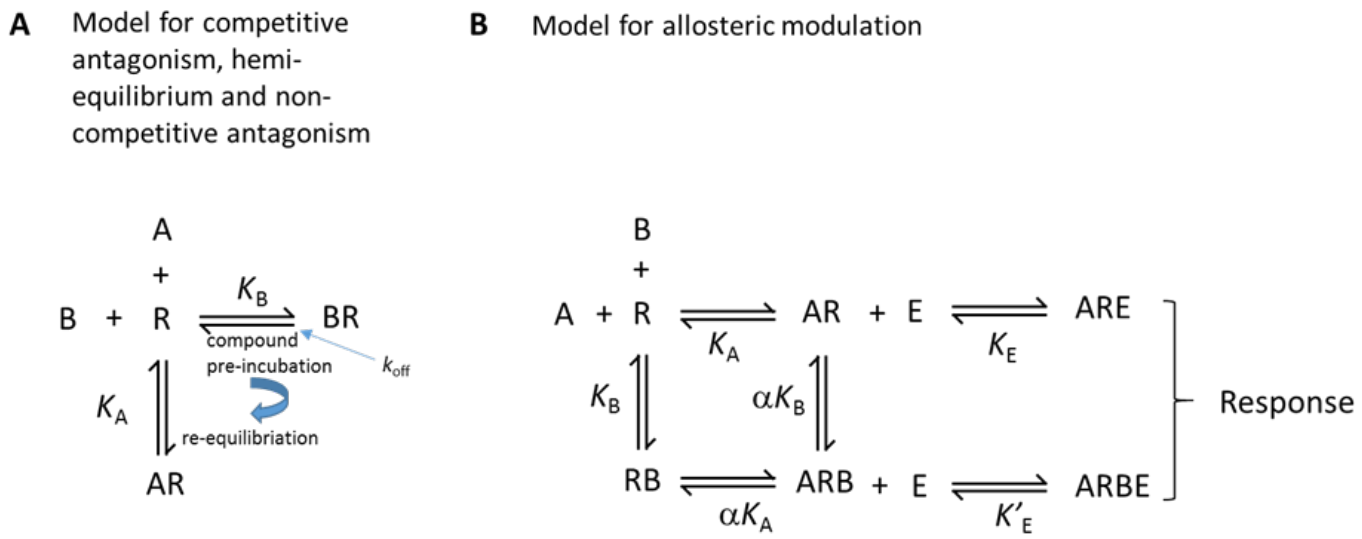
The contribution of binding kinetics to antagonistic activity can also be assessed by performing concentration-response assays in the antagonist mode after increasing the periods of incubation time, presumably when the antagonist potency improves to a certain degree (Figure 7D–F). Figure 7G shows a plot of the observed potency (IC_{50}) as a function of time t for the three different antagonist with differing binding kinetics. The observed time-dependent increase in measured antagonist potency reveals the underlying influence of binding kinetics on the observed antagonist potency and indicates the need to revise the assay procedure, in order to capture accurately the potency values of slow binders.

Questions to consider:

- Is the assay readout transient, cumulative or sustained?
- Do the measured potency values of the test compounds appear to reach a limit unanticipated from the SAR?
- Is low off-rate suspected of the compounds?
- Are there unexplained significant differences between cell-based antagonist potency and binding affinity values?
- Can the off-rates of compounds be determined from an alternative assay to estimate the necessary equilibration periods?
- For antagonist assays, is there a possibility that the agonist associates comparatively faster than the test antagonists?
- Does extending the equilibration time make the compound appear more potent?

Dose-Shift Assays

Dose-shift assays are powerful assays that offer diagnostic information for assays of antagonistic and allosteric function, and they provide information on the underlying mechanism of action of the test compounds in relationship to the agonist. In dose-shift assays, a full agonist concentration response curve is carried out in the absence and in the presence of a series of increasing fixed concentrations of test compound. The "dose-shift" is the ratio of the EC_{50} of the agonist in the presence vs. absence of test compounds. Dose-shift assays can reveal the presence of hemi-equilibrium conditions for orthosteric antagonism, or non-competitive antagonism. Moreover, the assay format also has the ability to differentiate orthosteric interaction from allosteric modulation. The models and equations used in describing these types of behavior are thoroughly discussed in Kenakin *et al.* (42).



Legend: A, agonist; B, test compound (antagonist or allosteric modulator); R, receptor; AR, agonist-receptor complex; BR, antagonist/modulator-receptor complex; ARB, ternary complex of agonist, receptor and allosteric modulator; E, effector of signaling; ARE, complex of agonist, receptor and effector; ARBE, complex of agonist, allosteric modulator, receptor and effector; K_A, agonist equilibrium dissociation constant; K_B, antagonist/allosteric modulator equilibrium dissociation constant; k_{off}, dissociation rate constant of antagonist; α, allosteric factor; K_E and K'_E, equilibrium constants for response coupling

The schematics for competitive antagonism, hemi-equilibrium, non-competitive antagonism (A), and allosteric modulation (B)

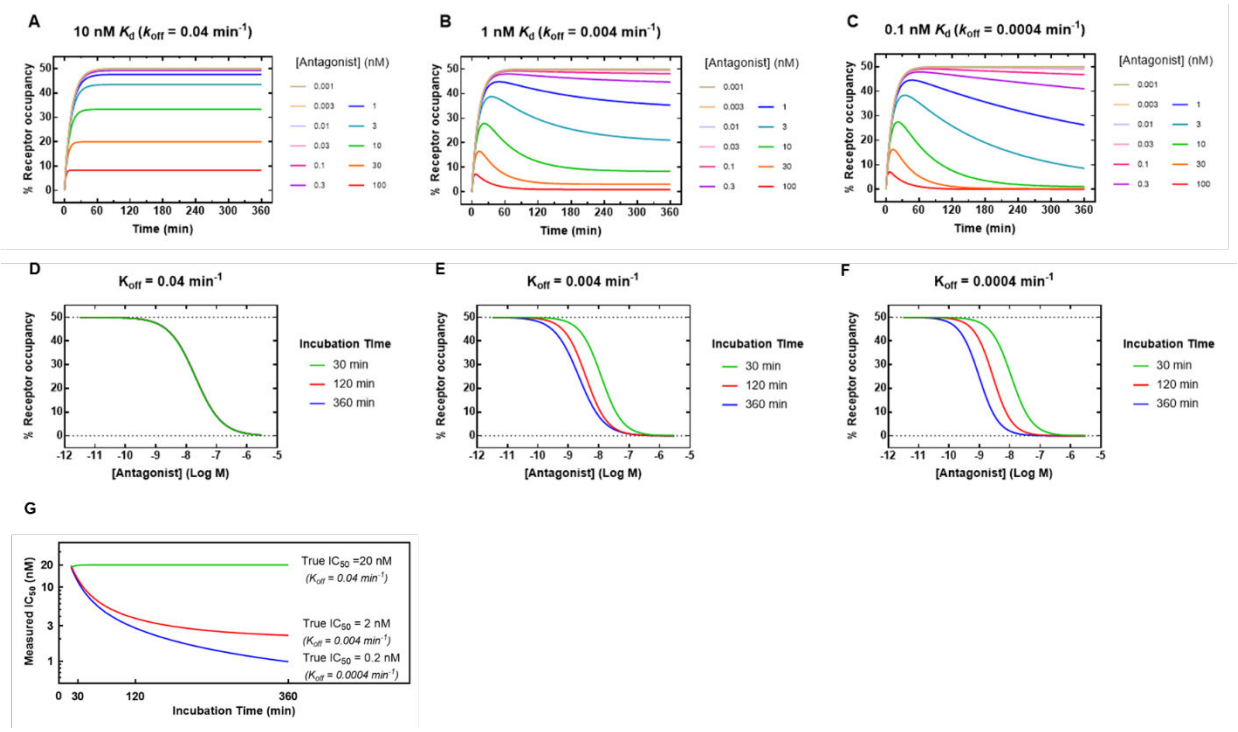


Figure 7. Influence of binding kinetics on the measured antagonist potency and the assessment of time-dependent antagonist response. The agonist used for stimulation has an EC_{50} of 10 nM and a concentration of 10 nM. Figure 7A–C. Time-dependent receptor occupancy of antagonists with the equilibrium dissociation constant (K_d) of 10, 1 and 0.1 nM and corresponding dissociation rate constant (k_{off}) of 0.04, 0.004 and 0.0004 min^{-1} in the co-presence of the agonist, as indicated on the graphs. Figure 7D–F. Concentration-response relationship of the three agonists measured after 30 min, 120 min and 360 min incubation and their corresponding potency (IC_{50}) values. Figure 7G. A plot of observed potency (IC_{50}) as a function of incubation time to assess the time-dependent antagonist response and the influence of binding kinetics on the observed potency values.

The equations for agonist response in the presence of an antagonist with consideration of off-rate of the antagonist is given as described in Riddy *et al.* (43):

$$Response = \frac{[A]/K_A\tau\left(1 - \bullet\left(1 - e^{-k_{off}\bullet\gamma\bullet t}\right) + \beta\bullet e^{-k_{off}\bullet\gamma\bullet t}\right)\tau E_{max}}{\frac{[A]}{K_A}\left(\left(1 - \left(\left(1 - e^{-k_{off}\bullet\gamma\bullet t}\right) + \beta\bullet e^{-k_{off}\bullet\gamma\bullet t}\right)\right)\tau + 1\right) + 1}$$

where

$$= \frac{[B]/K_B}{\frac{[B]}{K_B} + \frac{[A]}{K_A} + 1}$$

$$\beta = \frac{[B]/K_B}{\frac{[B]}{K_B} + 1}$$

$$\gamma = \frac{\frac{[B]}{K_B} + \frac{[A]}{K_A} + 1}{\frac{[A]}{K_A} + 1}$$

τ = operational efficacy of the agonist and E_{max} = the maximal system response.

The equation for agonist response in the presence of an allosteric modulator is shown below as described in Kenakin *et al.* where α = allosteric factor, $\xi = \tau'/\tau$, τ = operational efficacy of the agonist in the absence of

modulator, τ' = operational efficacy of the agonist in the presence of modulator, E_{max} = the maximal system response (42):

$$Response = \frac{[A]/K_A\tau\left(1+\frac{[B]}{K_B}\right)E_{max}}{\frac{[A]}{K_A}\left(1+\frac{[B]}{K_B}+\tau\left(1+\frac{[B]}{K_B}\right)+\frac{[B]}{K_B}+1\right.}$$

A dose-shift assay can be set up by performing a titration of the test compound and the agonist in a two-dimensional matrix format as shown in Figure 8. Serial dilutions of the test compounds, which can be orthosteric antagonists, negative allosteric modulators or positive allosteric modulators, are first pre-incubated with the cells, and then the cells are subsequently stimulated with the agonist that is serially diluted in the perpendicular direction. This generates a series of agonist concentration response curves carried out in the presence of, in this case, twelve fixed concentrations of test antagonist.

Figure 9 shows several possible patterns of behavior that could be observed from the results of such a dose-shift assay when the agonist concentration-response is plotted for each corresponding concentration of test compound. This helps identify their mode of interaction with the agonist, whether the interaction be competitive or non-competitive, or whether the interaction is allosteric or orthosteric in nature.

Orthosteric antagonism is seen as a parallel rightward shift of the agonist curves (Figure 9A). However, if the agonist curves are continuously showing a rightward shift, and at the same time the maximal response achievable by the agonist was depressed and converged to a lower level, this is indicative of a hemi-equilibrium condition (Figure 9B). Hemi-equilibrium occurs because the antagonist dissociates relatively slowly during the time-course of the experiment such that the inhibition becomes apparently insurmountable due to incomplete re-equilibration. Hemi-equilibrium conditions can occur in screening assays, as the off-rate of the test compound may become lower as the affinity of the compounds improves throughout the optimization process. The slower the off-rate of the antagonist, the more depressed the maximal response triggered by the agonist becomes. Hence, hemi-equilibrium offers a means to assess the off-rates of the test compounds either qualitatively or by fitting the hemi-equilibrium operational model as described in (43).

When the test compound binds so tightly to the receptor that it does not appreciably dissociate over the time course of the assay incubation (i.e., pseudo-irreversible), or when the test compound binds irreversibly to the receptor via covalent bonds, non-competitive antagonism results. Whether the test compound is orthosteric or allosteric, the effect is a complete suppression of maximal agonist response to basal level at high concentrations of the test compound (Figure 9C).

Allosteric modulations can be rather complex. An allosteric modulator can affect the affinity and/or efficacy of the orthosteric agonist independently. Figures 9D–F depict the functional effect of an allosteric ligand that decreases the affinity of the orthosteric agonist, while having no effect on agonist efficacy (Figure 9D), reducing the efficacy (Figure 9E), or enhancing the efficacy (Figure 9F). Figures 9G–I show the functional effect of an allosteric ligand that increases the affinity of the orthosteric agonist, while at the same time, having no effect on agonist efficacy (Figure 9G), decreases the efficacy (Figure 9H), or enhancing the efficacy (Figure 9I).

An allosteric modulator that simply decreases or increases the affinity of the orthosteric agonist produces a rightward shift or leftward shift respectively in the potency of the agonist in the presence of increasing fixed concentration of the test allosteric compound (Figure 9D and 9G). The extent of the shifts depends on the strength of the allosteric interaction, as described by the allosteric factor (α) in the model. An important feature of allosteric modulation is that there is a "ceiling" effect on this shifting behavior. Figure 9D shows the agonist curves carried out in the presence of increasing fixed concentration of an allosteric modulator that decreases agonist affinity, the EC_{50} for the agonist shifts rightward but stops at about 30 nM. Further increases in concentrations of the modulator fail to shift the curve further after all the allosteric binding sites are filled. Analogously, figure 7G shows the agonist curves carried out in the presence of increasing fixed concentration of

an allosteric modulator that increases agonist affinity, the EC_{50} for the agonist shifts leftward but stops at about 0.03 nM. Further increases in concentrations of the modulator fail to shift the curve further.

In addition to modulating the affinity, the allosteric ligand may or may not have an effect on the functional efficacy of the agonist. The effect on functional efficacy is reflected by a change in the maximal response of the agonist (Figure 9E, F, H and I). For an allosteric modulator that decreases agonist affinity and reduce its efficacy (Figure 9E), the agonist curves show a successive reduction in the maximal responses at the same time that the curves shift to the right. For an allosteric modulator that decreases agonist affinity but enhances its efficacy (Figure 9F), the mixed behavior can cause a rightward shift in potency that is counteracted by the increase in the efficacy of response. The results can be a diminished rightward shift, compared to Figure 9D, depending on the extent of efficacy enhancement. With an allosteric modulator that increases affinity while also reduces efficacy of the agonist, the opposing behavior may results in agonist curves as shown in Figure 9H. The potency enhancement becomes less prominent due to the negative effect on agonist efficacy, while at the same time the maximal respond reduces successively. In contrast, with an allosteric modulator that increases affinity while also enhances efficacy of the agonist, the combined effect may result in agonist curves as shown in Figure 9I, in which the potency enhancement is greater than by the effect on affinity alone.

Comparing Figure 9B and 9E, it is interesting to note that the pattern of dose-shift and maximal response change for a negative allosteric modulator can be indistinguishable to that seen with orthosteric antagonists under certain hemi-equilibrium conditions. Nevertheless, an important property of allosteric modulation is probe dependence, meaning that the degree of allosteric interaction between allosteric modulator and an agonist can be different with another agonist, whereas, the hemi-equilibrium effect from orthosteric antagonism should not change from one agonist to another. By comparing the behavior using multiple agonists, the nature of the underlying interaction can be identified.

Questions to consider:

- Is the shift of agonist potency values parallel and reaching the same maximal response at all antagonist concentrations?
- Does the shift of agonist potency values reach a limit?
- Is the pattern of potency shift caused by a test antagonist different with another agonist of similar potency (*i.e.*, probe dependent)?
- Is the maximal agonist response converging to a lower plateau at higher antagonist concentration?
- Is the maximal agonist response completely inhibited at high antagonist concentration, even in the presence of high agonist concentration of (*i.e.*, insurmountable)?
- Is the allosteric modulator affecting the agonist potency, maximal agonist response, or both?

Dose-Shift Assays in Screening Mode

Dose shift experiments require a significant number of assay wells as multiple agonist concentrations curves are constructed. In either an HTS screen or lead optimization campaign, it is more economical and convenient to assess pharmacological properties of a large number of antagonists or modulators by assessing the ability of increasing concentrations of test compounds in the presence of three to four fixed concentrations of agonist. The resultant EC_{50}/IC_{50} from this assay setup can be used to assess the mode of interaction. The setup of such a workflow is shown in Figure 10. Prior to running the assays, the concentration-response of the agonist alone is assayed, to determine concentrations that give rise to 20%, 50% of the full response (i.e., EC_{20}, EC_{50}, EC_{80}). Subsequently, the desirable agonist concentrations, such as EC_{20}, EC_{50}, EC_{80}, or 100X EC_{50}, can be applied as the agonist trigger in the assay setup.

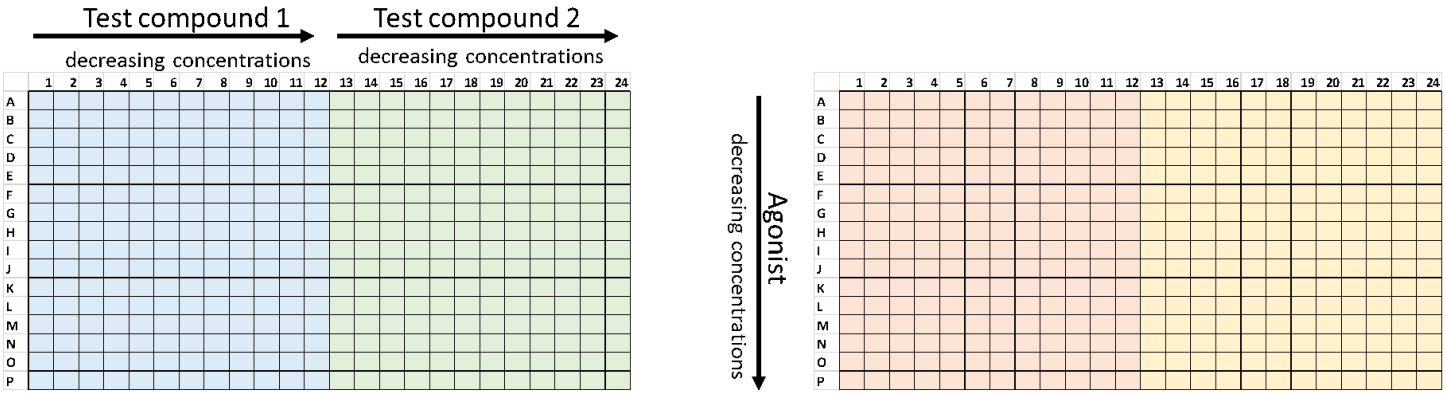


Figure 8. Compound plate setup for dose-shift assays. A 384-well plate can accommodate two different test compounds. The test compounds are serially diluted from left to right, and the same dilution series placed in every row from top to bottom. The agonist used for stimulation is diluted in another 384-well plate from top to bottom, and the same dilution series placed in every column from left to right.

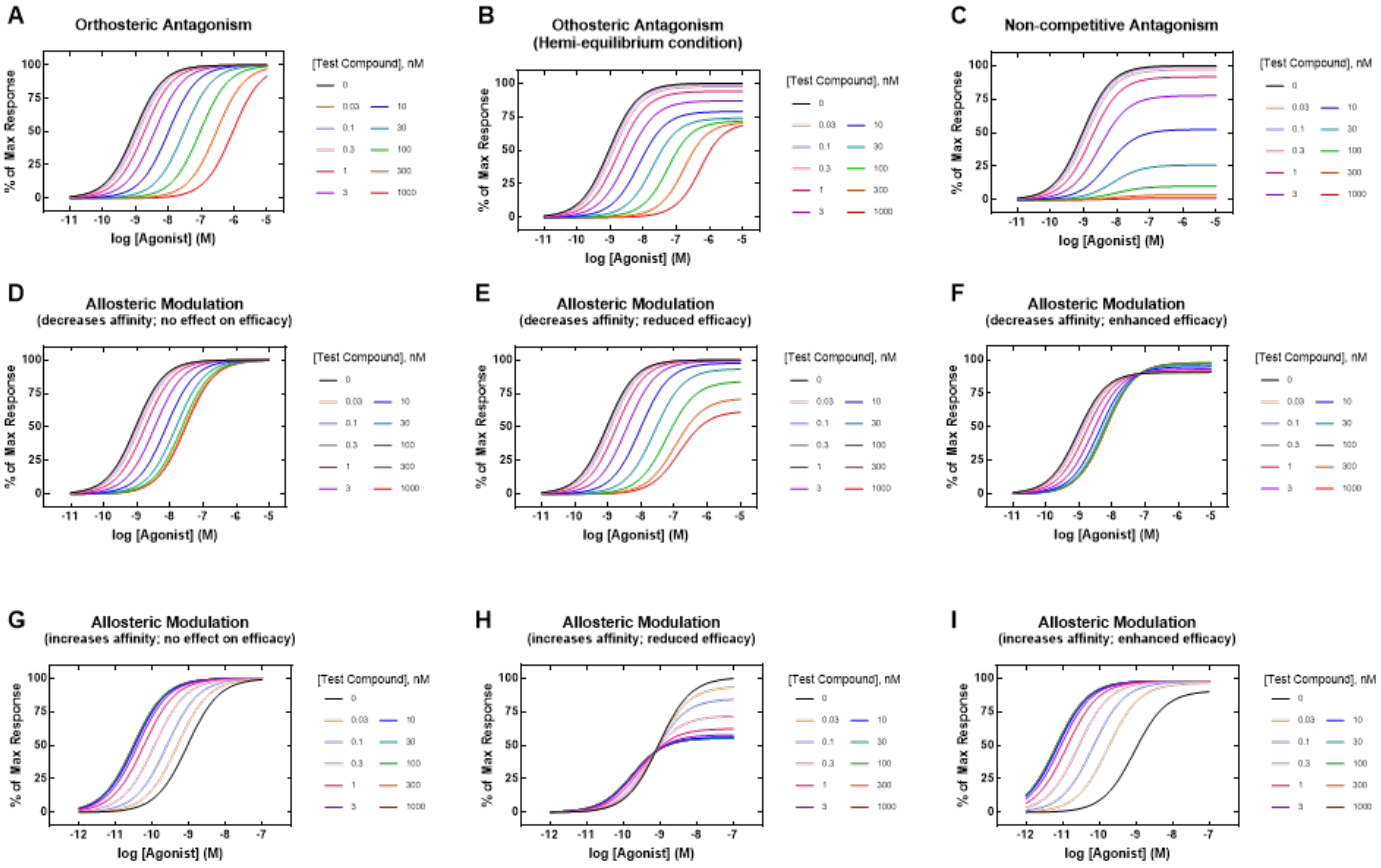


Figure 9. Characteristic patterns of agonist concentration curves in a dose-shift assay in the presence of different pharmacological agents. Figure 9A–C: (9A) competitive antagonist under equilibrium condition, (9B) competitive antagonist in hemi-equilibrium, and (9C) non-competitive antagonists/modulators. Figures 9D–F: Allosteric modulator that decreases affinity which does not affect agonist efficacy (9D), reduces agonist efficacy (9E), or enhances efficacy (9F). Figures 9G–I: Allosteric modulator that increases affinity which does not affect agonist efficacy (9G), reduces agonist efficacy (9H), or enhance efficacy (9I).

Figure 11 shows the pattern of behavior of the same pharmacological agents when assessed by this screening mode using three to four concentrations of agonists. Screening for antagonist or negative allosteric modulators are usually carried out using the EC_{50} or EC_{80} concentration of the agonist. Figure 11A shows the typical competitive antagonism behavior in which the IC_{50} values are shifted proportionally to the applied agonist trigger concentrations. Non-competitive antagonism clearly demonstrates a lack of or non-proportional IC_{50} shift (Figure 11C). Under hemi-equilibrium condition as shown in Figure 11B, the inhibition curves may resemble something in between competitive antagonism and non-competitive antagonist, dependent on the extent of re-equilibration during the agonist stimulation step. It is important to note that with low concentrations of agonist, the IC_{50} values of the test compounds are similar (Figures 11A-C, red curves), whether they are competitive, under hemi-equilibrium conditions, or non-competitive inhibitors. Using a fewer concentrations of agonist and as high as possible, one can increase the diagnostic power of the assay, as shown with the blue curves (Figure 11A, Figure 11B, Figure 11C, blue curves).

Using such setup, negative allosteric modulations can be shown to display a notable pedestal effect (Figure 11D, green and blue curves) depending on the strength of the allosteric interaction (α). The weaker the allosteric interaction, the more pronounce is the pedestal, which also occurs more likely with a higher concentration of agonist. As noted in the above section, due to the resemblance of negative allosteric modulation to certain hemi-equilibrium conditions, follow-up experiments with different agonists should be pursued to identify any probe dependence, which may be used to distinguish allosteric interaction from behavior owing to hemi-equilibrium condition. For positive allosteric modulators, the EC_{20} concentration of the agonist is usually used to assess the potentiation effect of the test compounds and to maximize the ability to observe the potentiation effect (Figure 11G). Although it may be easier to detect allosteric modulations that affect only the agonist binding affinity as described above, modulators that affect also the agonist efficacy can result in a complicated IC_{50} curves. Figure 11E, Figure 11F, Figure 11H and Figure 11I show examples of each of these possible behaviors. When it becomes unclear, it is best to perform follow up experiment in a full dose-shift assay.

Questions to consider:

- Is there a right shift of compound potency values at high levels of agonist concentration (e.g., 100X EC_{50})?
- Is the compound concentration reaching a pedestal above the baseline?
- Is the compound behavior at lower agonist concentrations opposite to that at high agonist concentrations?

Biased Signaling

Due to the potential for signaling bias, modern day GPCR drug discovery cannot simply rely on a single pathway or assay readout to progress compounds through the discovery pipelines. The signaling properties of the test compounds should be assessed with as many relevant pharmacological assays as possible. The screening assays for biased ligands are not different from traditional pharmacological assays as described above. However, as suggested by Charlton and Vauquelin (38), careful consideration should be paid to the presence of pre-equilibrium conditions as noted above; otherwise, incorrect determination of potency values will misrepresent signaling bias.

The selection of compounds with the desirable signaling profiles or bias requires a means to represent such profiles. A number of methods with different pros and cons have been described in the literature (9,44-46). A simple visualization of the effect from two pathways can be done in a model-independent manner as described by Kenakin and Christopoulos (9). The Bias Plot, as illustrated in Figure 12, is constructed by taking the responses triggered by a given compound at all tested concentrations and putting them on the x and y coordinates of the Cartesian coordinates and connecting the coordinates by a line or curve. The plot can be overlaid with the plots obtained with other compounds to compare their signaling profiles.

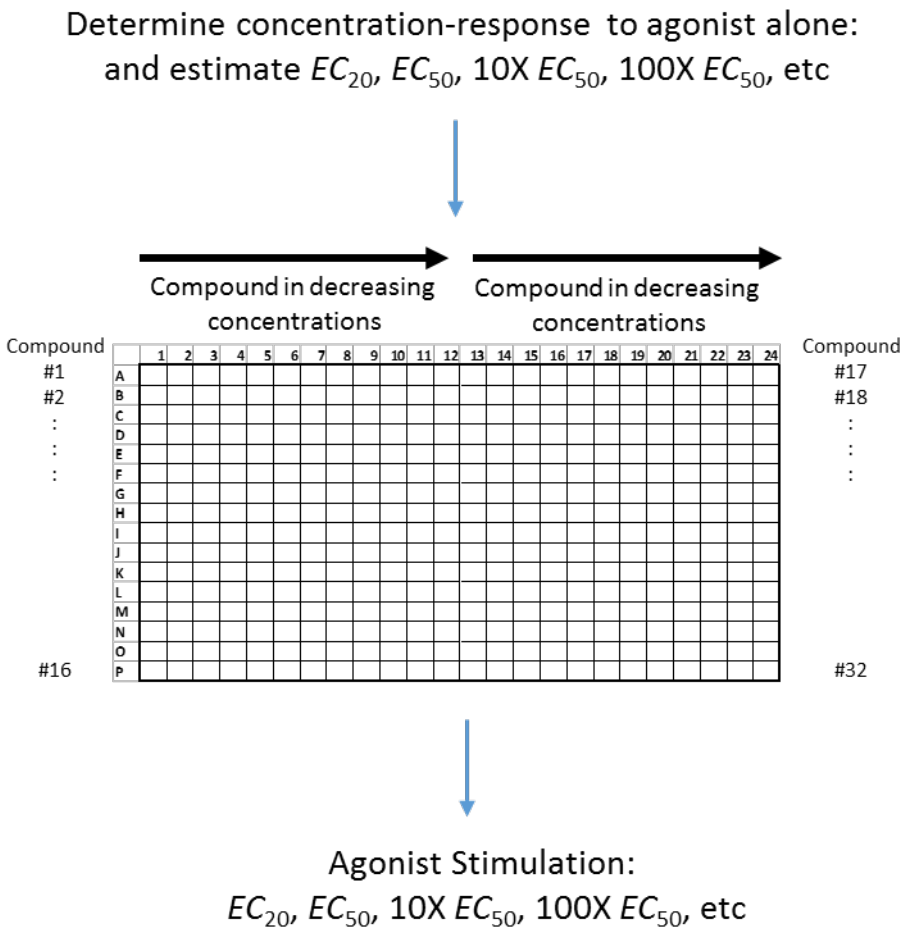


Figure 10. An example for setting up dose-shift assays in antagonist mode.

To visualize the signaling profiles across all test compounds, when more than two pathways are compared, a radar plot can be constructed in which each vertex on the plot is taken from the quantified compound activity (e.g., relative activity, transduction ratio) from each assay readout. The profiles can be compared across compounds by overlaying the plot obtained for each compound (Figure 13).

To quantify the signaling bias, the two most commonly used methods are the relative activity calculation (45) or the more sophisticated operational model (9,44). With the relative activity model, the compound activity in any given pathway is expressed as the ratio of maximal response and potency value. With the operational model, the compound activity is expressed as the transduction ratio (τ/K_A), which is obtained upon fitting the operational model to the data. Figure 14 shows a general scheme for the process to determine the signaling bias across test compounds. These quantified parameters for all compounds can then be used for the radar plot to visualize similarities or differences in their profiles. Alternatively, the activity of a given compound relative to the reference compound can be compared by subtracting the logarithmic values (Δlog) to obtain the ligand bias in each pathway. Additionally, the logarithmic values from a signaling pathway can be compared against a particular pathway by taking the difference between the pathways for every ligand ($\Delta\Delta$log). The resulting values will be quantified signaling or pathway bias.

Lefkowitz and his colleagues reviewed in detail the considerations for selecting assays to best detect signaling bias, such as receptor reserve, signal amplifications, proximity of the readout to the receptor, expression levels of the effectors leading to measured readouts (47). High receptor reserve, strong signal amplification or high expression levels of signaling intermediates can drive up the observed activity the compounds and may make it

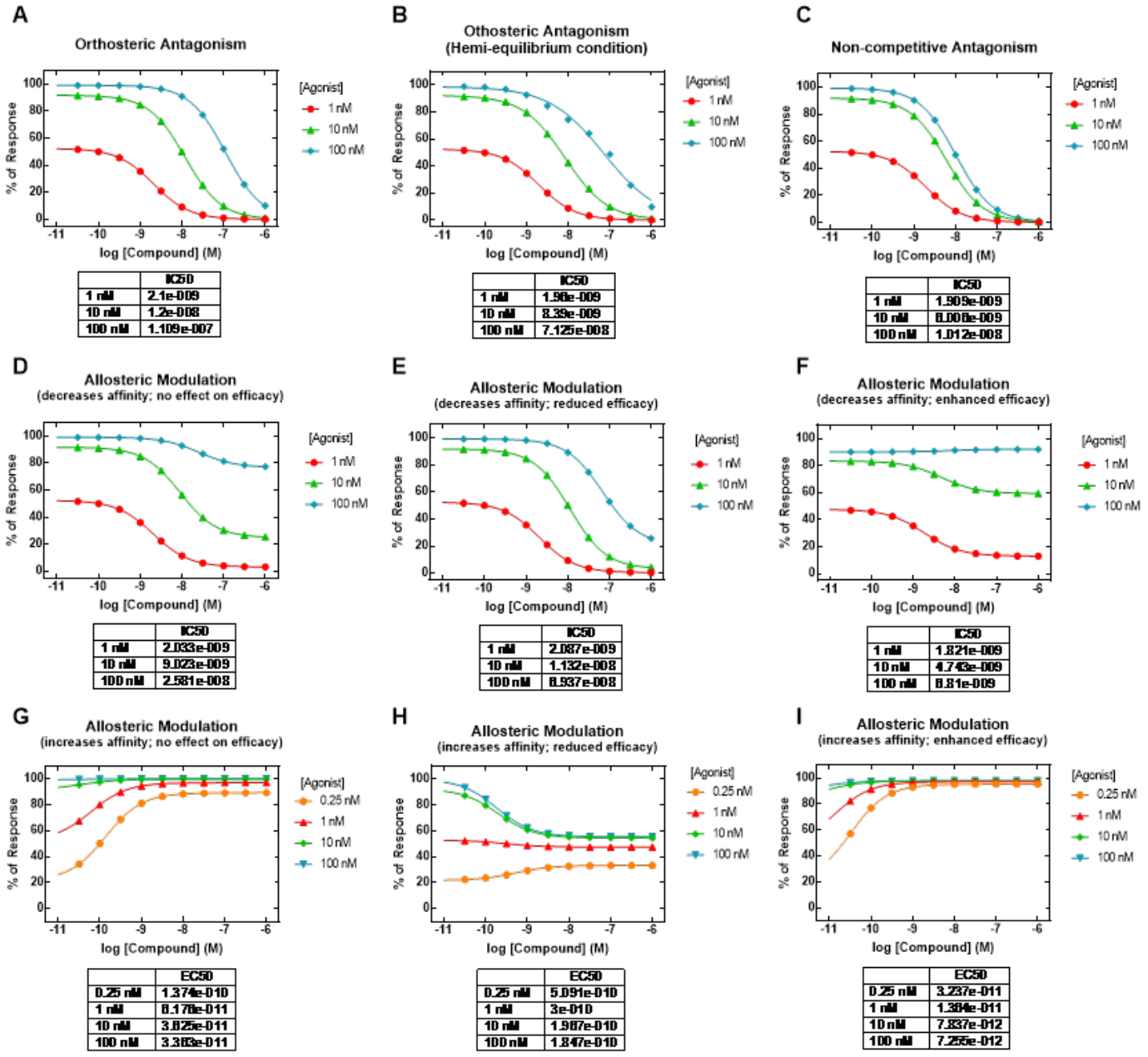


Figure 11. Patterns of behavior of the same pharmacological agents given in Figure 9 assessed by their concentration-response in the presence of three to four agonist concentrations. Figures 11A–C: competitive antagonist under equilibrium condition (11A), competitive antagonist in hemi-equilibrium (11B) and non-competitive antagonists/modulators (11C). Figures 11D–F: Allosteric modulator that decreases affinity which does not affect agonist efficacy (11D), reduces agonist efficacy (11E), or enhances efficacy (11F). Figures 11G–I: Allosteric modulator that increases affinity which does not affect agonist efficacy (11G), reduces agonist efficacy (11H), or enhances efficacy (11I).

difficult to discern biases. Ideally, it is best to choose assays with similar levels of amplification, or at least a lower amplification assay to detect the bias for the particular pathway of interest. Therefore, assay technologies that measure the most proximal response such as the BRET-based G protein and arrestin biosensors will be the most valuable in this regard (18).

From a drug discovery perspective, it is important to ensure that the observed *in vitro* bias translates to the *in vivo* situation. This involves more than simply rank ordering the compound potencies as now more than one

signaling function is compared. Transgenic knock out studies can be useful for determining what signaling readouts are relevant to the *in vivo* biological response (e.g., G proteins vs. β-arrestin). Often, the actual signaling effect that mediates the *in vivo* response can be difficult to dissect or the response may be mediated by multiple effectors. Moreover, the cellular context of the *in vitro* assay is unlikely going to be the same as *in vivo*. Consequently, instead of determining the "right" *in vitro* bias, it may be more useful to tune the *in vitro* assays to reflect as close as possible to the *in vivo* observations, for example, by altering the levels of receptor and/or effectors, to establish the *in vitro-in vivo* connectivity for the desired level of bias.

Questions to consider:

- What is the signaling effector responsible for the *in vivo* response?
- What assay readouts are available for the desired signaling pathways?
- Is the signaling bias obscured by signal over amplification?
- Is the measured bias predictive of the desired *in vivo* outcome?

Conclusion

Due to their extensive involvement in regulating human pathophysiology and their pharmacological tractability, GPCRs are currently the most intensively studied drug targets. The continued evolution of our knowledge about GPCR signaling has brought new opportunities for developing novel drugs targeting more specific diseases with fewer adverse effects. New developments in screening assay technologies further enable the capability to identify and characterize candidate compounds with different biased signaling and/or allosteric modulation. However, the identification and optimization of compounds with biased signaling and/or allosteric modulation poses significant challenges across drug discovery labs in industry and academia. By describing the details, design, and execution of bioassays needed for hit identification, lead optimization, and mechanistic studies, the information in this chapter is relevant to new and experienced GPCR drug hunters.

Acknowledgement

The authors would like to thank Mr. Nathan Cheadle for the careful reading of the manuscripts and the thoughtful comments.

Keywords

G-Protein Coupled Receptor, GPCR, agonist, antagonists, allosteric modulator, biased ligand, drug discovery, lead optimization, pharmacology

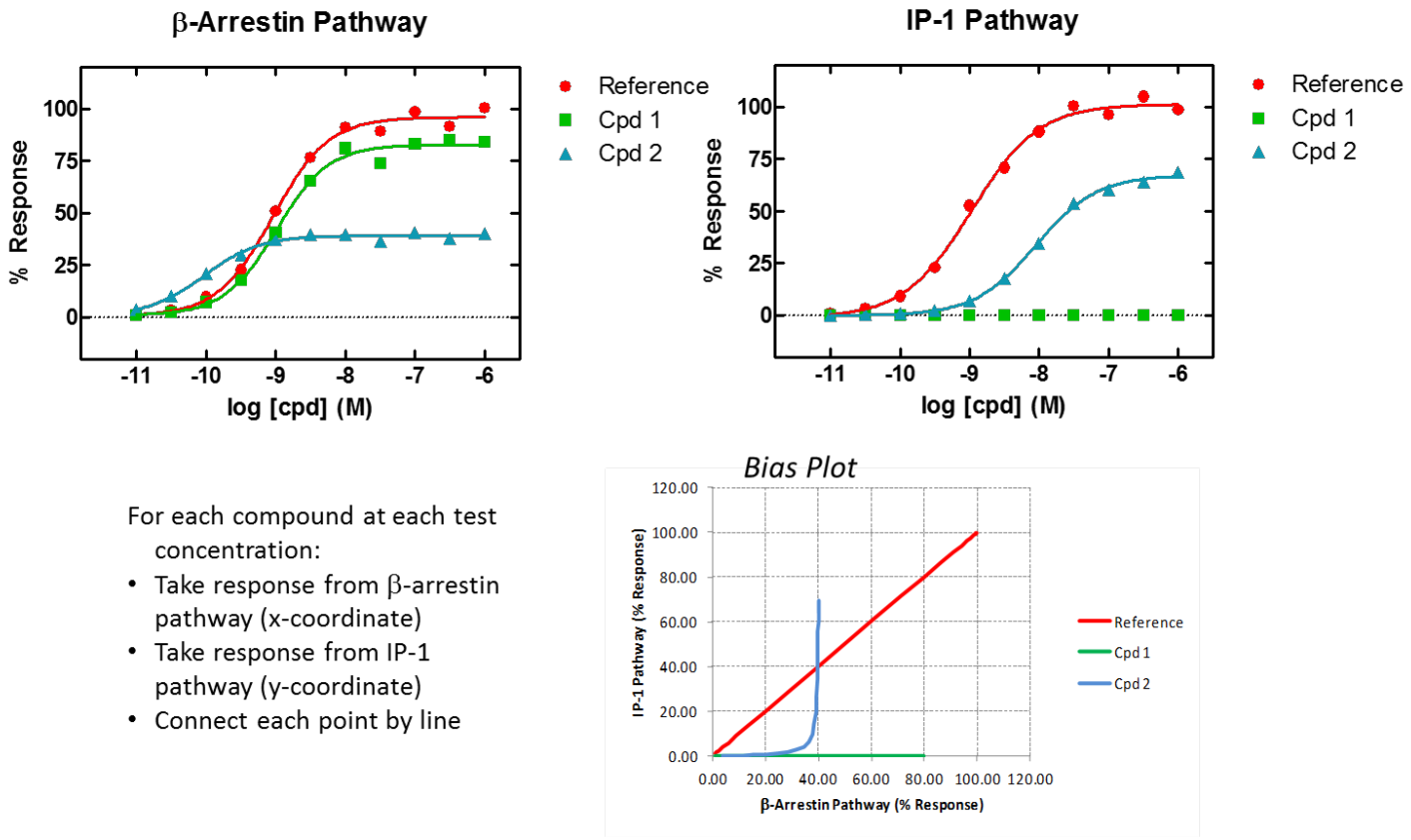


Figure 12. An example to visualize signaling bias across two pathways by Bias Plot.

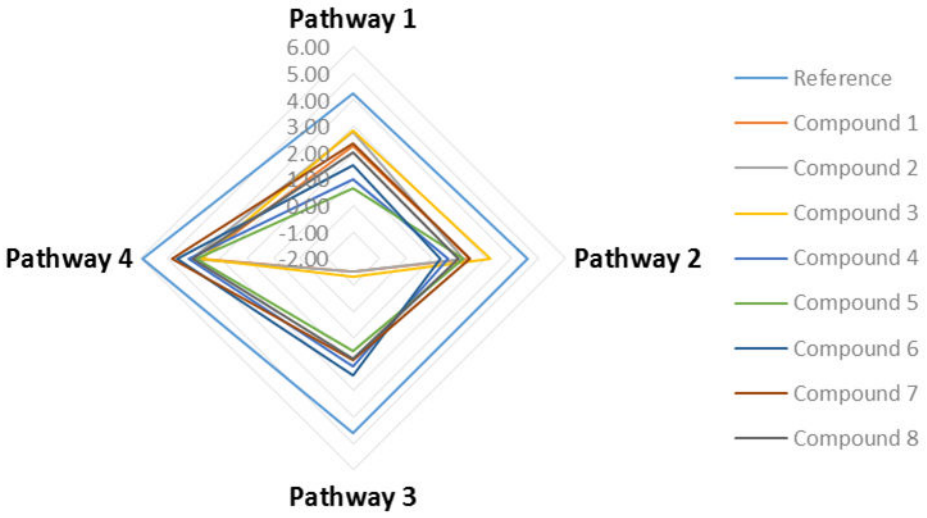


Figure 13. An example of radar plot depicting the signaling profiles of various compounds across four pathways.

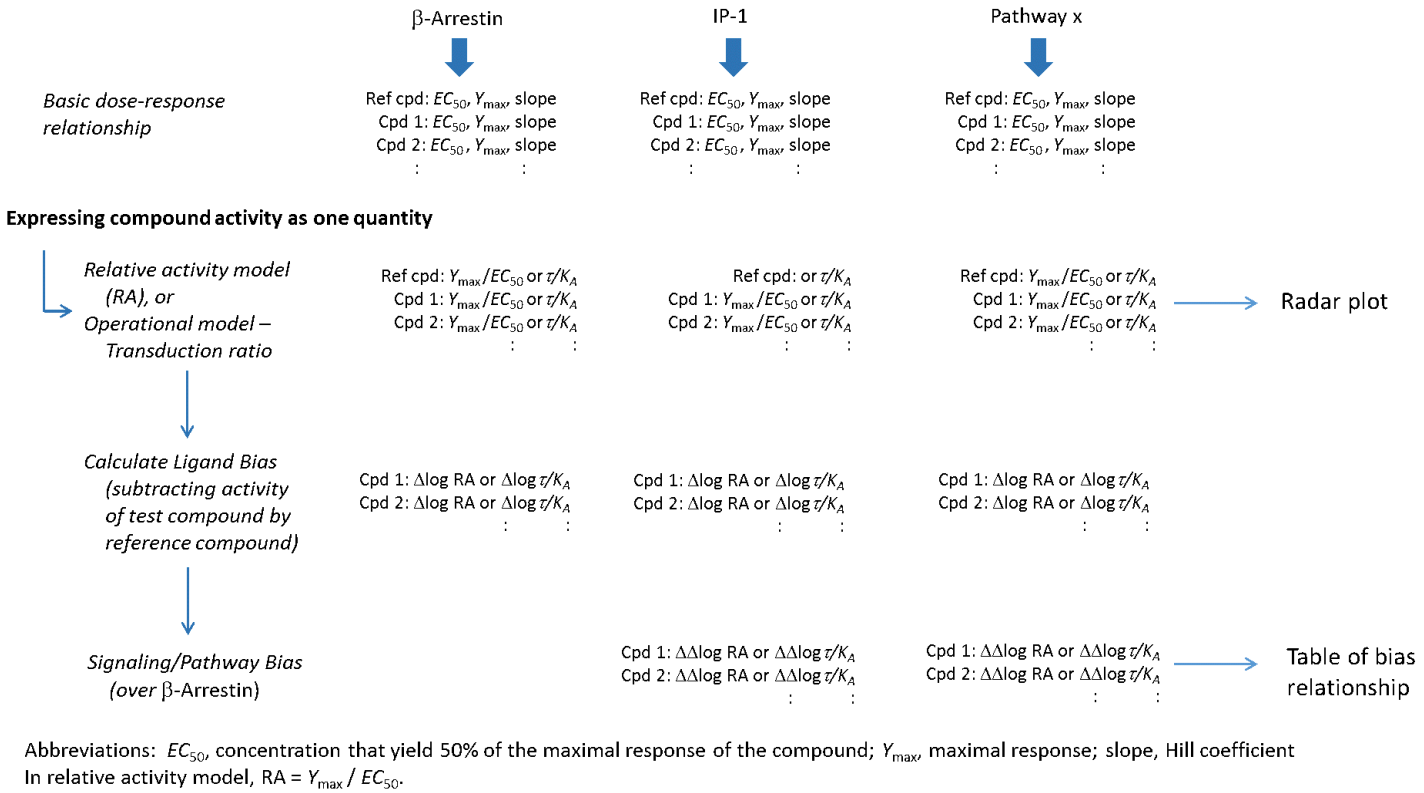


Figure 14. A general scheme to determine the signaling bias across multiple test compounds and signaling pathways.

References

1. Santos R., et al. A comprehensive map of molecular drug targets. Nat Rev Drug Discov. 2016. PubMed PMID: 27910877.
2. Gentry P.R., Sexton P.M., Christopoulos A. Novel Allosteric Modulators of G Protein-coupled Receptors. J Biol Chem. 2015;290(32):19478–88. PubMed PMID: 26100627.
3. Wild C., Cunningham K.A., Zhou J. Allosteric Modulation of G Protein-Coupled Receptors: An Emerging Approach of Drug Discovery. Austin J Pharmacol Ther. 2014;2(1) PubMed PMID: 27148592.
4. Conn P.J., Christopoulos A., Lindsley C.W. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat Rev Drug Discov. 2009;8(1):41–54. PubMed PMID: 19116626.
5. Rocher, J.P., Recent Advances in Drug Discovery of GPCR Allosteric Modulators. MedChem News, 2011. **3**.
6. Kenakin T. Functional selectivity and biased receptor signaling. J Pharmacol Exp Ther. 2011;336(2):296–302. PubMed PMID: 21030484.
7. Viscusi E.R., et al. A randomized, phase 2 study investigating TRV130, a biased ligand of the mu-opioid receptor, for the intravenous treatment of acute pain. Pain. 2016;157(1):264–72. PubMed PMID: 26683109.
8. Gesty-Palmer D., Luttrell L.M. Refining efficacy: exploiting functional selectivity for drug discovery. Adv Pharmacol. 2011;62:79–107. PubMed PMID: 21907907.
9. Kenakin T., Christopoulos A. Signalling bias in new drug discovery: detection, quantification and therapeutic impact. Nat Rev Drug Discov. 2013;12(3):205–16. PubMed PMID: 23411724.
10. Zhang L., et al. Case study: technology initiative led to advanced lead optimization screening processes at Bristol-Myers Squibb, 2004-2009. Drug Discov Today. 2012;17(13-14):733–40. PubMed PMID: 22425710.
11. Cvijic M.E., et al. GPCR profiling: from hits to leads and from genotype to phenotype. Drug Discov Today Technol. 2015;18:30–7. PubMed PMID: 26723890.

12. Zhang, G., Cvijic., M.E., *Screening and Characterization of G-Protein-Coupled Receptor Ligand for Drug Discovery*, in *Handbook of Drug Screening*, R. Seethala, Zhang, L., Editor. 2009, Informa Healthcare USA, Inc.
13. Wang, T., et al., *Measurement of cAMP for Galphas- and Galphai Protein-Coupled Receptors (GPCRs)*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
14. Wang, T., et al., *Measurement of beta-Arrestin Recruitment for GPCR Targets*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
15. Garbison, K.E., B.A. Heinz, and M.E. Lajiness, *IP-3/IP-1 Assays*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
16. Arkin, M.R., et al., *FLIPR Assays for GPCR and Ion Channel Targets*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
17. Cheng Z., et al. Luciferase Reporter Assay System for Deciphering GPCR Pathways. Curr Chem Genomics. 2010;4:84–91. PubMed PMID: 21331312.
18. Picard L.P., et al. Bioluminescence resonance energy transfer-based biosensors allow monitoring of ligand- and transducer-mediated GPCR conformational changes. Commun Biol. 2018;1:106. PubMed PMID: 30271986.
19. Zhang R., Xie X. Tools for GPCR drug discovery. Acta Pharmacol Sin. 2012;33(3):372–84. PubMed PMID: 22266728.
20. Kostenis E., Waelbroeck M., Milligan G. Techniques: promiscuous Galpha proteins in basic research and drug discovery. Trends Pharmacol Sci. 2005;26(11):595–602. PubMed PMID: 16183138.
21. Zhu T., Fang L.Y., Xie X. Development of a universal high-throughput calcium assay for G-protein- coupled receptors with promiscuous G-protein Galpha15/16. Acta Pharmacol Sin. 2008;29(4):507–16. PubMed PMID: 18358098.
22. Luttrell L.M. Minireview: More than just a hammer: ligand "bias" and pharmaceutical discovery. Mol Endocrinol. 2014;28(3):281–94. PubMed PMID: 24433041.
23. Jacobson K.A. New paradigms in GPCR drug discovery. Biochem Pharmacol. 2015;98(4):541–55. PubMed PMID: 26265138.
24. Thanawala V.J., et al. Ligand bias prevents class equality among beta-blockers. Curr Opin Pharmacol. 2014;16:50–7. PubMed PMID: 24681351.
25. Van Der Ryst E. Maraviroc - A CCR5 Antagonist for the Treatment of HIV-1 Infection. Front Immunol. 2015;6:277. PubMed PMID: 26097475.
26. Torres P.U. Cinacalcet HCl: a novel treatment for secondary hyperparathyroidism caused by chronic kidney disease. J Ren Nutr. 2006;16(3):253–8. PubMed PMID: 16825031.
27. Jones, E., S. Michael, and G.S. Sittampalam, *Basics of Assay Equipment and Instrumentation for High Throughput Screening*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
28. Haas, J.V., et al., *Minimum Significant Ratio - A Statistic to Assess Assay Variability*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
29. Campbell, R.M., et al., *Data Standardization for Results Management*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
30. Vempati, U.D. and S.C. Schurer, *Development and Applications of the Bioassay Ontology (BAO) to Describe and Categorize High-Throughput Assays*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
31. Cheng X., et al. Studies on repository compound stability in DMSO under various conditions. J Biomol Screen. 2003;8(3):292–304. PubMed PMID: 12857383.
32. Engeloch C., et al. Stability of screening compounds in wet DMSO. J Biomol Screen. 2008;13(10):999–1006. PubMed PMID: 19029012.
33. Inglese J., et al. Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. Proc Natl Acad Sci U S A. 2006;103(31):11473–8. PubMed PMID: 16864780.

34. Swinney D.C., et al. The Role of Binding Kinetics in GPCR Drug Discovery. Curr Top Med Chem. 2015;15(24):2504–22. PubMed PMID: 26126905.
35. Black J.W., Leff P. Operational models of pharmacological agonism. Proc R Soc Lond B Biol Sci. 1983;220(1219):141–62. PubMed PMID: 6141562.
36. Kenakin, T., *A Pharmacology Primer: Techniques for More Effective and Strategic Drug Discovery*. 4th ed. 2014: Academic Press. 456.
37. Charlton S.J. Agonist efficacy and receptor desensitization: from partial truths to a fuller picture. Br J Pharmacol. 2009;158(1):165–8. PubMed PMID: 19719779.
38. Charlton S.J., Vauquelin G. Elusive equilibrium: the challenge of interpreting receptor pharmacology using calcium assays. Br J Pharmacol. 2010;161(6):1250–65. PubMed PMID: 20977466.
39. Masuho I., et al. Distinct profiles of functional discrimination among G proteins determine the actions of G protein-coupled receptors. Sci Signal. 2015;8(405):ra123. PubMed PMID: 26628681.
40. Salahpour A., et al. BRET biosensors to study GPCR biology, pharmacology, and signal transduction. Front Endocrinol (Lausanne). 2012;3:105. PubMed PMID: 22952466.
41. Sprenger J.U., Nikolaev V.O. Biophysical techniques for detection of cAMP and cGMP in living cells. Int J Mol Sci. 2013;14(4):8025–46. PubMed PMID: 23584022.
42. Kenakin T., Jenkinson S., Watson C. Determining the potency and molecular mechanism of action of insurmountable antagonists. J Pharmacol Exp Ther. 2006;319(2):710–23. PubMed PMID: 16857731.
43. Riddy D.M., et al. Label-Free Kinetics: Exploiting Functional Hemi-Equilibrium to Derive Rate Constants for Muscarinic Receptor Antagonists. Mol Pharmacol. 2015;88(4):779–90. PubMed PMID: 26243731.
44. Kenakin T., et al. A simple method for quantifying functional selectivity and agonist bias. ACS Chem Neurosci. 2012;3(3):193–203. PubMed PMID: 22860188.
45. Griffin M.T., et al. Estimation of agonist activity at G protein-coupled receptors: analysis of M2 muscarinic receptor signaling through Gi/o,Gs, and G15. J Pharmacol Exp Ther. 2007;321(3):1193–207. PubMed PMID: 17392404.
46. Gregory K.J., et al. Identification of orthosteric and allosteric site mutations in M2 muscarinic acetylcholine receptors that contribute to ligand-selective signaling bias. J Biol Chem. 2010;285(10):7459–74. PubMed PMID: 20051519.
47. Smith J.S., Lefkowitz R.J., Rajagopal S. Biased signalling: from simple switches to allosteric microprocessors. Nat Rev Drug Discov. 2018. PubMed PMID: 29302067.

In Vivo Assay Guidelines

Viswanath Devanarayan, PhD and Philip W. Iversen, PhD

In Vivo Assay Guidelines

Joseph Haas, MS,[1,*] Jason Manro, MS,[1] Harlan Shannon, PhD,[2] Wes Anderson, MS,[1,**] Joe Brozinick, PhD,[1] Arunava Chakravartty, PhD,[3] Mark Chambers, PhD,[1] Jian Du, PhD,[1] Brian Eastwood, PhD,[1] Joe Heuer, PhD,[1] Stephen Iturria, PhD,[4] Philip Iversen, PhD,[1,***] Dwayne Johnson, BS,[1] Kirk Johnson, PhD,[1] Michael O'Neill, PhD,[1] Hui-Rong Qian, PhD,[1] Dana Sindelar, PhD,[1] and Kjell Svensson, PhD[1]

Created: May 1, 2012; Updated: October 1, 2012.

Abstract

This document is intended to provide guidance for the design, development and statistical validation of *in vivo* assays residing in flow schemes of discovery projects. It provides statistical methodology for pre-study, cross-study (lab-to-lab transfers and protocol changes), and in-study (quality control monitoring) validation. Application of the enclosed methods will increase confidence in data from *in vivo* assays in the critical path and enable better decisions about SAR directions and compound prioritization. Screens using both single dose and multiple dose *in vivo* assays are discussed, along with acceptance criteria at different stages of validation.

Flow Chart

This document is intended to provide guidance for the development and statistical validation of *in vivo* assays residing in flow schemes of discovery projects. It provides statistical methodology for pre-study, cross-study (lab-to-lab transfers and protocol changes), and in-study (quality control monitoring) validation. Application of the enclosed methods will increase confidence in data from *in vivo* assays in the critical path and enable better decisions about SAR directions and compound prioritization.

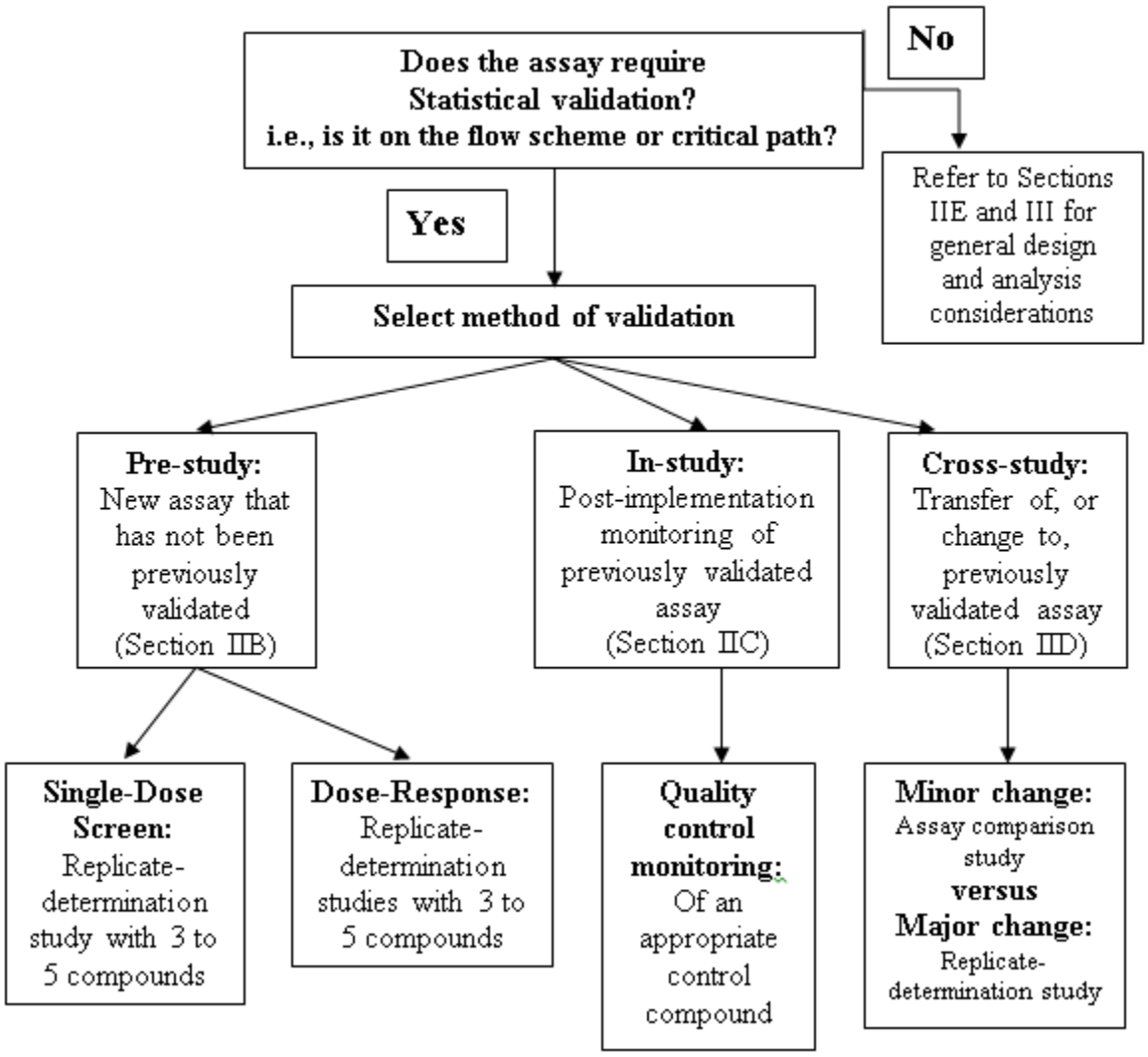
Author Affiliations: 1 Eli Lilly & Company, Indianapolis, IN. 2 Indiana University Purdue University at Indianapolis, Indianapolis, IN. 3 Abbott Laboratories, Chicago, IL. 4 Monsanto, Ankeny, IA.

Corresponding author.

*Principal Author

**Co-Author

***Editor



The document is organized into three sections, and includes several examples. The Introduction (Section 1) provides general definitions of biological assays and provides general concepts in assay development and validation. Assay Validation (Section 2) contains the procedures and statistical details for *in vivo* assay validation, while Background Material (Section 3) covers general statistical concepts related to the design and analysis of *in vivo* experiments.

1. Introduction

This document is written to provide guidance to investigators who are developing and statistically validating *in vivo* assays for the evaluation of structure-activity relationships and/or compound collections to identify chemical probes that modulate the activity of biological targets. Specifically, this manual provides guidelines for:

a. Identifying potential assay formats of *in vivo* models compatible with **Single Dose Screens** (SDSs) and **Dose-Response Curves** (DRCs) for evaluating structure-activity relationships (SAR).
b. Statistical validation of the assay performance parameters (pre-study, in-study, and cross-study validation).
c. Optimizing assay protocols with respect to sensitivity, dynamic range, and stability.

1.1. General definition of biological assays

A biological assay is defined by a set of methods that produce a detectable signal allowing a biological process to be quantified. In general, the quality of an assay is defined by the *robustness and reproducibility* of this signal in the absence of any test compounds or in the presence of inactive compounds. This robustness will depend on the type of signal measured (biochemical, physiological, behavioral, etc.), and the analytical and automation instrumentation employed. The quality of the SDS is then defined by the behavior of this assay system when screened against a collection of compounds. These two general concepts, assay quality and screen quality, are discussed with specific examples in this manual.

1.2. General Concepts in Method (Assay) Development and Validation of an *In vivo* Model

The overall objective of any method validation procedure is to demonstrate that the method is *acceptable for its intended purpose*. Usually, the purpose is to determine the biological and or pharmacological activity of new chemical entities (NCE). The acceptability of a measurement procedure or bioassay method begins with its design and construction, which can significantly affect its performance and robustness.

The validation process originates during identification and/or design of a model and method development and continues throughout the assay life cycle (Figure 1). During method development, assay conditions and procedures are selected that minimize the impact of potential sources of invalidity (e.g. so-called false positives or false negatives) on the measurement of analyte or the biological end point (e.g. biochemical, physiological or behavioral changes). There are three fundamental general areas in method development and validation: (a) Pre-study (Identification and Design phase) validation (b) In-study (Development and Production phase) validation, and (c) Cross-validation or method transfer validation. These stages encompass the systematic scientific steps in an assay development and validation cycle.

1.2.1. Pre-study validation:

This validation occurs prior to implementing the assay. At this stage the choice of an assay format is made. Close attention must be paid to factors such as the selection of methods with appropriate specificity and stability. It is important that the assay be designed and the protocol written to facilitate the statistical analysis (i.e. appropriate design and analysis method, and adequate sample size), and that proper randomization techniques be used. This requires the generation and statistical analysis of confirmatory data from planned experiments to document that analytical results satisfy pre-defined acceptance criteria. Within-run statistical measures of assay performance such as the Minimum Significant Difference (MSD) for SDSs and Minimum Significant Ratio (MSR) for DRCs are calculated. If available, the assay sensitivity and pharmacology is evaluated using control compounds.

1.2.2. In-study validation:

These procedures are needed to verify that a method remains acceptable *during its routine use*. In order to compare data for compounds tested at different times, the pre-study statistical measures of assay performance (MSD or MSR) are updated to include between-run variability. Each run of the assay should contain appropriate {maximum and minimum} control groups or treatments to serve as quality controls of each run and to check overall performance. (Maximum and/or minimum quality controls differ conceptually from an “active” positive or negative controls; see Control Compounds or Treatment Groups.) This will allow the investigator to check for procedural errors and to evaluate stability of the method over time. Control Charts illustrates procedures which may be used to evaluate assay performance over time (i.e., control chart monitoring).

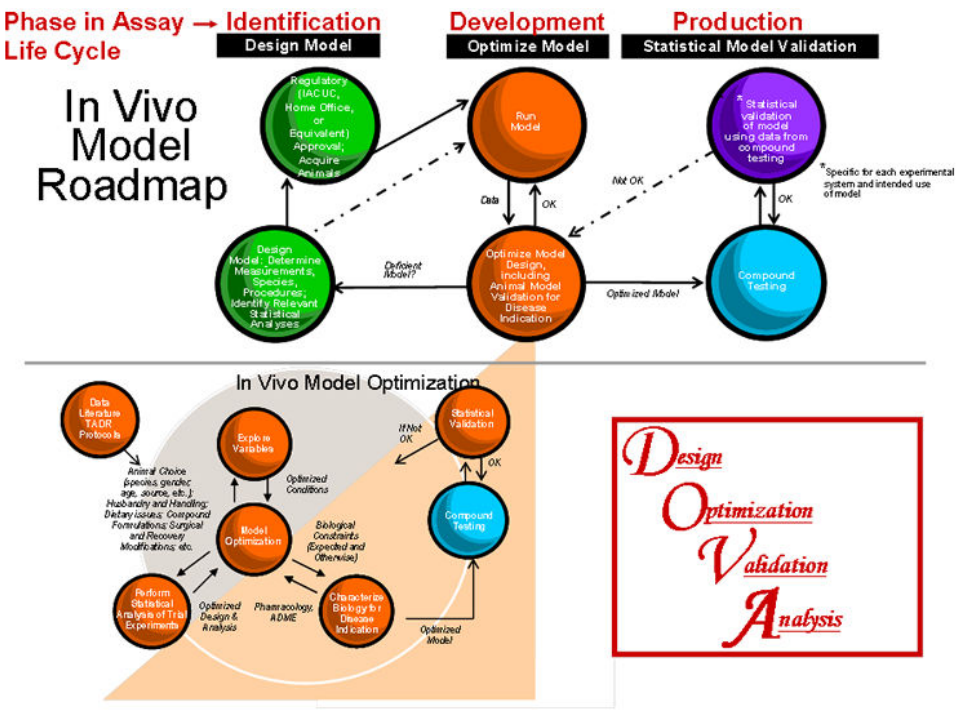


Figure 1: The *In vivo* Model Assay Development Cycle.

1.2.3. Cross validation:

This portion includes the *assay hand-off* from the individual investigator's team to another laboratory or a screening center. More broadly, this procedure is used at any stage to verify that an acceptable level of agreement exists in analytical results before and after procedural changes in a method as well as between results from two or more methods or laboratories. Typically, each laboratory assays a subset of compounds and the agreement in results is compared to predefined criteria that specify the allowable performance of the assay.

1.2.4. Resources:

The validation guidelines described here should be applicable to most *in vivo* models encountered in drug discovery research. However, situations could arise in which their verbatim application would be impractical given resource constraints, intended use of the assay, or other reasons. In these situations, and in general, the following principles should apply:

- Some form of statistical validation should always be performed and is better than no validation.
- The amount of resources, including time, spent on validation should be kept to a reasonably small fraction of the total resources to be used for testing compounds. What is "reasonable" will be have to be determined by the key personnel involved with each project.

These guidelines are intended to be used as "guidelines" and not exact "requirements." The specified assay performance measures serve to quantify how well the assay is performing and should be used to guide proper interpretation of the data. Determining whether an assay is "fit for its intended purpose" should be based on a combination of these assay performance measures and sound scientific judgment of the team.

2. Assay Validation Procedures

2.1. Overview

The statistical validation requirements for an assay vary, depending upon the prior history of the assay. The four main components of the statistical validation are:

1. Adequate study design and data analysis method.
2. Proper randomization of animals.
3. Appropriate statistical power and sample size.
4. Adequate reproducibility across assay runs.

Assays should be designed so that all biologically meaningful effects are statistically significant. In an exploratory study, this "meaningful effect" might correspond to any effect that is pharmacologically relevant. For a project/program team, it might correspond to an effect that meets the critical success factors (CSFs) defined in the compound development flow scheme. Power and sample size analyses are especially relevant for experiments that are designed to address key endpoints in a flow scheme. It is not acceptable to set a CSF equal to the effect size that is statistically significant since that effect size may or may not be biologically relevant. A CSF should be established based upon its biological relevance to the discovery effort. The assay is designed, optimized and validated so that biologically meaningful effects (e.g., CSFs) are statistically significant. Quantifying the reproducibility of an *in vivo* flow scheme assay will enable a team to discern whether compounds tested in different runs of the assay are exhibiting differential activity. This will result in better decisions about SAR directions and compound prioritization.

If the assay is new, or has never been previously validated relative to the targeted purpose, or mechanism of action, of the assay, then full validation should be performed. If the assay has been previously validated in a different laboratory, and is being transferred to a new laboratory, then a Replicate-Determination study (Pre-Study Validation) and a formal comparison between the new assay and the existing (old) assay (i.e., cross-study validation; Cross Validation) should be performed. If an assay is being transferred, it is considered validated if it has previously been assessed by all the methods in Section 2.2, and is being transferred to a new laboratory without undergoing any substantive changes to the protocol. If the intent is to store the data under the same Assay Method Version (or AMV) in an electronic database as the previous laboratory's AMV, then an assay comparison study (Cross Validation) should be done as part of the Replicate-Determination study. Otherwise only the intra-laboratory part of the Replicate-Determination study (Pre-Study Validation) is recommended.

If the assay is updated from a previous version run in the same facility then the requirements vary, depending upon the *extent of the change*. Major changes require a validation study equivalent to the validation of a new assay. Minor changes require bridging studies that demonstrate the equivalence of the assay before and after the change. See Section 2.4 - Cross Validation for examples of major and minor changes.

An assay methodology which has been previously validated for a different target or mechanism of action should be validated in full when used for a new or different target or mechanism as the variability, in particular, and reproducibility may be quite different for different mechanisms even though the methods may be very similar or identical. This concept is analogous to separately validating receptor binding assays for different receptors.

2.2. Pre-Study Validation: Replicate-Determination Study for Single-Dose Screens and Dose-Response Curves

2.2.1. Overview and Rationale

It is important to verify that assay results from multiple determinations or assay runs have acceptable reproducibility with no material systematic trends in the key endpoints. In this section, we define how to

quantify assay variability and determine assay equivalence. We also explain the rationale for the statistical methods employed in calculating reproducibility of activity and potency. *We strongly recommend consultation with a statistician before designing experiments to estimate variability, and the sources thereof, described below*. In particular, you should discuss with a statistician alternatives for assays with significant time, resource, or expenditure constraints as well as assays which will be used to test a minimal number of compounds to properly balance validation requirements with these constraints.

Replicate-Determination studies with two runs are used to formally evaluate the *within-run* assay variability (i.e., pre-study validation), or to formally compare a new assay to an existing (old) assay (i.e., cross-study validation). Replicate-Determination studies also allow a preliminary assessment of the *overall* or *between-run* assay variability, but two runs are not enough to adequately assess overall variability. In-study methods (In-Study Validation) are used to formally evaluate the overall variability in the assay in routine use. Note that the Replicate-Determination study is a diagnostic and decision tool used to establish that the assay is ready to go into production by showing that the endpoints of the assay are reproducible over a range of efficacies or potencies. It is not intended as a substitute for in-study monitoring or to provide an estimate of the *overall* MSD or MSR.

It may seem counter-intuitive to call the differences between two independent assay runs "within-run." However, the terminology results from the way those terms are defined. Experimental variation is categorized into two distinct components: *between-run* and *within-run* sources.

Consider the following examples:

- Between-run variation: If there is variation in the concentrations of the components in the vehicle between two runs then the assay results could be affected. However, assuming that the same vehicle is used with all compounds within the run, each compound will be equally affected and so the difference will only show up when comparing the results of two runs: one run will appear higher on average than the other run. This variation is called *between-run* variation.
- Within-run variation: If the concentration of one compound in the vehicle varies from the intended concentration (or dose) then all animals receiving that compound will be affected. However, animals receiving other compounds will be unaffected. This type of variation is called *within-run* as the source of variation affects different compounds in the same run differently. Therefore, it is necessary to compare results for one compound on each of two occasions.
- Some sources of variability affect *both within- and between-run variation*. For example, environmental conditions in an animal room have the potential to contribute to both types of variability. Suppose within a run of a particular *in vivo* assay, the temperature in the animal room exhibits spatial variation and animal response is sensitive to temperature. Animals will then respond differently depending upon their location in the room, and these differences are *within-run* as not all animals are equally affected. In comparison, suppose that during a particular run of the assay the room temperature on average is higher or lower than in previous runs. In this instance, the animals will respond differently on average during this run relative to other runs, and since all animals are affected this is *between-run* variation. Thus, a variable such as animal room temperature can be a source of both within- and between-run variation.

The *total variation* is the sum of both sources of variation. When comparing two compounds across runs, one must take into account both the within-run and between-run sources of variation. But when comparing two compounds in the same run, one must only take into account the within-run sources, since, by definition, the between-run sources affect both compounds equally.

In a Replicate-Determination study, the between-run sources of variation cause one determination to be on average higher than another determination. However, it would be very unlikely that the difference between the two determinations would be exactly the same for every compound in the study. These individual compound

"differences from the average difference" are caused by the within-run sources of variation. The higher the within-run variability the greater the individual compound variation within assay runs.

The analysis approach used in the Replicate-Determination study is to estimate and factor out between-run variability, and then estimate the magnitude of within-run variability.

Note: The between- and within-run sources of variability assessed during assay validation calculations apply to the treatment group means, not to the animal level data. Animal-to-animal variability is obviously present in *in vivo* experiments and this variability affects the reproducibility of the resulting treatment group means, but as shown subsequently, the animal-to-animal variability is not directly used in the calculation of the MSD and MSR. It is used to assess whether a data transformation is needed, and to assess sample size adequacy. See Section 2.5 for additional information about sources of variability in an *in vivo* assay.

2.2.2. Procedure (Steps)

All assays should have a reproducibility comparison (Steps 1 – 3). Single-dose screens should be validated separately from dose-response curve determinations, but it may be possible to validate both methods concurrently (consult with a statistician for options; also see Section 2.2.7 for an example). If the assay is to replace an existing assay using the same AMV code then an assay comparison study should also be done (Step 4).

1. Select a minimum of 3 to 5 compounds that have activities covering the effect range of interest and, if applicable, potencies that cover the dose-range to be tested. The compounds should be well spaced over these ranges.
2. All of the compounds should be tested in each of two runs of the assay.
3. Compare the two runs (as per Section 2.2.5 – 2.2.6.2.)
4. If the assay is to replace an existing assay:
    a. All compounds should be tested in a single run of the former assay as well as in two runs of the new assay. If a single run of the former assay already exists that meets the requirements in Step 1, it can be used for validation as long as the run is reasonably contemporaneous in time with the run of the new assay.
    b. Compare the results of the two assays, or labs, by analyzing the first run of the new assay with the single run of the former assay.

2.2.3. Summary of Acceptance Criteria (discussed in detail below)

1. For new assays, in Step 3 conduct reproducibility and equivalence tests for activity (single-dose screens and dose-response assays) and/or potency (dose-response assays) comparing the two runs. The assay should have an MSD < 20% and both Limits of Agreement on Differences (LsAd) between -20% and +20% for % activity. For potency results, recommendations are an MSR < 3 and both Limits of Agreement (LsA) between 0.33 and 3.0.
2. For assay transfer purposes, in Step 3 conduct reproducibility and equivalence tests for activity (single-dose screen and dose-response assays) and/or potency (dose-response assays) comparing the two runs in the new lab. The assay should have an MSD < 20% and both Limits of Agreement between -20% and +20% for % activity, and an MSR < 3 and both Limits of Agreement between 0.33 and 3.0 for potency.
3. For assay transfer purposes, in Step 4b conduct reproducibility and equivalence tests for activity and/or potency comparing the first run of the new lab to the single run of the old lab. The assays should have Limits of Agreement between -20% and +20% for % activity and Limits of Agreement between 0.33 and 3.0 for potency to be declared equivalent.

2.2.4. Notes

1. The Replicate-Determination study as laid out in Sections 2.2.2 and 2.2.3, assumes that the entire study is accomplished in just two runs of the assay (i.e. selected compounds are tested in each of two runs). While this design is preferred, it may not be possible for low-throughput assays in which only a small number of compounds can be tested per run. Alternatives are available in which each half of the replicate determination spans multiple runs (i.e., selected compounds are tested in each of a set of assay runs). Teams should discuss possible design alternatives and the appropriate analysis with their statistician. See also Section 2.2.7 for one possible alternative design and analysis.
2. *The acceptance criteria summarized above should be considered guidelines*, and may in fact be too stringent for some assays. Failure to meet the acceptance guidelines does not necessarily mean that the assay is unusable. Teams should consult with their statistician to understand the ramifications of missing the recommended criteria, and how it affects setting the CSF for the project and making decisions about compounds. For example, if the CSF is >80% and the MSD is 30%, then the assay will fail too many efficacious compounds, since even a 90%-active compound will fall below the CSF some of the time when the MSD is 30%. A more appropriate CSF in this situation might be 70 or even 60%. Furthermore, a 30% MSD indicates that a compound would need an activity test result of at least 80% to be considered to have differential activity from a compound with an activity test result of 50%. If the team desires increased power to discriminate among test compounds, the assay may need to be re-optimized to identify and reduce sources of variability (see Section 2.5 for additional details).
3. If a project is very new, there may not be 3 to 5 unique compounds with activity *in vivo* (where activity means some measurable activity above the minimum threshold of the assay). In that case it is acceptable to test compounds more than once, and/or at different doses, to get an acceptable sample size. For example, if there are only 2 active compounds then test each compound twice at the same two or three doses in each determination. However, when doing so, (a) it is important to biologically evaluate them as though they were different compounds, including independent sample preparation (i.e., weighing and solubilization), and (b) label the compounds and/or doses as "a", "b" etc. so that it is clear in the Replicate-Determination analysis which results are being compared across runs.
4. Dose-response assays need to be compared for both potency (ED_{50}) and efficacy (% maximum response). It is acceptable to use the same compounds for both single-dose and dose-response assays.

An assay may pass the reproducibility assessment (Steps 1-3 in the procedure [Section 2.2.2]), but may fail the assay comparison study (Step 4 in the procedure [Section 2.2.2]). The assay comparison study may fail either because of an absolute Mean Difference (MD) > 5% (or Mean Ratio (MR) different from 1), or a high MSD (or MSR) in the assay comparison study. If it's the former then there is an activity or potency shift between the assays. You should assess the values in the assays to ascertain their validity (e.g., which assay's results compare best to those reported in the literature?). If it fails because the assay comparison study MSD (or MSR) is too large (but the new assay passes the reproducibility study) then the old assay lacks reproducibility. In either case, if the problem is with the old assay, then the team should consider re-running key compounds in the new assay to provide comparable results to compounds subsequently run in the new assay.

2.2.5. Activity/Efficacy; Single Dose Screens

2.2.5.1. Analysis: Activity\Efficacy; Single Dose Screens

The points below describe and define the terms used in the acceptance criterion summarized in Section 2.2.3 and discussed in the Diagnostic Tests (Section 2.2.5.2). Individual animal data should be normalized to the positive (maximally effective) and negative (minimally effective) control averages to yield % activity for each animal. The computations that follow should then be made on the % activity averages for each compound, not the individual animal % activities.

1. Compute the difference in activity (first minus second) between the first and second determination for each compound. Let $\bar{d}$, and s_d be the sample mean and standard deviation of the difference in activity.
2. Compute the **Mean-Difference**: $MD = |\bar{d}|$ This is the average difference in activity between the two determinations.
3. Compute the **Difference Limits**: $DLs = \bar{d} \pm 2s_d/\sqrt{n}$, where n is the number of compounds. This is a 95% confidence interval for the Mean-Difference.
4. Compute the **Minimum Significant Difference**: $MSD = 2s_d$. This is the smallest activity difference between two compounds that is statistically significant.
5. Compute the **Limits of Agreement**: $LsAd = \bar{d} \pm 2s_d$. Most of the compound activity differences should fall within these limits (approximately 95%).
6. For each compound compute the **Difference** (first minus second) of the two activities, and the **Mean** activity (average of first and second).

Items 2-6 can be combined into one plot: the Difference-Mean plot (see example below).

2.2.5.2. Example

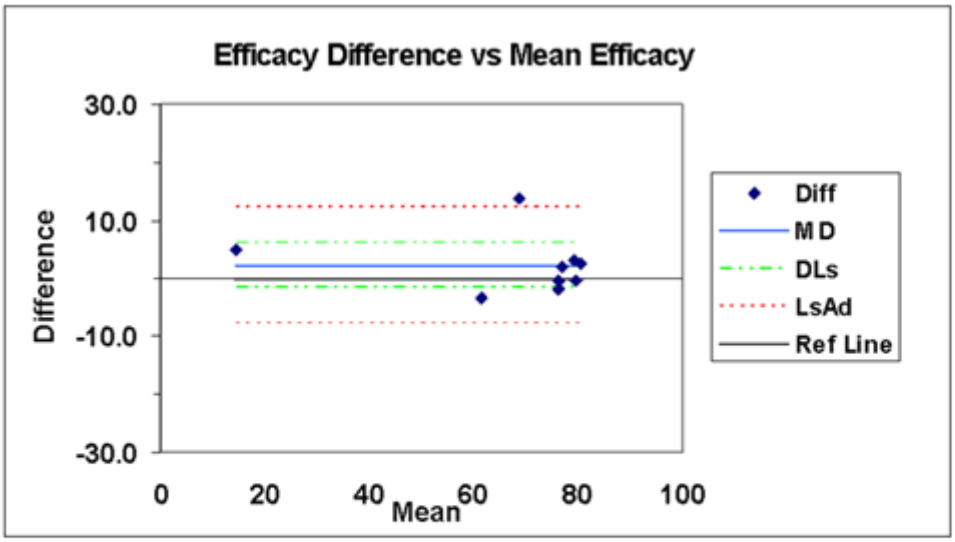
What follows is an example of an *in vivo* assay which measures the inhibition of EnzymeX Desaturase Index in rats, and which was transferred to Lab B from Lab A. Nine compounds were tested in each of two runs in Lab B with n=5 rats per compound. Table 1 displays the % inhibition averages of 5 rats per compound, along with the differences in compound averages. The MD, MSD, and Limits of Agreement are also shown which are computed from the average and standard deviation of the differences. This particular assay easily meets the acceptance criteria summarized in Section 2.2.3.

The compound averages can also be entered directly into the Replicate-Experiment template developed for in vitro assays and available on the NIH website. This template will perform the MD, MSD, and LsAd calculations and display them in the Difference-Mean plot.

Figure 2 shows the desired result of pure chance variation in the difference in activities between runs. The blue solid line shows the Mean Difference, i.e. the *average* relationship between the first and second run. The green long-dashed lines show the 95% confidence limits (or Difference Limits) of the Mean Difference. These limits should contain the value 0, as they do in this case. The red short-dashed lines indicate the Limits of Agreement between runs. They indicate the *individual compound* variation between the first and second run. You should see all, or almost all, the points fall within the red dashed lines. The lower line should be above -20, while the upper line should be below +20, which indicates a 20% difference between runs in either direction. The MD should be less than 5%, as it is in this example.

As this assay originated in Lab A, the lab comparison described in Step 4 of Section 2.2.2 was also completed. Eight of the nine compounds tested in the Lab B Replicate-Determination were also tested in Lab A. The eight compound averages of n=5 rats were entered into the template along with the corresponding eight compound averages from the first run in Lab B. Figure 3 shows the results, which meet the acceptance criteria summarized in Part 3 of Section 2.2.3.

Note: Although this assay achieved a successful transfer from Lab A to Lab B, demonstrating reproducibility within Lab B, as well as equivalence between labs, the validation has a shortcoming with respect to the selected compounds. *The selected compounds are not well spaced over the activity range of interest*, with most compound activities clustered together and falling well above the CSF of 50%. Since the reproducibility of a compound result from a given assay could depend on the activity level of the compound, it is very important to make an effort to cover the range of potential activity, and in particular to include compounds with activity bracketing the CSF (recognizing that this may be difficult depending on the maturity of a particular project).



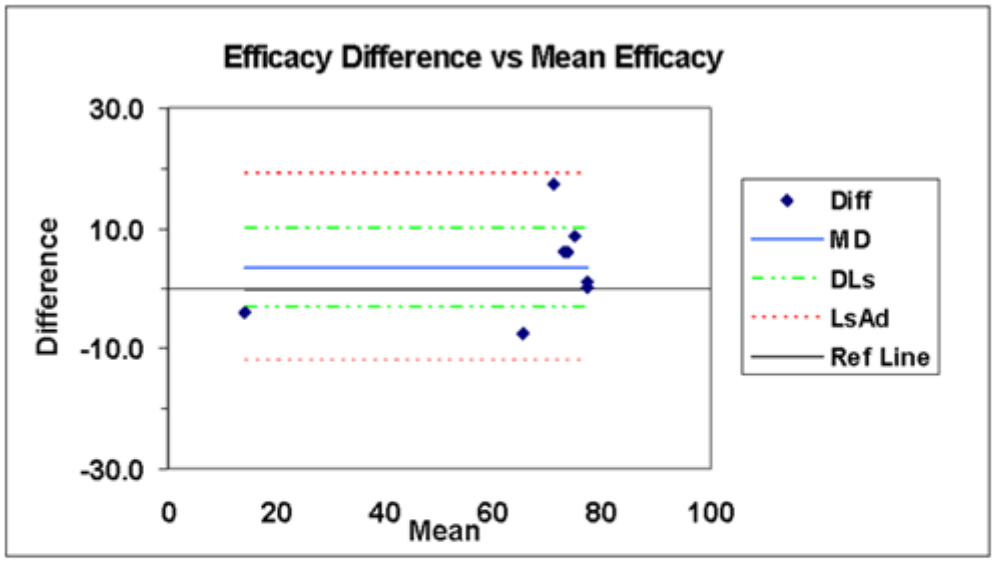
Equivalence Test Results

N	9
MD	2.20
DLs	-1.67
	6.07

Reproducibility Test Results

MSD (Within-Run)	10.06
LsAd	-7.87
	12.26

Figure 2: Result of reproducibility and equivalence tests for activity comparing the two runs in the new lab, Lab B. This example shows the desired result of pure chance variation in the difference in activities between runs.



Equivalence Test Results

N	8
MD	3.49
DLs	-3.03
	10.01

Reproducibility Test Results

MSD (Within-Run)	15.60
LsAd	-12.10
	19.09

Figure 3: Result of reproducibility and equivalence tests for activity comparing the first run of the new lab (Lab B) to the single run of the old lab (Lab A). This example shows results that meet the acceptance criteria.

Table 1: EnzymeX Desaturase Index Assay Replicate-Determination in Lab B.

Compound	Test1 %Inh	Test2 %Inh	Difference
A	76.19	76.68	-0.49
B	17.01	12.21	4.79
C	80.78	77.76	3.02
D	75.49	61.77	13.72
E	77.87	76.04	1.83
F	59.78	63.13	-3.35
G	79.47	79.84	-0.36
H	75.28	77.26	-1.98
I	81.92	79.30	2.61
MD = AVG of diffs		**2.20**	
MSD = 2 x STD of diffs		**10.06**	
LsAd = MD ± MSD (2.20 ± 10.06)		**-7.87**	**12.26**

2.2.5.3. Diagnostic Tests and Acceptance Criterion: Activity\Efficacy; Single Dose Screens

1. If the MSD ≥ 20% then there is poor agreement for **individual** compounds between the two determinations. This problem occurs when the *within-run* variability of the assay is too high. An assay meets the MSD acceptance criterion if the (within-run) MSD < 20%.
2. If the Difference Limits do not contain the value 0, then there is a statistically significant **average** difference between the two determinations. Within a lab (Step 3) this is due to high *between-run* assay variability. Between labs (Step 4), this could be due to a systematic difference between labs, or high between-run variability in one or both labs. Note that it is possible with a very "tight" assay (i.e., one with a very low MSD) or with a large set of compounds to have a statistically significant result for this test that is not very material, i.e., the actual MD is small enough to be ignorable. If the result is statistically significant then examine the MD. If it is between -5% and +5% then the average difference between runs is deemed immaterial.
3. The MD and the MSD are combined into a single interval referred to as the **Limits of Agreement**. An assay that either has a high MSD and/or an MD different from 0 will tend to have poor agreement of results between the two determinations. An assay meets the Limits of Agreement acceptance criterion if both the upper and lower limits of agreement are between -20 and +20.

2.2.6 Potency; Dose-response curves

2.2.6.1. Analysis: Potency; Dose-response curves

The points below describe and define the terms used in the acceptance criterion summarized in Section 2.2.3 and discussed in the Diagnostic Tests (Section 2.2.6.2). See Section 3.5 for guidelines for fitting *in vivo* dose-response curves.

1. Compute the difference in log-potency (first minus second) between the first and second determination for each compound. Let $\bar{d}$, and s_d be the sample mean and standard deviation of the difference in log-potency. Since ratios of ED_{50} values (relative potencies) are more meaningful than differences in potency (1 and 3, 10 and 30, 100 and 300 have the same ratio but not the same difference), we take logs in order to analyze ratios as differences.
2. Compute the **Mean-Ratio**: $MR = 10^{\bar{d}}$. This is the geometric average fold-difference in potency between two determinations.
3. Compute the **Ratio Limits**: $RLs = 10^{\bar{d} \pm 2s_d/\sqrt{n}}$, where n is the number of compounds. This is the 95% confidence interval for the Mean-Ratio.
4. Compute the **Minimum Significant Ratio**: $MSR = 10^{2s_d}$. This is the smallest potency ratio between two compounds that is statistically significant.
5. Compute the **Limits of Agreement**: $LsA = 10^{\bar{d} \pm 2s_d}$. Most of the compound potency ratios (approximately 95%) should fall within these limits.
6. For each compound compute the **Ratio** (first divided by second) of the two potencies, and the **Geometric Mean** potency: $GM = \sqrt{\text{first} \times \text{second}}$.

Items 2-6 can be combined into one plot: the Ratio-GM plot. The plot is very similar to the Difference-Mean plot described previously in Section 2.2.5 except that both axes are on the log scale instead of the linear scale.

2.2.6.2. Diagnostic Tests and Acceptance Criterion: Potency; Dose-response curves

1. If the MSR ≥ 3 then there is poor agreement for **individual** compounds between the two runs. This problem occurs when the *within-run* variability of the assay is too high. An assay meets the MSR acceptance criterion if the (within-run) **MSR < 3.**
2. If Ratio Limits do not contain the value 1, then there is a statistically significant **average** difference between the two determinations. Within a lab (Step 3) this is due to high *between-run* assay variability.

Between labs (Step 4), this could be due to a systematic difference between labs, or high between-run variability in one or both labs. Note that it is possible with a very "tight" assay (i.e., one with a very low MSR), or with a large set of compounds, to have a statistically significant result for this test that is not very material, i.e., the actual MR is small enough to be ignorable. If the result is statistically significant then examine the MR. If it is between 0.67 and 1.5 then the average difference between runs is less than 50% and is deemed immaterial. Note that there is no direct requirement for the MR, but values that are outside 0.67 and 1.5 are unlikely to pass the Limits of Agreement criterion in step 3 below.

3. The MR and the MSR are combined into a single interval referred to as the **Limits of Agreement**. An assay that either has a high MSR and/or an MR different from 1 will tend to have poor agreement of results between the two determinations. An assay meets the Limits of Agreement acceptance criterion if both the upper and lower limits of agreement are between 0.33 and 3.0.

2.2.7. Other Approaches for Assessing Reproducibility

2.2.7.1 Retrospective Assessment of Reproducibility

The preferred method for validating an in-vivo assay is by the Replicate-Determination Study described in Sections 2.2.1 – 2.2.6.2. This prospective approach provides an up-front assessment of the capability of an assay to reproducibly identify active compounds prior to placing the assay into production.

It is also possible to validate an assay retrospectively, that is, to collect validation data while screening test compounds. For example, a Single Dose Screen that has been adequately designed and powered, and that will use proper randomization and analysis techniques could be placed immediately into production. One could then include a "quality control" (see Sections 2.3.1 and 3.1 for a description) in each run of test compounds in order to assess the reproducibility component of assay validation. After six or more runs of the assay, the quality control data could be used to calculate an MSD for the assay.

One obvious risk in using the retrospective approach is that six or more runs of test compound data will be generated before knowing whether or not the reproducibility of the assay is adequate for a team's needs. This can be an efficient approach provided the resulting MSD is acceptably small; if it is not, the assay may need to be re-optimized to identify and reduce sources of variability (see Section 2.5 for additional details).

We strongly recommend consultation with a statistician before embarking on this approach to assessing the reproducibility component of assay validation. Teams should also consult a statistician regarding the calculation and interpretation of the MSD once the data are collected. Note also that this approach produces an "overall" MSD (see Section 2.1 for a description), since in addition to within-run variation, the resulting MSD also encompasses between-run variation from more than six runs.

2.2.7.1.1. Example

What follows is an example of an *in vivo* assay run in Lab B, which measures the inhibition of EnzymeX Desaturase Index in rats, and which was transferred to Lab B from Lab A and discussed in Section 2.2.5. The assay was retrospectively validated in Lab A by including a quality control in each of seven runs. These data are shown in Figure 4, which displays % inhibition data for individual animals as well as the averages for each run (connected by the blue line in the plot). The overall MSD of 16% is calculated from the standard deviation of the seven % inhibition averages and meets the validation criteria of MSD < 20%. The variability chart also indicates a decrease in animal-to-animal variation over time, as well as a slight increase in the activity of the quality control over time.

The % inhibition averages can also be entered into the control chart template (see Section 2.3.2 for additional details), which will calculate the overall MSD and allow monitoring of the averages over time. Output from the template is displayed in Figure 5.

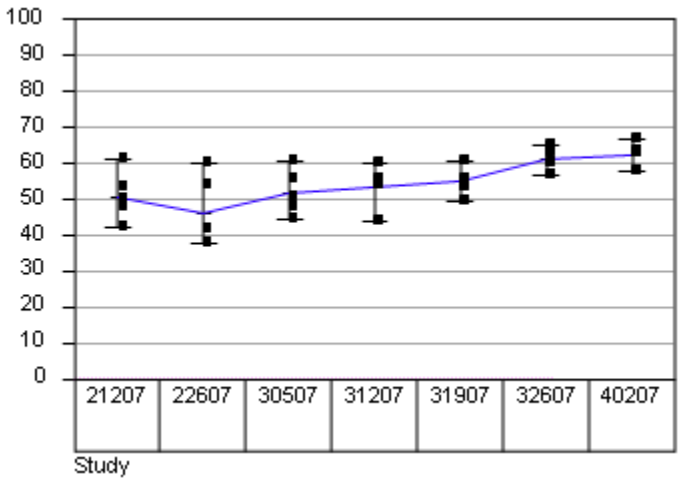


Figure 4: Variability chart of EnzymeX Desaturase Index versus Study.

2.2.7.2. Concurrent Validation of SDSs and DRCs

As mentioned in Section 2.2.2, SDSs should be validated separately from DRCs, but it is possible to validate both methods concurrently. Rather than beginning with the SDS validation via a Replicate-Determination Study (described in Section 2) using several compounds at a single dose, one could proceed directly to the DRC Replicate-Determination Study using a smaller number of compounds tested in dose-response format. The pairs of % activity determinations for the individual doses of each compound can be used to calculate an MSD that is applicable to the assay in single-dose format, while the pairs of ED_{50} determinations can be used to calculate an MSR that is applicable to the assay in dose-response format.

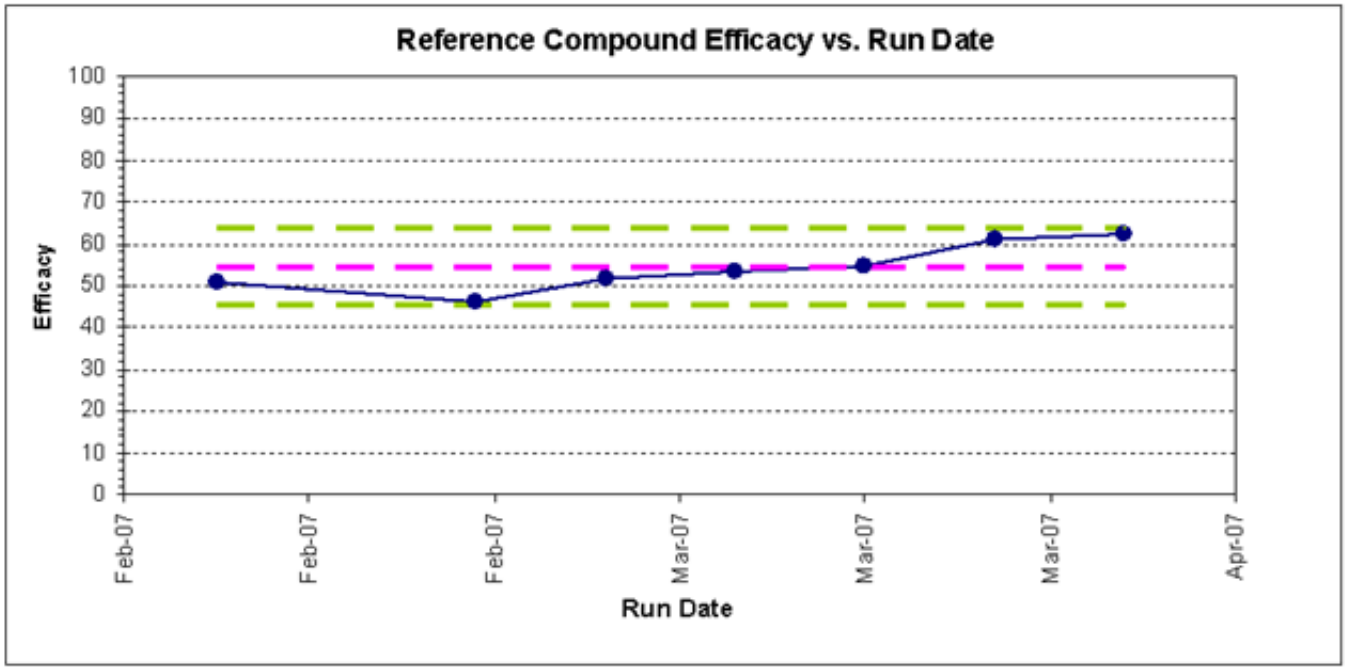
2.2.7.2.1. Example

The Capsaicin-Induced Eye Wipe Assay evaluates the ability of compounds to alleviate conjunctival irritation caused by exposure to capsaicin, as measured by inhibition of eye wiping in rats. The assay was developed and validated in Lab A, and upon completion of these activities, the assay was transferred to Lab B. Since the operator in the receiving laboratory was well-trained with respect to running the assay, and since no other major changes to the protocol were made, a bridging study or "Single-Determination Study" was utilized to validate the transfer (see Section 2.4 for additional details).

Three compounds were selected and tested in dose-response format with n=6 rats per each of four doses plus a capsaicin control for a total of 30 animals. The three compounds were tested once in the originating lab and once in the receiving lab. Percent inhibitions of the capsaicin-induced numbers of eye wipes are displayed in the variability chart (Figure 6). One compound was lost for technical reasons in one of the labs, so only two are charted.

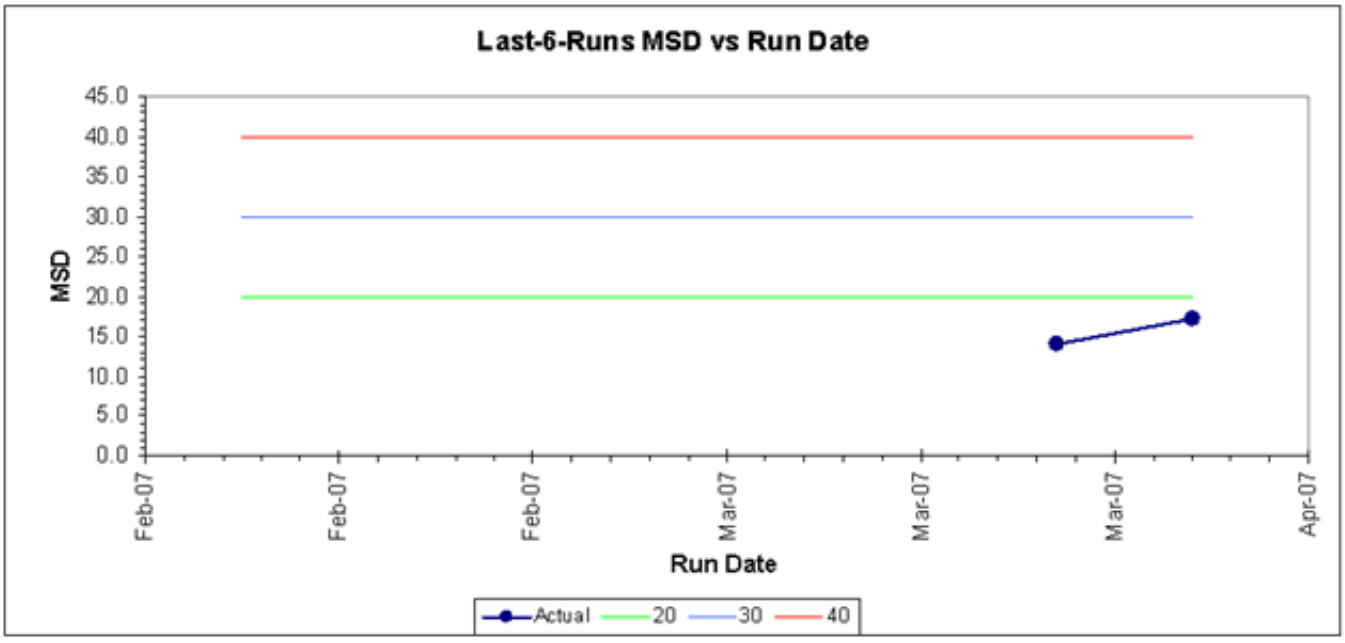
The variability chart indicates similar-dose response relationships for the two compounds between labs. There does however appear to be a consistent shift in the % inhibition for the individual compounds between labs. Also note that there is not complete overlap in the tested doses between labs, but each lab has four doses plus capsaicin available for estimation of the ED_{50} and 6 pairs of single-dose % inhibitions for calculation of the Mean Difference, Minimum Significant Difference, and Limits of Agreement.

This Single-Determination Study is also an example in which each lab's set of three determinations spanned multiple runs of the assay (i.e., each compound was tested in a separate run) as described in Section 2.2.4 Note 1. Teams should consult their statistician for appropriate MD, MSD, and LsAd calculations when determinations span multiple runs.



Control Limits

Mean (Center Line)	54.38
Upper Control Limit	63.60
Lower Control Limit	45.16



Mean Summary		**MSD Summary**	
High	62.48	High	17.15
Low	46.20	Low	14.02
Overall	54.38	Overall	16.31
Current	62.48	Current	17.15

Figure 5: Output from entering percent inhibition averages into the control chart template.

For the data in this example, the MD is 17%, the MSD is 11%, and the LsAd are 5.1% and 28%. While the MSD is excellent at 11%, the LsAd do not fall between -20% to +20 due to the 17% average shift in % inhibition between labs. While ED_{50}s from just two compounds are not enough to calculate an MSR, the ED_{50}s estimated from a four-parameter logistic fit also exhibited a shift in potency between labs, but the magnitude of the shift was less than 3-fold for both compounds.

Even though the LsAd did not meet the within ±20% criterion, the transfer validation was accepted and the assay was implemented in the receiving laboratory. This decision was based on discussion with scientists regarding possible explanations for the shift, as well as implications for the projects the assay supports. While the LsAd failed the ±20% criterion, the limits did fall within ±30%, and the ED_{50}s from the two compounds fell within 3-fold, which taken together was considered acceptable for the teams' needs. No MSR could be calculated, but since each compound's dose-response curve consisted of 30 independently tested animals (thereby providing confidence in the resulting ED_{50}s), and the ED_{50}s agreed to within 3-fold, the transfer was considered to be validated in both single-point and dose-response formats. Close monitoring of a quality control was also recommended for the assay going forward.

Another alternative approach is to include the dose used for the single-dose screen in dose-response determinations, and using the resulting repeats to calculate an MSD.

2.3. In-study Validation (Single-Dose Screens)

As mentioned previously in Section 2.2.1, the Replicate-Determination study is used to formally evaluate *within-run* assay variability. The study also provides a preliminary assessment of the *overall* assay variability, but typically does not involve enough assay runs to adequately assess *overall* variability. Overall assay variability consists of both within-run and *between-run* sources of variability, and hence needs to be estimated from several runs of the assay.

Post-implementation monitoring of the measured activity of a control compound that is regularly included in runs of the assay is an effective means for estimating the overall variability in the assay, which can then be used to calculate an overall MSD. The resulting MSD will likely be larger than the Replicate-Determination MSD, since the between-run variability is now included, but it is the appropriate MSD for comparing and prioritizing compounds tested in different runs of the assay.

Continuous monitoring of a regularly tested control compound is also an effective means of tracking assay performance over time, thereby ensuring high quality data for compound prioritization. Control charting the activity of the same compound over time ensures that the activity remains stable without appreciable "assay drift," and enables identification of suspect runs. Tracking assay performance over time will also ensure that the reproducibility of the assay (i.e., the MSD) remains at an acceptable level.

2.3.1. Control Compounds or Treatment Groups

Control compounds in *in vivo* experiments serve three purposes: (i) as comparator for the test compounds, (ii) to normalize responses across assay runs, and (iii) as a quality control marker. A ***negative* (minimally effect) *control*** is typically included in all runs of an assay and serves as the comparator for test compounds. A ***positive* (maximally effective) *control*** is also often included to establish that the assay is working and to normalize the response over assay runs. While these so-called minimum and maximum controls can be monitored over time to ensure that adequate separation is maintained, they may not be the best choice for monitoring assay performance and reproducibility as described in Section 2.3. Since these controls lie at the extremes of the dynamic range of the assay, they may not accurately represent the level of variability present within the interior of the range (i.e., assay variability may be different for low, medium, and high activity levels). Since interest often lies in prioritizing compounds whose activities fall in the middle to upper end of the dynamic range (but not necessarily at the top of the range), an additional control compound with a normalized (to positive and negative

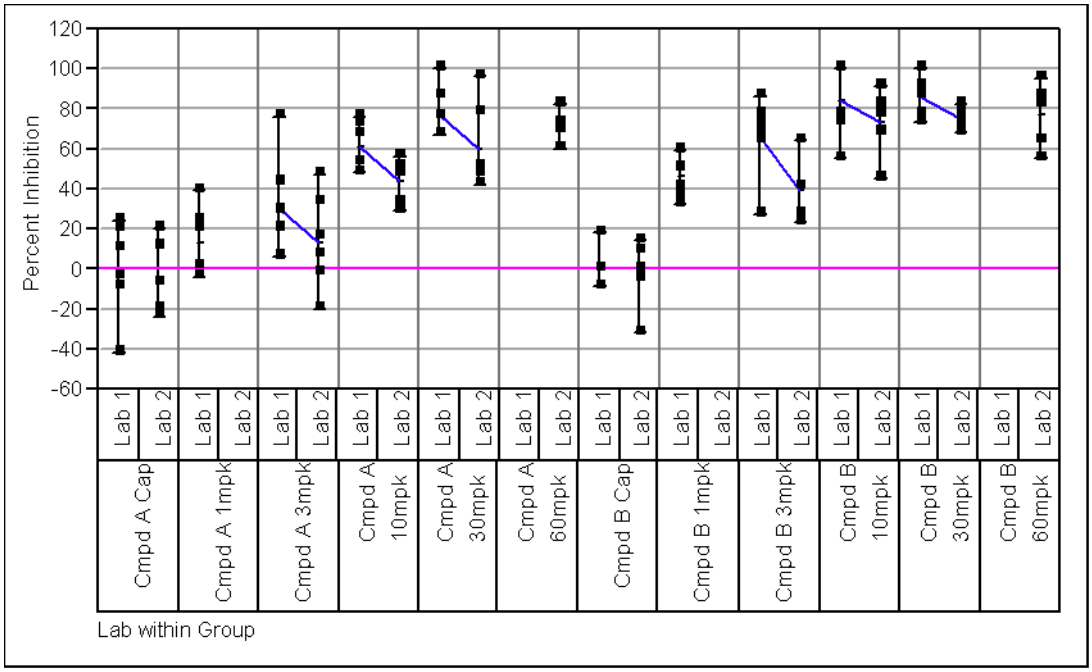


Figure 6: Variability chart for percent inhibition in a Single-Determination Study.

controls) activity of 50 to 70% should also be included to adequately monitor assay performance and reproducibility. This ***quality control*** should be included in each run of the assay, if possible, but could be included periodically (e.g. every other or every third run) if practical limitations (cost, time, resources etc.) exist for the assay.

2.3.2. Control Charts

Monitoring assay performance can be accomplished with a simple scatter plot of the % activity of the quality control versus run date, which should be updated after each run of the assay. Assay drift (trends up or down) can be identified visually, and problems investigated and corrected as they occur. Suspect runs should also be investigated and repeated if warranted.

Full-fledged control charts that calculate and display control limits are preferred, since the limits can assist with identifying trends or steps in the data. Control chart monitoring templates that produce scatter plots, as well as the preferred control charts, for *in vivo* assays are available (check with your statistician). Both templates can be used for *in vivo* data by first normalizing the individual animal data to the positive and negative control averages in each run, and entering the % activity averages for the quality control from each run into the templates.

After six runs, both templates will also calculate the overall MSD, which incorporates both within-run and between-run sources of variability (in contrast to the Replicate-Determination MSD, which encompasses only within-run variability). As such, it is expected that inclusion of the within-run variability will inflate the overall MSD as compared to the Replicate-Determination MSD. Teams should consult with their statistician about the interpretation of the overall MSD and how it affects comparing activities of compounds (and subsequent prioritization) tested in different runs of the assay. After each subsequent run, a running MSD (i.e., last-6-runs MSD computed from last six runs) is also computed and displayed graphically to enable monitoring of the reproducibility of the assay over time to ensure that it remains at an acceptable level.

2.3.2.1. Example

What follows is an example of an *in vivo* assay which measures the inhibition of EnzymeX Desaturase Index in DIO rats, and which was discussed previously in Section 2.2.5. A control compound was included in 14 runs of the assay. After normalizing the individual animal EnzymeX DIs to the positive and negative controls within each run, the average % inhibition (n=5 animals per group) of EnzymeX DI for the control compound in each run was entered into the control chart template. The output is shown in Figure 7.

Note the very stable activity of the quality control over the first 14 runs of the assay in the control chart of % activity versus run date. The % activity of the quality control tracks closely to the average of 79% over the 14 runs, with no material drifts or jumps in activity apparent in the chart.

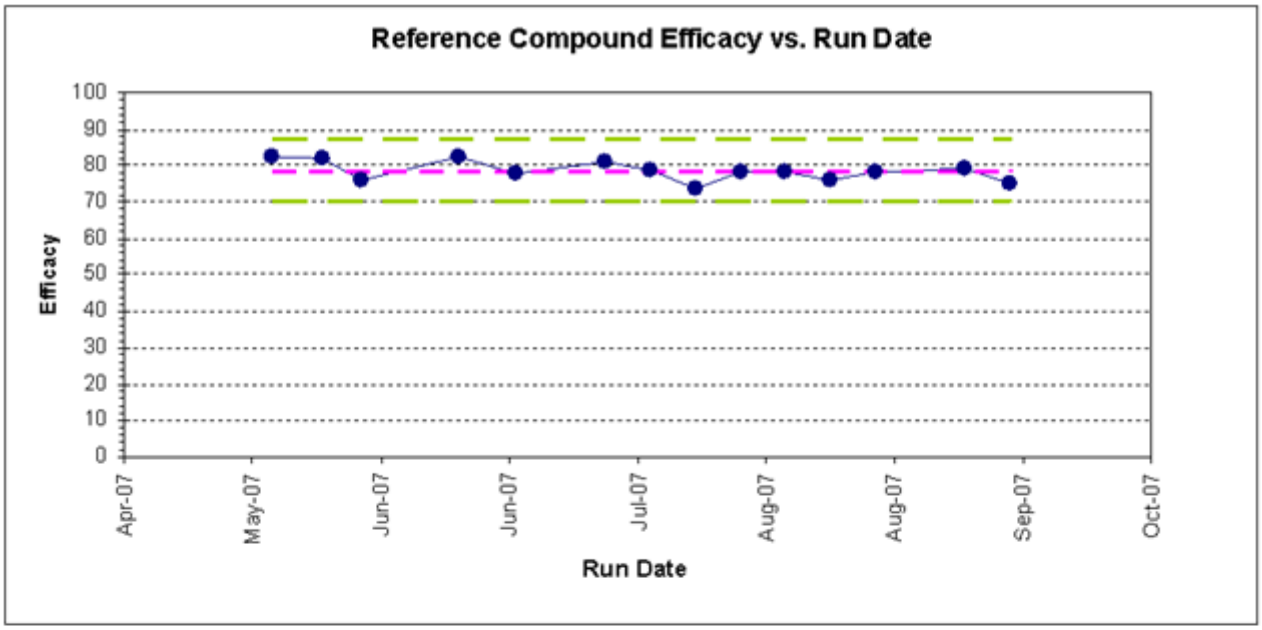
Also note the very stable level of reproducibility in the Last-6-Runs MSD versus run date chart (in fact, the reproducibility looks to be improving slightly over time). The running MSDs are well below 20% (actually fall below 10%), and the overall MSD is excellent at 7.6%, suggesting plenty of power to discriminate among compounds tested in different runs of the assay. However, it is important to note that the overall MSD of 7.64, which incorporates both within-run and between-run variability, is actually smaller than the Replicate-Determination MSD of 10.06 (Section 2.2.5), which incorporates only the within-run variability. This could be the result of the operator simply getting better at running the assay (the slight improvement in the running MSDs noted above provides some evidence for this), or it could be due to the fact that the quality control is quite active. While not maximally effective, the % activity of the quality control was certainly outside the recommended range of 50 to 70% activity for a quality control. Since variability (and hence reproducibility) may depend on the activity level of the control (with more active compounds expected to be less variable), the 79% active control may in fact be overestimating the true reproducibility of the assay.

2.4. Cross Validation: Bridging Studies for Assay Upgrades/Minor Changes versus Major Changes

Sections 2.2 and 2.3 cover the validation of entirely new assays, or assays that are intended to replace existing assays. The replacement assays are "different" from the original assay, either because of facility changes, personnel differences, or substantively different measurement and recording or automation equipment. Assay upgrades and changes occur as a natural part of the assay life cycle. Requiring a full validation for every conceivable change is impractical and would serve as a barrier to implementing assay improvements. Hence, full validation following every assay change is not recommended. Instead, bridging studies or "mini-validation" studies are recommended to document that the change does not degrade the quality of the data generated by the new assay. In addition, if the assay is to report results in an Assay Method Version (AMV) previously reported into by another assay then it has to be verified that the two labs produce equivalent results.

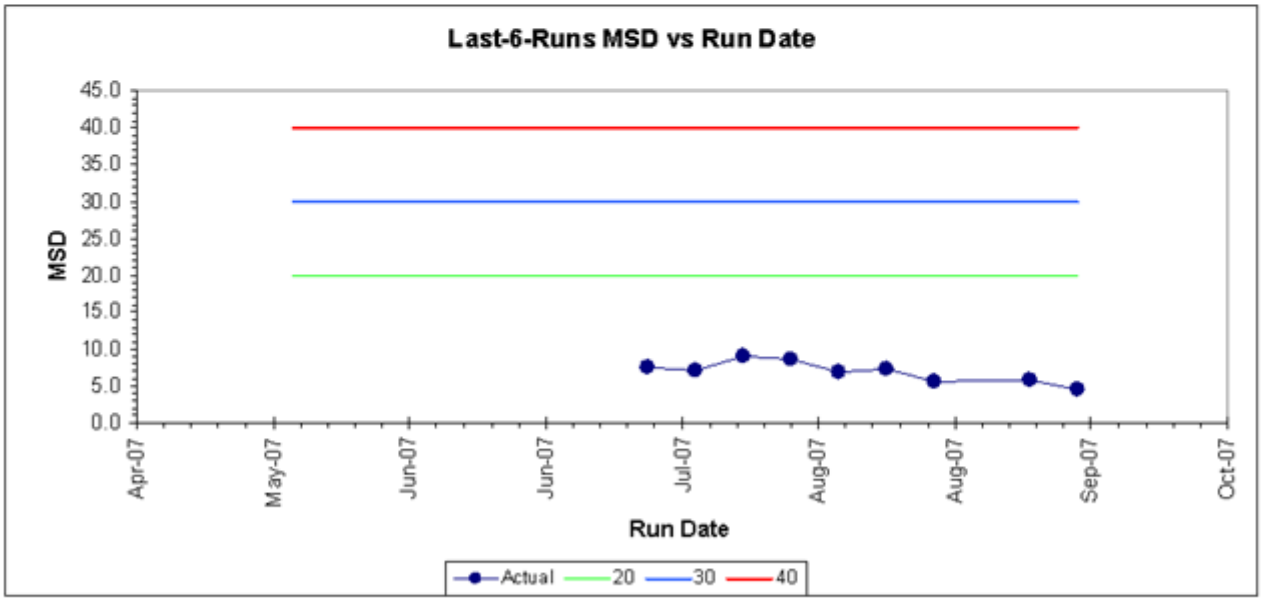
The level of validation recommended has two tiers, either a Single-Determination Study (Tier I), or a Replicate-Determination Study (Tier II) similar to the full validation package of Section 2.2. Examples of changes within each Tier are given below, along with the recommended validation study for that tier. Note that if the study indicates the change will have an adverse impact on assay quality (i.e., the study indicates there are problems), then the cause should be investigated, corrected, and a full validation should be done. If the results from that study indicate the assays are not equivalent, but the new assay is acceptable, then a new AMV should be established for the assay.

The following guidelines apply principally to changes in biological components of the protocol. If changes are made to the data analysis protocol then these can ordinarily be validated without generating any new data, but rather by comparing the results using the original and new data analysis protocols on a set of existing data. Discuss any changes with a statistician. If changes are made to both the data analysis and biological components



Control Limits

Mean (Center Line)	78.54
Upper Control Limit	87.24
Lower Control Limit	69.83



Mean Summary		**MSD Summary**	
High	82.46	High	9.07
Low	73.77	Low	4.45
Overall	78.54	Overall	7.64
Current	75.31	Current	4.45

Figure 7: Output from entering average percent inhibition of EnzymeX DI into the control chart template.

of the protocol then the appropriate tier should be selected according to the severity of the biological change as discussed below. The data analysis changes should be validated on the new validation data.

2.4.1. Tier I: Single Step and/or Minor Changes to the Assay

Tier I modifications are single changes in an assay such as a change in the person running the assay, an assay condition, instrumentation, or to a *reagent,* that is made either to improve the assay quality or increase the capacity without changing the assay quality. The changes can also be made for reasons unrelated to assay throughput or performance (e.g., change of a supplier for cost savings). Examples of such changes are

- Changes in supplier of animals
- Change in barrier facility used by the supplier
- Changes in recording instruments with similar or comparable electronics. E.g.: blood pressure recording instruments, clinical chemistry equipment, HPLCs, spectrophotometers, behavioral testing equipment such as locomotor activity instruments or operant conditioning chambers. A performance check for signal dynamic range, and signal stability is recommended prior to switching instruments.
- For *ex vivo* analysis of tissues, changes in liquid handling equipment with similar or comparable volume dispensing capabilities. Volume calibration of the new instrument is recommended prior to switching instruments. [Note that plate and pipette tip materials can cause significant changes in derived results (IC_{50}, EC_{50}). This may be due to changes in the adsorption and wetting properties of the plastic material employed by vendors. Under these conditions a full validation may be required].
- Changes in dilution protocols covering the same concentration range for the concentration–response curves. A bridging study is recommended when dilution protocol changes are required.
- Lot changes of critical reagents such as a new lot of receptor membranes or a new lot of serum antibodies.
- Assay moved to a new laboratory without major changes in instrumentation, using the same reagent lots, same operators (or operators with similar experience and/or training), and assay protocols.
- Assay transfer to an associate or technician within the same laboratory having substantial experience in the assay platform, biology and pharmacology. No other changes are made to the assay.

The purpose of the validation study is to document that the change does not reduce the assay quality.

2.4.2.1. Protocol and Analysis

Conduct the assay comparison portion of the Replicate-Determination Study discussed in Section 2.2, i.e., compare one run of a minimum of 3 to 5 compounds using the existing assay to one run of the assay under the proposed format. If the compound set used in the original validation is available then one need to only run the set again in the new assay protocol, and compare back to Run 1 of the original Replicate-Determination Study. The acceptance criterion is the same as for the assay comparison study: the assay should have Limits of Agreement between -20% and +20% for % activity, and an MSR < 3 and Limits of Agreement should between 0.33 and 3.0 for potency.

2.4.2. Tier II: Substantive Changes

Substantive changes requiring full assay validation: when substantive changes are made in the assay procedures, measured signal responses, target pharmacology and control compound activity values may change significantly. Under these circumstances, the assay should be re-validated according to methods described in Section 2.2. The following changes constitute substantive changes, particularly when multiple changes in factors listed below are involved:

- Changes in strain of animals: e.g., SD to Fischer
- Transfer of the assay to a different laboratory location, with distinctly different instrumentation, QB practices or training.

- Changes in detection instruments with significant difference in the optics and electronics. For example, blood pressure monitors, behavioral test equipment, counting equipment, spectrophotometers, and plate readers.
- Changes in assay platform: e.g.: filter binding to LS/MS detection for ex vivo binding assays.
- Changes in assay reagents (including lot changes and supplier) that produce significant changes in assay response, pharmacology and control activity values. For example, changes in enzyme substrates, isozymes, cell-lines, label types, control compounds, calibration standards, (radiolabel vs. fluorescent label), plates, tips and bead types, major changes in buffer composition and pH, co-factors, metal ions, etc.
- Changes in liquid handling equipment with significant differences in volume dispensing capabilities.
- Changes in liquid handling protocol with significant differences in volume dispensing methods.
- Changes in assay conditions such as shaking, incubation time, or temperature that produce significant change in assay response, pharmacology and control activity values.
- Major changes in dilution protocols involving mixed solvents, number of dilution steps and changes in concentration range for the concentration-response curves.
- Change in analyst/operator running the assay, particularly if new to the job and/or has no experience in running the assay in its current format/assay platform.
- Making more than one of the above-mentioned changes to the assay protocol at any one time.

Substantive changes typically require full validation, i.e., a complete Replicate-Determination Study. If the intent is to report the data in the same AMV then an assay comparison study must be conducted as part of the Replicate-Determination study.

2.5. How to Deal with High Assay Variability

As mentioned previously, the sources of variation in an assay include the between- and within-run sources of variability. In addition, the animal-to-animal variability as well as variability in measuring the response of the subject also contributes to the overall variability in the assay. In order to optimize an assay and/or when an assay fails to meet the acceptance criteria, it is important to specifically assess each of these sources of variability. The variance, which is the square of the standard deviation, can be used to estimate the magnitude, or relative contribution, of each of these sources of variability. Variance is useful because the sources of variation are additive on the variance scale, but not on the standard deviation scale. When an assay fails to meet the acceptance criteria, it is necessary to determine the source of the high assay variability in order to be able to make changes to reduce the relevant variability. For example, simply increasing the number of animals per group may not necessarily reduce between-run variability.

2.5.1. Example: Analyzing Variability in an Ex Vivo Binding Assay

In an *ex vivo* receptor binding assay, subjects (5 rats per dose group) were administered a dose of a test compound orally, sacrificed 1 hour later, the cerebellum removed and stored at -70° until used in the binding assay. For the *ex vivo* receptor binding assay, the tissue was homogenized and incubated for 30 minutes at 37°C in a water bath. The tissue from each animal was aliquoted into 8 tubes, together with radioligand and buffer, and incubated for 2 hours. Four tubes were used to measure total binding and four tubes to measure non-specific binding. Each tube was centrifuged at 12000 x g for 5 minutes, washed and then placed in a gamma counter and counted. Counts for each tube were converted to DPM and recorded. A plot of the data from each of the vehicle treated animals for each of 3 experiments is presented in Figure 8. The variance in the assay is summarized in Table 2.

The total variance was partitioned statistically into study-to-study, animal-to-animal, and tube-to-tube variability. In this example, the result for each of the four tubes is the difference between the DPM in one of the four tubes used to estimate total binding minus the mean of the four tubes used to estimate nonspecific binding. Thus, each of the four tubes represents a different measurement of one subject and the variability among the four

tubes from one animal represents variability in the measurement of binding in that animal. In this example, study-to-study variability accounts for approximately 68% of the total (Table 2, far right column). In order to optimize the assay and reduce the overall variability, reducing the study-to-study variability would have the greatest impact. One way of reducing study-to-study variability is to normalize the data within each study in order to compare results across each study. In this example, the radiolabel was ^{125}I, and its relatively short half-life is the major source of the study-to-study variability; normalizing the data within each study would reduce study-to-study variability due to isotopic decay. The need to normalize, rather than using raw counts, is well known and accepted for *in vitro* binding studies, but this practice is not common for *in vivo* assays. It is appropriate to compare raw signals across studies only when the study-to-study variation is negligible. Note that increasing the sample size would not reduce this source of variability. The second largest source of variability, approximately 18%, is tube-to-tube, that is, the variability in measuring each animal. In order to reduce this source of variability, additional measurements of each animal would be needed. Using 5 tubes to assess specific binding, rather than 4, and only 3 tubes to assess nonspecific binding (which is typically not highly variable) would reduce this source of variability. If only one measurement on an animal is made, the measurement-to-measurement variability still exists, but cannot be calculated. (Another example is that of measuring tumor volume with calipers – measuring only once at a given assessment versus measuring multiple times at a given assessment.) Note that increasing the sample size would not reduce this source of variability. Multiple measurements of the response for one subject are the exception rather than the rule for *in vivo* experiments, and yet variability in the measurement process may be a major source of variance in the study. It is important during assay development and optimization to assess the potential contribution of measurement variability on the total variability in the assay and address this potential source as needed, for example by taking the average (or median) of multiple measures on a given subject. The smallest source of variability in the present example, approximately 13% of the total, was due to animal-to-animal variability. Thus, increasing sample size would have impacted the smallest source of variability. The relative contributions of "animal" and "measurement" variation can be used to determine the optimal number of measurements per animal, as well as number of animals. In conclusion, the relative contribution from different sources of variability needs to be directly assessed during assay development and optimization. This assessment allows one to directly address the most relevant sources of variability in order to optimize and statistically validate an assay.

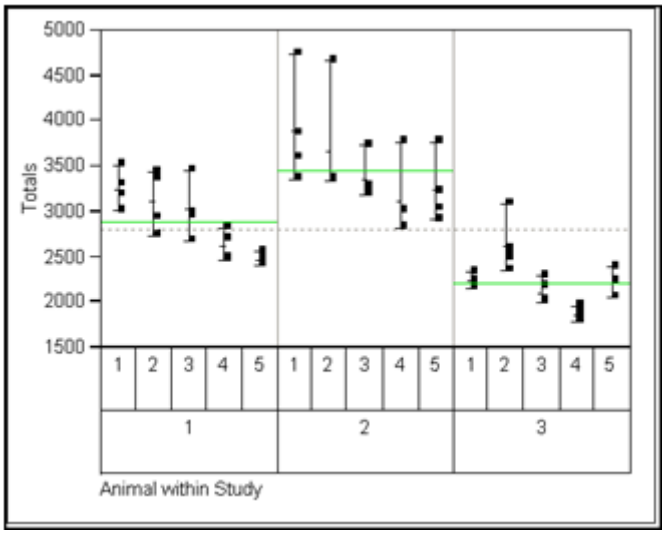


Figure 8: Plot of the vehicle treated animals in an *ex vivo* binding assay. The ordinate (y-axis) is total specific binding in DPM in each of the four tubes used to assess specific binding. Data for 5 animals in each of three studies is represented.

Table 2: Estimated variance contributed by study-to-study variability, animal-to-animal variability, and tube-to-tube variability in an *ex vivo* binding assay.

Source of Variation	Estimated Variance	Estimated Std Dev	Pct of Total (%)
Study	373088	610.8	68.3
Animal	72819	269.9	13.3
Tube	100033	316.3	18.3
Total	545941		100

2.5.2. High Variation in Single Dose Determinations

Table 3 below can be used as a reference to determine the sample size necessary for single dose or dose-response assays with high variability. For a given **coefficient of variation** (CV) of the raw data values based on a sample size of 1 subject, the table shows the number of subjects per dose needed for the CV of a mean to be less than or equal to 10 or 20%.

Increasing sample size to reduce variability will also reduce the capacity (i.e., throughput) of the assay to test compounds. Further optimization of the assay could reduce variability and maintain or increase its capacity. The decision to further optimize or to increase sample size will have to be made for each assay.

Table 3: Sample size necessary to reduce the coefficient of variation (CV) to less than 10 or 20%, given a known CV when the sample size is one.

CV for Individual Subjects	Number of Subjects so that CV Mean < 10%	Number of Subjects so that CV Mean < 20%
<10	1	1
10.1-14.1	2	1
14.2-17.3	3	1
17.4-20.0	4	1
20.1-22.3	5	2
22.4-24.4	6	2

Table 3: continued from previous page.

24.5-26.4	7	2
26.5-28.2	8	2
28.3-30	9	3
30.1-31.6	10	3
31.7-33.1	11	3
33.2-34.6	12	3
34.7-36.0	13	4
36.1-37.4	14	4
37.5-38.7	15	4
38.8-40.0	16	4

2.5.3. High Variation in Dose-Response Determinations

If in Section 2.2 the assay fails either test (MSR > 3 or Limits of Agreement outside the interval 0.33-3) then the variability of the assay is too high for typical purposes. The following options should be considered to reduce the assay variability:

1. Optimizing the assay to lower the variability in the raw data values. Check that the dose range is appropriate for the compound results. Increasing the number of doses and/or subjects per dose may improve the results. A minimum of 6 doses at 2X intervals, and analyzing the data using nonlinear curve-fitting techniques, is recommended. In general, it is better to have more doses rather than more subjects per dose. The doses should cover the expected range of the assay, e.g., 0-100%, as much as possible.
2. Consider increasing sample size as discussed below. Note that the impact of increasing sample size may decrease capacity, and so the Replicate-Determination Study, and a detailed analysis of the sources of variation, should be used to assess whether increasing the number of subjects per dose will achieve the objective.
3. Adopt as part of the standard protocol to re-test compounds. For example, each compound may be tested 2 or more times in different runs of the assay. Averaging the results from multiple runs will reduce the assay variability (**NB**. In such cases the individual run results may be stored in the database and then the data mining/query tools are used to average the results).

To investigate the impact of increasing sample size in the dose-response assay you should conduct the Replicate-Determination Study with the maximum number of subjects contemplated (e.g., 5 subjects / dose). The data can be analyzed first using all available subjects. Then one subject per group can be removed at random and the data re-analyzed. This step is repeated until the smallest sample size is found that still meets the acceptance criteria. An example below will illustrate this idea.

An *in vivo* receptor occupancy assay was run using 1 subject per dose and the Replicate-Determination Study did not meet the acceptance criteria. To examine if replication (i.e., increasing the number of subjects per dose) would help, a new Replicate-Determination Study was conducted using 4 subjects per dose. Table 4 shows the results of fitting ED_{50} curves and re-evaluating the MSR and LsA for 2, 3, or 4 subjects per group:

From Table 4 we can see that it takes all 4 subjects to meet the MSR acceptance criteria, and more than 4 subjects would be needed to meet LsA acceptance criterion. It should be noted that the LsA results are close to being acceptable with 4 subjects per group.

Table 4: Results of fitting ED_{50} curves and re-evaluating the MSR and LsA for 2, 3, or 4 subjects per group.

Subjects	MSR	LsA
2	3.62	0.35 – 4.59
3	3.32	0.43 – 4.74
4	2.44	0.53 – 3.16

3. Design, Sample Size, Randomization, and Analysis Considerations

3.1. Assay (Experimental) Design Considerations

Good experimental design is important in order to answer the research question of interest in a way that is free of bias, can be generalized to the desired or targeted population, and is of sufficient size to properly answer the question. This includes such things as determining the measurements to make, timing, dosing frequency and route, the species to use, etc. It also includes identifying the relevant statistical analyses, determining appropriate sample sizes, and determining a randomization scheme.

Several expectations for good experimental design and analysis include:

- An a priori statement of objectives/hypotheses
- Appropriate experimental design methods
- Sufficient, but not excessive, sample size to ensure statistical significance of biologically relevant effects
- A priori determination of appropriate statistical methods

3.1.1. Objectives and/or Hypotheses

State the objectives and/or the hypotheses you wish to accomplish or test with an assay before beginning an experiment. If a hypothesis is formed after collecting the data, the results may be biased. The objectives should be defined in terms of well defined end-point(s). Examples of objectives include comparison of food consumption between a test compound and control, comparison of survival rates between the treated and untreated groups, comparison of the effect level or the ED_{50} between a test compound and control, etc.

3.1.2. Design Strategy

In vivo studies should be designed in such a way that all meaningful biological effects are statistically significant. In an exploratory study, this "meaningful effect" might correspond to any effect that is pharmacologically relevant. For a project/program team, this meaningful effect might correspond to an effect that meets the CSFs defined in the flow scheme. Power and sample size analysis is especially relevant for assays that are designed to address key endpoints and make decisions as to whether a compound meets the assay's CSF. Biologically meaningful effects are not always well-known in advance, in which case a range of plausible effects could be considered.

To a great extent, experimental design is about identifying and determining a strategy to deal with different kinds of variables. The types of variables encountered in research include:

- Manipulated variable ***(independent/explanatory variable)***
- Response variable ***(dependent/outcome variable)***
- Extraneous variables ***(uncontrolled/random)***

The manipulated variable is a purposeful attempt to introduce variability into the experiment by, for example, administering different doses of a drug. If the manipulated variable were the only source of variability, then research design would be quite simple.

Extraneous variables can encroach upon an experiment and can change the results in ways that we do not want or of which we are unaware. Examples of extraneous variables could include inherent animal variation, time of day, baseline body weights or glucose levels, operator ability or skill, lab noise or activity, etc. **To a great extent, experimental design is about having a strategy to deal with extraneous variables**. To ignore them can all too often lead to biased results and the requirement of larger sample sizes. Fixing (holding them constant) or eliminating them, such as by considering only a subgroup of animals, can reduce bias and sample sizes, but can also reduce the generalizability of the results to only those conditions considered in the experiment. Another approach is to control for them by incorporating them into the experimental design, ideally at the design stage, or at the statistical analysis stage if the former is not possible.

Some additional design considerations include:

- Appropriate random allocation of animals to treatment groups.
- Blinding of observers to drug treatment allocation, whenever possible, especially when subjective evaluations are to be made by observers.
- Proper selection of dose levels.
- Optimal selection of control groups.
- Optimal time points to collect samples.
- Proper statistical methodology.

Different design strategies should be carefully considered to minimize variability and maximize information from the experiment.

The design issues stated above should be addressed in the context of the key endpoints (or summary measures) from the study. Examples of such endpoints may include survival rate, glucose normalization, etc. When there are several endpoints of interest from an assay, certain design questions such as the power of the study should be assessed with respect to the endpoints that are considered to be most important by the scientist and the project team.

3.1.3. Endpoints

The key endpoints from the study must be identified first, as all other design choices should be tailored to these outcomes. Typical outcomes include:

- Statistical significance from control using Analysis of Variance (ANOVA).
- ED_{50} (either absolute or relative) from a dose-response model such as the 4-parameter logistic model (4PL)

Note that the latter can be determined using a variety of methods, and it is important to define the pharmacologically "effective dose," as well.

3.1.4. Control Groups

Control groups serve three purposes: (i) a comparison to the test groups, (ii) as a quality control marker, and (iii) to normalize the response for comparison across studies. A "negative" control is used in all studies and serves as the comparison group. An active or "positive" control is a compound that has a different response from the negative control, and normally represents the maximal response of a standard treatment. A positive control is used when normalization is necessary to stabilize the response across runs of the assay or to illustrate the range of signal available in a particular run of the assay. If a positive control fails to separate from the negative control

then there is an increased chance of a false negative outcome. Since normalized responses may still lack reproducibility across runs, a third control may be employed to monitor the reproducibility of the normalized response across runs of the experiment. This control would be a second positive control, but at an activity level lower than the first positive control. These are used as a quality control marker to establish that each run of the experiment is performing as expected, and hence are often called a "quality control." The activity of a "quality control" should be at a level of activity desired for the advancement of test compounds (see Section 2.3).

3.1.5. Statistical Analysis Plan and Implementation

Before developing and validating an assay to be used on a flow scheme, appropriate statistical methods including data transformations and software for analyzing the data from these experiments should be determined.

There are several statistical methods available to analyze any given experiment/data set, and the choice of these methods and the way a certain class of methods is implemented can significantly impact the conclusions from the experiment. For example, there are certain statistical considerations one should take into account when using the analysis of variance (ANOVA) method, including the distribution of the data, equality of variances, baseline variables, methods for comparing different groups, etc. Also, a two-sample *t*-test might often seem appropriate for several types of experiments, but upon careful examination of the study design, the *t*-test might turn out to be less appropriate than some of the other statistical analysis methods for such experiments.

If the study design is changed at any time during a series of experiments, appropriate analysis methods and implementation strategy should once again be examined in light of these changes.

The basic types of experimental designs are:

- Parallel group
- Randomized block
- Repeated measures
- Cross-over design

3.1.5.1. Parallel Groups Design

In a parallel groups design, subjects are randomly assigned to groups and each group receives one level of a treatment, for example, one dose of a drug, so each group is independent of every other group. This basic design assumes that there are no important *extraneous variables* that we can identify which will influence or bias the results.

Features of a parallel groups design:

- Simplest design.
- Subjects are randomly assigned to groups, and groups are typically, but not necessarily, of equal size.
- Each group receives one level of a treatment, e.g., one dose of a drug
- Use when randomization is possible.
- Does not account for extraneous variables that influence or bias the results. The variation caused by extraneous variables is attributed to the overall assay variability.

3.1.5.2. Randomized Block Design

A second basic design type is called a Randomized Block Design. Randomized block designs are used when an extraneous variable can be identified prior to randomization and subjects can be divided into subgroups based on values of the extraneous variable. Like a parallel groups design, each treatment group is an independent group of subjects. However, subjects are not assigned to treatment groups in an entirely random manner. Rather, subjects are first placed into one of several subgroups based on a blocking or matching factor (such as baseline

values, time of day, gender, baseline body weight, etc.) and then subjects in each block are randomized to the treatment groups.

It is necessary for each subgroup (or "block") to contribute equally to each treatment. Each subgroup must contribute an equal number of subjects to each treatment. Thus, there must be at least as many subjects in each subgroup as there are treatments. For example, if there are 4 treatments, there must be at least 4 or ideally a multiple of 4 subjects in each subgroup.

It is important that a separate randomization be performed on each block so that high or low values of each subgroup are not always placed into one treatment. This strategy forces the extraneous variable to be balanced across treatment groups.

One disadvantage of the randomized block design is that it is not always logistically feasible. For example, the investigator may be aware of an extraneous variable, but not be able to measure it before the randomization process, such as the size of a cardiac or cerebral infarct.

Features of a randomized block design:

- An extraneous variable can be identified and measured before starting the experiment.
- Subjects can be divided into subgroups based on values of the extraneous variable.
- Each subgroup ("block") has as many subjects as there are treatment levels.
- Within each block, treatments are randomly allocated to subjects.
- A separate randomization is performed in each block.
- Forces the extraneous variable to be balanced across treatment groups.
- Not always logistically feasible.

Note: we refer to a factor as a "blocking" factor when it is a continuous measure, for example baseline blood glucose levels, that we can divide into different levels, such as high, medium and low. We refer to a factor as a "stratification" factor when it is not continuous, such as gender and we can stratify the groups on the basis of that parameter.

In a randomized block design, it is necessary for all subjects within a block to be as similar as possible. There are several ways in which blocking can be accomplished:

One of the most common ways to "match" subjects is to rank all of the subjects (e.g., 1 through 20) according to each subject's value of the blocking factor. Subgroups of subjects are then grouped into individual blocks (e.g., subjects 1 – 4 as block 1, subjects 5 – 8 as block 2, etc.). So, within a block, there are similar values of the blocking factor. This minimizes the variance due to the blocking factor ("extraneous variable").

3.1.5.3. Analysis of Covariance (ANCOVA)

Sometimes it is not possible to identify or account for an extraneous variable at the beginning of a study or design phase. In these instances, it may be still be possible to remove the effect of an extraneous variable during the analysis. One technique is to employ an Analysis of Covariance or ANCOVA. Note that if the treatment affects both the response and the covariate then ANCOVA must not be used, because any observed effect may be due to actions of the treatment on the covariate rather than on the response variable.

An ANCOVA model is most useful when there is a linear relationship between the response and the covariate within each group, and when these slopes are similar. If there is sufficient sample size, it is relatively straightforward to test for these conditions. If the slopes aren't different from zero then there is little to no benefit in the ANCOVA model. If the slopes aren't parallel then the interpretation of the treatment comparisons depends on the level of the covariate.

It is also very important that the values of the covariate overlap among the treatments. If the values for the covariate don't overlap, then you are extrapolating into regions where you have no data and the results could be incorrect. Use of proper randomization techniques will usually prevent this situation.

If the experiment was designed as a randomized block design, it is generally best to analyze it as a randomized block design and avoid using ANCOVA.

You may be able to avoid an ANCOVA model by using the baseline to normalize the response (i.e., change from baseline, ratio, etc.; Figure 9).

3.1.5.4. Repeated Measures and Cross-over Designs

Another popular design is to use each subject as a block and test each subject at each of several time points (repeated measures design) or under each treatment condition (cross-over design). With this approach there is only one subject within a block and this minimizes the variance by using each subject as its own control. Repeated measures and crossover designs are just special cases of a randomized block design.

Some examples:

- When a subject receives a dose of drug and is tested at multiple time points.
- When a subject receives all doses of drug. This type of design is also a repeated measures design, but it is a special case of repeated measures and is referred to as a crossover design because each subject is "crossed over" to each treatment or treatment level.
- In a crossover design, different sequences of treatments are identified, and each subject is randomized to one sequence.
- Cross-over designs assume that the effects of each treatment dissipate or don't interfere with the response of the next treatment (i.e., no carryover). If this is not the case, then the cross-over is not an appropriate design.

Repeated measures designs allow us to separate out the variability due to individual differences (that is, to use each individual as their own control) and therefore better evaluate the effects of the manipulated, or independent variable. This increases the power of the analysis and means that fewer subjects are needed to have adequate statistical power.

3.2. Statistical Analysis Considerations

Once the experimental design has been selected, the appropriate statistical analysis then follows:

Parallel groups

- One factor with **only two levels**: t-test or ANCOVA
- One factor with more than two levels: one-way ANOVA or ANCOVA
- Two factors: two-way ANOVA or ANCOVA

Randomized blocks

- One factor: two-way ANOVA
  - Block as the second factor
- Two factors: three-way ANOVA
  - Block as the third factor
- Crossover & repeated measures: two-way and repeated-measures ANOVA
- Stratified designs: two-way ANOVA

The main analysis issues deal with how well the response data match the probability model assumed by the statistical analysis.

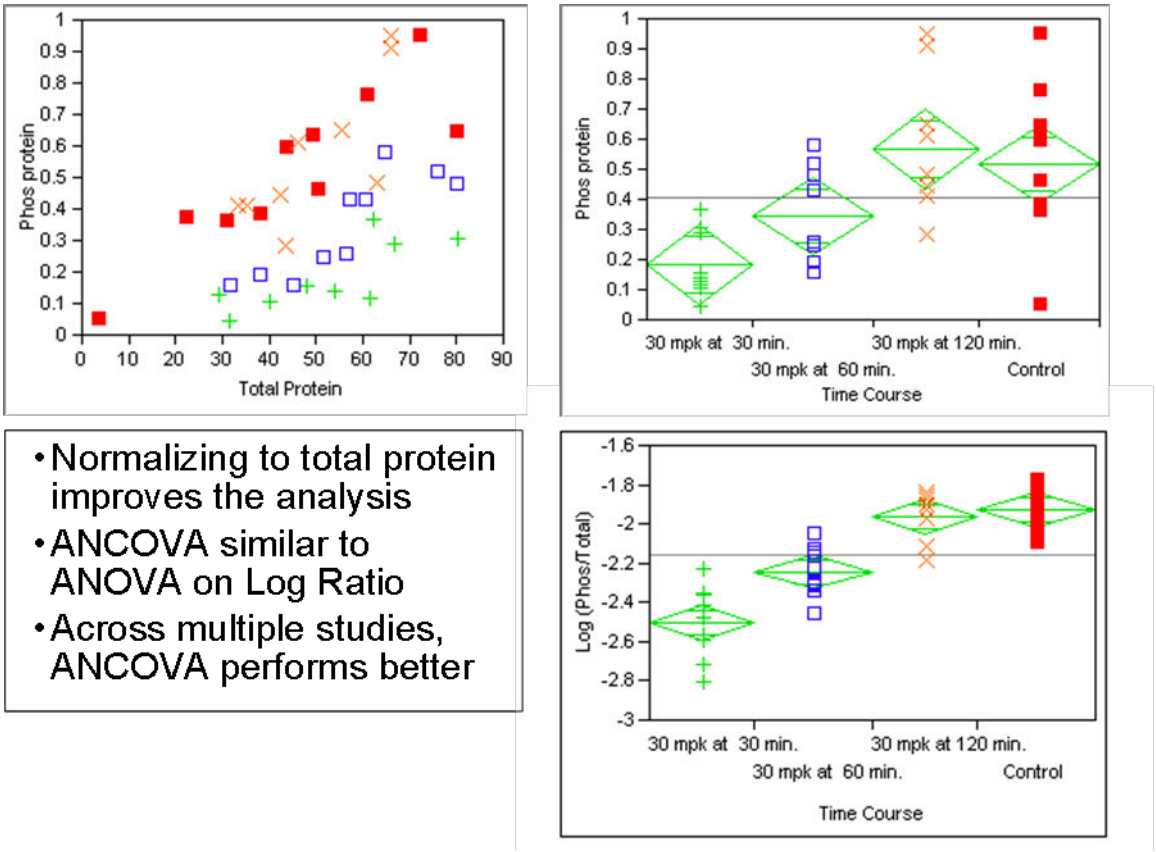


Figure 9: Examples of Analysis of Covariance (ANCOVA).

For ANOVA to identify statistically significant treatments or doses, the main issues are to verify the following.

- The residuals of the response or dependant variable are normally distributed with constant variability across groups
- Techniques to handle outliers are appropriate (see below)
- Appropriate multiple comparison techniques are employed for the study objectives
- If concomitant variables such as baseline measures are used, the analysis should appropriately adjust for these variables

For dose-response studies to estimate an ED_{50}, the main issues to be examined are the following:

- Whether the variability is constant or varies with the magnitude of the mean response
- Suitability of the dose range for the dose-response model
- Suitability of the dose-response model for the biological effects being examined.

For example, to fit a 4 parameter logistic model successfully, a wide dose-range allowing for well-defined tops and bottoms is required. In practice that is difficult to achieve, and often the top or the bottom is fixed at some pre-specified value. This imposes additional biological assumptions that should be assessed.

3.2.1. Outliers

The occurrence of an occasional outlier is not uncommon with *in vivo* experiments, and their presence can sometimes have a large impact on calculated means and, in particular, standard deviations. An observation isolated far above (or below) the bulk of the rest of data can "pull" the calculated mean toward it and result in an overestimate (or underestimate) of the true mean. Since the standard deviation involves a squared term, the impact of an outlier can be even more dramatic, making the standard deviation larger.

In addition to negatively affecting the calculation of summary statistics, outliers can also affect the accuracy of the p-values generated by statistical tests, such as the paired and two-sample t-tests, and the ANOVA. One possible remedy is to perform the statistical test on transformed data (typically the square root or log transform). Transformation is indicated when variability is not constant (e.g., across treatment groups in a one-way ANOVA) and/or when the data are skewed (i.e., longer tail) to the right, but a transformation can also eliminate apparent outliers in some cases (Figure 10).

Another remedial approach is to employ a non-parametric or rank-based statistical test, in which raw data are replaced by their ranks. These are indicated when the data at hand are not adequately modeled by the normal (symmetric, bell-shaped) distribution. Statistical tests based on ranks also down-weight any outliers present in the data. Non-parametric analogs of the paired and two-sample t-tests, and the ANOVA are respectively the Wilcoxon signed rank test, the Mann-Whitney U test, and the Kruskal-Wallis test.

It is acceptable to report the usual summary statistics when presenting data from an *in vivo* experiment for which a transformation or nonparametric test was used for the statistical analysis. However, as mentioned above, outliers can distort the mean and standard deviation, so a better approach is to report summary statistics consistent with the chosen remedial measure (e.g., use the antilog of the mean for log-transformed data; use the median for rank-based tests). Consult a statistician for the appropriate method of reporting summary statistics when outliers are present.

Another approach to dealing with outliers is to simply remove them, but this of course requires one to be able to discern which observations are truly erroneous and which simply represent the underlying variability in the assay. The outlier boxplot is a commonly employed and effective tool for identifying outliers. **However, detection of an outlier by this (or other methods) does not automatically mean the observation can be removed**. It simply identifies observations that have the potential to disrupt the statistical analysis and that should be investigated.

If a valid, assignable cause can be identified for an outlying result, or if the result is simply inconsistent with what is being measured and suggests an error was made, the observation can be removed. Otherwise, the analysis should be performed with and without the outlier(s), with both results reported. **If an observation is removed for cause, it should be documented that the data point existed along with the reason for removing it**. Ad hoc rules based on distance in standard deviations units from the mean should not be used.

Outlier box-plots are shown above the histograms in Figure 10. The box is formed by the first and third quartiles and represents the middle 50% of the data. The length of the box (i.e., third quartile minus first quartile) is the inter-quartile range (IQR). The vertical line within the box is the median, and the horizontal lines connected to the box extend to the extremes of the data that fall within 1.5 times the IQR. Any observations falling outside 1.5 times the IQR from the first or third quartiles appear as points on the box-plot and are potential outliers.

A minimum of 10 observations are recommended to generate a box-plot, but in the context of an experiment comparing several compound activities, each individual compound may be tested in fewer than 10 animals. In that case, the desired statistical model (e.g., a one-way ANOVA) can be fit and the outlier box-plot generated on the residuals from the model fit.

Residuals are automatically calculated in most statistical software, and in some cases are very simple to calculate. For example, the residuals from a one-way ANOVA of compound activities are simply the data "centered" by subtracting the respective compound mean. Consult a statistician for help with generating residuals from statistical models.

Figure 10 is an example of dog plasma exposures for 8 formulations of a single compound (4 dogs per formulation). The residuals from a one-way ANOVA fit to the raw AUCs appear in the box-plot and histogram

in Figure 10A. Note that the histogram suggests right skew for the AUCs, and there are potential outliers in the box-plot, particularly on the upper end of the distribution.

Right skew with outliers (along with increasing variability with increasing mean) are telltale signs that a log transform may be needed. The residuals from a one-way ANOVA fit to log-transformed AUCs appear in Figure 10B. Note that the distribution is much more symmetric, and the box-plot does not identify any outliers on the log scale. The bell-shaped normal distribution (superimposed on the histogram) appears to more adequately model the log-transformed AUCs.

In summary:

- Outlier box-plots can be used to identify potential outliers. Other methods, such as number of standard deviations from the mean, are not recommended.
- An outlier can be removed from an analysis if it has an assignable cause or is clearly erroneous, but this should be documented.
- When there are outliers that do not have a known cause and are not clearly erroneous, analyze the data with and with the suspected outliers and report all results.
- Transforming the response variable can sometimes make outliers become non-outliers and satisfy standard analysis assumptions better than non-transformed data.
- Non-parametric analysis methods can be used to minimize the impact of outliers on analysis results without removing them from the analysis.

3.3. Randomization

There are a number of randomization techniques that are available. Certain study designs require specific randomization techniques, (e.g. randomized block designs, stratified designs, etc.). Random numbers should be obtained from an acceptable random number generator, (e.g., Excel, JMP, SAS, random number table, randomization web tool, etc.).

It is also very important to appropriately randomize subjects to treatment groups, as this reduces opportunities for bias. Randomization requires extra time and effort, but it can be more costly to not use it. Non-random strategies for assigning animals to treatment groups, when applied consistently across studies, will tend to introduce flaws, or bias, into the study results. By looking at the performance of a strategy across a series of studies we can then examine how well or poor a particular strategy is working. A truly random selection process is one where each subject is equally likely to be selected for any treatment group. The word “random” has a specific meaning that implies a formal random process such as a random number table or a computer generated random list has been used to make animal selections. **Using a computer-generated random list or method is the preferred method.**

Intentionally assigning subjects to different groups in order to balance one or more baseline variables is not an acceptable randomization method. While it might appear that the goals of randomization have been achieved, other unknown biases could be introduced.

- Randomization:
- Prevents implicit/explicit bias in treatment assignment
    - Approximately balances groups with respect to all extraneous variables affecting response
    - Does add to logistical complexity
    - Is the primary technique for assigning animals to groups
- Work with a statistician to
    - Design a workable randomization
    - Incorporate known covariate information

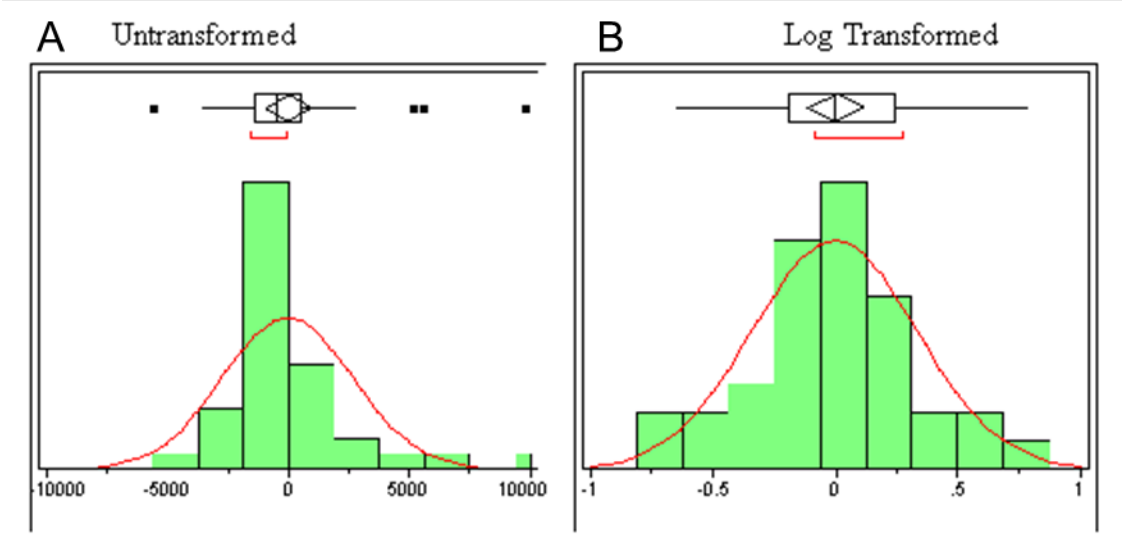


Figure 10: Results of a transformation eliminating outliers. (A) The residuals from a one-way ANOVA fit to the raw AUCs. (B) The residuals from a one-way ANOVA fit to log-transformed AUCs.

3.4. Power

The study should be planned so that the minimum pharmacologically relevant change or the flow scheme CSF has a high probability (usually 80%) of being statistically significant. The probability that a relevant change will be statistically significant is referred to as statistical power.

Note that since the total number of animals available for an assay run may be constrained by practical considerations such as a processing capacity, etc., the power analysis often determines how many dose groups may be examined within a single study.

The statistical power is a function of the following key elements.

- Assay variability (the lower the assay variability the higher the statistical power for a given sample size and effect size)
- Effect size (the larger the effect size the larger the statistical power for a given sample size and level of assay variability)
- Total number of animals in the test group
- Total number of groups and the number of animals per group in the study (in both cases the more animals the higher the statistical power).

Standard convention is to use 80% power to detect the minimum biologically significant effect with a false positive rate of 5% when evaluating or setting sample size. Declaring statistical significance when p-values are less than 0.05, along with appropriate multiplicity corrections, ensures the false positive rate is 5%. In general, a minimum of three runs of an assay with the format to be used in production should be used to estimate the experimental error. The power calculation should account for multiplicity corrections. The number of animals used should be sufficient so that all relevant drug effects are statistically significant.

When setting CSFs for an *in vivo* study, the CSF should be set above the minimum detectable difference if the CSF is defined on the response scale (e.g. percent inhibition > x% at y dose, body weight change > x grams, etc.). For potency CSFs, the study should be powered to exceed the minimum biologically significant effect. Following this paradigm, it is unlikely that you will have sufficient sample size to declare statistical significance at the

minimum biologically significant effect, but you should power the study to be able to detect statistical significance at some dose.

In a dose-response study to estimate an ED_{50}, two compounds will have statistically different ED_{50}s if the ratio of the ED_{50}s (larger::smaller) exceeds the minimum significant ratio (MSR) of the assay (see Section 2.2). Thus, the latter should be small enough to discriminate all pharmacologically relevant differences. The MSR depends upon the number of animals, the number of concentrations used, and the spacing of the doses with respect to the ED_{50}. The number of doses should either be large enough to estimate an ED_{50} over a reasonably wide range and/or adjusted for each compound based on the efficacy demonstrated in a single dose screen. If a large number of doses are used, the number of animals per dose may be quite small.

3.5. Analysis of Dose-Response Curves: Principles and Practice

In considering how to statistically evaluate dose-response curves, it is informative to review some of the historical context of how dose-response curves have been analyzed. When pharmacologists were first determining dose-response curves, in the very early days of pharmacology, they often plotted the data on an arithmetic scale (Figure 11A), that being the simplest and, perhaps, most familiar, to them.

In order to summarize and compare different dose-response curves, scientists attempted to describe these curves mathematically. However, the curves, a type of parabola, require rather complicated equations to describe. Prior to the invention of calculators and computers, it was too laborious and time-consuming to solve these equations. Scientists therefore searched for other ways of plotting and summarizing their data which might be readily described by equations. One way of doing this was to plot the dose-response curve on a semi-logarithmic plot, as shown for the same data in Figure 11B.

The semi-log plot of the data is the familiar sigmoidal (S-shaped) plot of many dose-response curves. This type of plot had the important advantage that **the middle part of the curve was approximately linear** (Figure 12). The portion of the curve from 16% efficacy to 84% efficacy can be described by a linear equation ($y = mx + b$) which can be solved without the aid of calculators or computers. However, data outside of this range (below 16% and above 84%) was off the linear portion of the curve and therefore was excluded from the analysis since the inclusion of such data would alter the slope of the line.

Pharmacologists wanted to have a single number to describe the dose-response plots of their data. For this purpose, the dose which produced a 50% effect (ED_{50}) seemed ideal since it was in the middle of the linear portion of the curve and could be calculated from the linear regression. These types of considerations led pharmacologists to design experiments which emphasized the middle, linear portion of the curve. Pharmacologists therefore primarily designed studies to have three groups (the minimum number of points needed to describe a line), and to have large numbers of animals in each group (to have a robust estimate of each mean).

A problem with this approach was that if the means of any groups fell outside of the 16 to 84% range, they were of little use in solving the linear equation. Many text books taught that data outside of the 16 – 84% range should be excluded, particularly the results of any groups where the mean was 0% or 100%. However, data at the extremes, or asymptotes, were very important data as they defined the top and bottom of dose-response curves. Moreover, it was difficult to calculate confidence limits on the slope and ED_{50}. Thus, using only linear regression and the linear portion of the curve had substantial limitations.

Since sigmoidal curves are nonlinear, a non-linear regression algorithm should be used to fit the data. Today, with computers to solve complex equations very rapidly, we can use non-linear curve fitting techniques to model, or mathematically describe sigmoidal-shaped dose-response curves. From the non-linear curve-fit, specific parameters are estimated which describe the dose-response curve. The parameter estimates can then be used to compare dose-response curves for different compounds.

Sigmoidal dose-response curves can be described using four different parameters:

- The top, or maximum effect
- The bottom, or minimum effect
- The ED_{50}, or mid-point on the curve
- The slope of the curve

Using four parameters to describe a non-linear curve is called a 4 parameter logistic model, as illustrated in Figure 13.

The 4-Parameter Logistic (4PL) Model generates a family of curves using the four parameters of top, bottom, middle and slope. With these parameters, and using nonlinear regression, we can describe most sigmoidal curves. On occasion, there may be a practical or theoretical reason to define what the bottom and/or the top of a curve will be; for example, it may be known from the experimental methods or pharmacological theory that the bottom will be 0% or the top will be 100%. In such cases, only three parameters may be needed to describe the data. In such a case, a 3-parameter logistic model may be used to describe the data. Three parameter models are used most commonly when the dose range is imperfect (too high or too low) with respect to the potency of the compound, and we do not have doses which yield data near the top or bottom.

One also could fix both the top and bottom at constant values (a 2-parameter logistic model). However this approach makes some strong assumptions and one should let the data estimate the top and bottom whenever possible.

3.5.1. Analyzing Dose-Response Data

One way of analyzing data from dose-response determinations is to use ANOVA with a Dunnett's test to determine if there is an overall effect of the drug, and which doses produced an effect was statistically different from the control group. However, this analysis doesn't allow one to estimate a dose producing a 50% effect nor estimate the minimal and maximal effects or the slope. The dose producing a 50% effect will typically lie between two of the doses tested, and therefore requires interpolation to estimate. Interpolation requires some type of regression analysis.

In certain situations, for example with a limited number of dose groups, linear regression may be the best approach. However, the linear regression approach has several limitations:

1. Interpolating between two doses does not use all the data and hence can be inefficient. If one uses more than two doses, a straight-line fit may not be appropriate and the result will be a distorted (biased) estimate of the ED_{50}.
2. It is not trivial to quantify the precision (i.e. calculate a standard error) of the estimate of the ED_{50} from linear regression. Thus, it is not trivial to determine if two ED_{50} values are statistically significantly different.
3. There is no ability to identify the minimum and maximum effect, nor the precision of the estimates, with linear regression.

By using a nonlinear regression model, one can get a better fit to **all** of the data, as well as calculate estimates and 95% confidence limits for the ED_{50}, slope, maximum effect and minimum effect.

In the process of assay validation, methods for analyzing dose-response data must be selected and the method should remain constant for a validated assay. However, there may be situations where deviations are appropriate. Consult with a statistician for how to most appropriately deal with specific situations, or datasets, where deviations from pre-determined data analysis approaches may be appropriate.

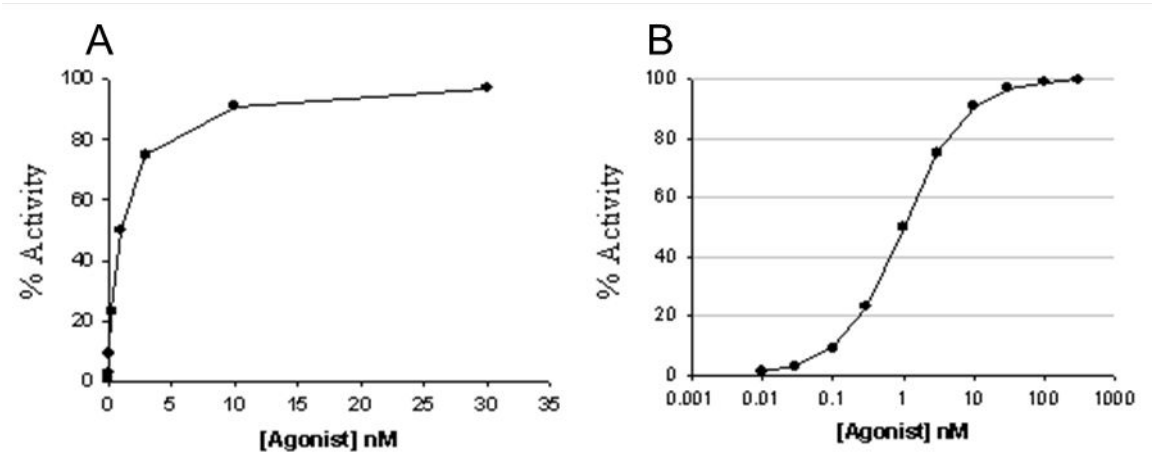


Figure 11: (A) Linear versus (B) semi-logarithmic plots of the same set of theoretical data.

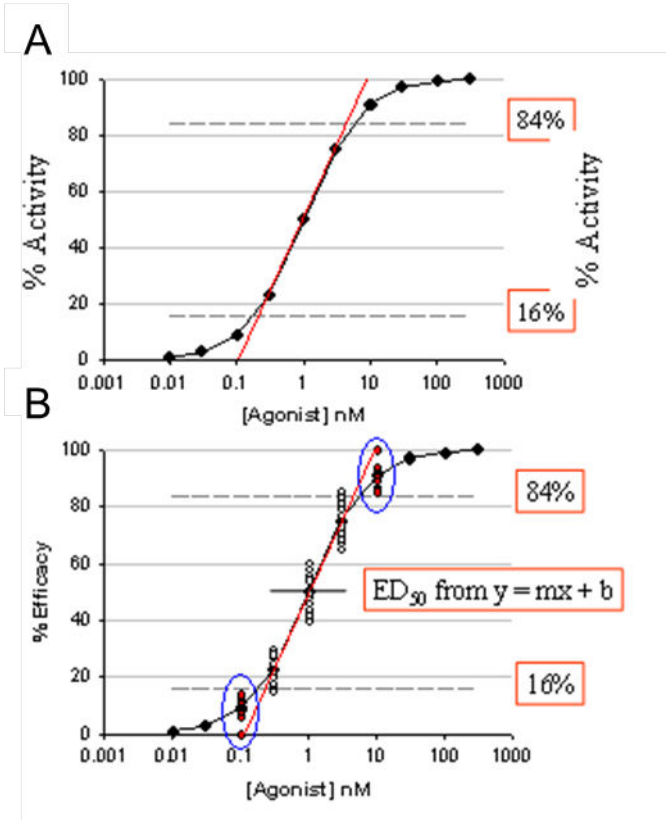


Figure 12: (A) Illustration of the linear portion of a sigmoidal dose-response curve and (B) the influence of data points which lie outside of the 16 – 84% region of the dose-response curve on the calculation of the slope of the line and potentially on the dose producing a 50% effect (ED_{50}).

3.5.2. Experimental design requirements for linear vs. nonlinear analysis

For analysis by linear regression or ANOVA followed by a Dunnett's test, 3 or 4 groups of usually 5 to 10 animals per group traditionally has been required to provide precise estimates of the mean and variation for each treatment group and thus, sufficient power to identify all important treatment effects. To estimate potency results, such as ED_{50}, the requirements are quite different; having data over the entire range of the dose-response curve is required, and estimates of the mean at each individual dose do not need to be as precise. That is, with

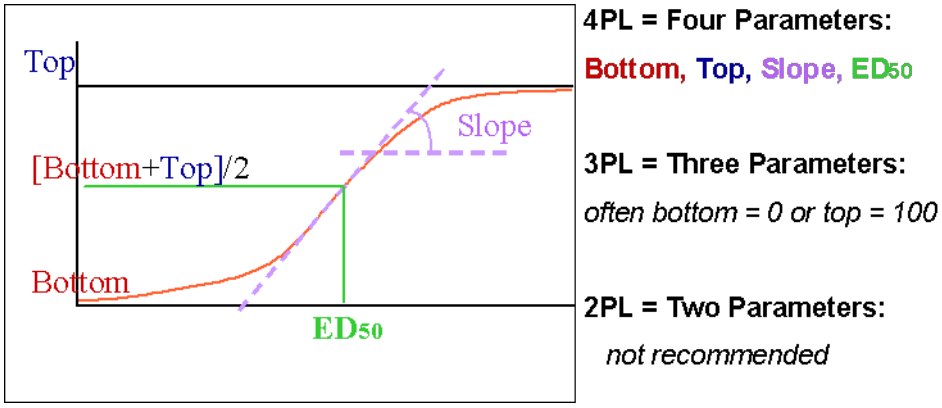


Figure 13: Illustration of the parameters of a 4-parameter logistic model of sigmoidal-shaped curves.

nonlinear regression, one is not estimating a given point (or mean), but rather the parameters of the entire curve; modest deviations in a single point are less likely to substantially impact the parameters of the overall curve. On the other hand, not having data points at, for example, the top or bottom, can substantially impact the analysis. Therefore, it is necessary to have a larger number of groups with fewer animals per group. With experimental designs for nonlinear analysis, it may even be possible to reduce the total number of animals needed. Diagnostic checking of the resulting curve fit is essential; the ED_{50} and asymptotes must make sense. Further, the confidence limits should be sufficiently precise to meet research objectives. Therefore, it is critically important to carefully consider the selection of doses in an experimental design with nonlinear regression analysis. For example, using 6 dose levels with 4 animals per group would be far more appropriate than using 3 dose levels with 8 animals per group. Also for nonlinear regression analysis, it is best to have doses that are equally-spaced on the log scale. *When an assay is statistically validated, the number of groups and animals per group needed should be verified with a statistician.*

As an illustration, the effects of a drug on decreasing ethanol consumption in a behavioral assay were compared under two different experimental designs. For these experiments, animals were trained to drink an ethanol solution over a period of several weeks. Animals were then randomized to treatment groups which receive either vehicle or a dose of the test drug. The drug produced a dose-related decrease in ethanol consumption (Figure 14). In the more traditional experimental design (Figure 14A), vehicle and 4 doses of drug were administered to groups of 6 animals each, for a total of 30 animals. When the data were analyzed by ANOVA and a Dunnett's test, doses of 3, 10 and 30 were significantly different from vehicle. There were only 2 doses on the linear portion of the dose-response curve (16 – 84%), an insufficient number to properly use linear regression, and therefore an ED_{50} and confidence limits could not be calculated with any degree of robustness. If nonlinear regression analysis is applied to the data in the left panel, an ED_{50} value of 0.6 mg/kg is obtained (a value which appears to be too low based on visual inspection of the data) and confidence limits can be obtained only for the parameter Top. The bottom could be fixed at zero, which provides a better estimate of the ED_{50}, but its lower confidence limit could not be calculated using JMP software. When a nonlinear-compatible design was used (Figure 14B), 6 doses of drug were administered to groups of 3 animals each for a total of 18 animals. A low, inactive dose of drug was used in place of vehicle, allowing for a better estimate of Bottom in this experiment. A vehicle group could also have been included. (Note: results from a vehicle treated group can, if appropriate, be incorporated into the nonlinear analysis; however, since there is no "zero" point on a log scale, vehicle is typically assigned a dose-value two to three orders of magnitude below the lowest dose tested; consult with a statistician for ways to incorporate a vehicle group into a nonlinear analysis). From nonlinear regression, an ED_{50} value of 1.5 mg/kg is obtained (which appears reasonable based on visual inspection of the data) as well as 95% confidence limits (0.91 to 2.4 mg/kg). In addition, estimates of the Top, Bottom and Slope, along with their respective confidence

limits, are obtained. Thus, if the primary goal is to obtain an ED_{50} value together with confidence limits, a nonlinear-compatible design yields more results with far greater precision, and may also require fewer animals.

3.5.3. Key points in the analysis of dose-response curves

Analysis Key Points:

- Regression is needed for interpolation
- 4 parameter logistic model is primary model for dose-response work
- Diagnostic checking is essential
  - Reasonable asymptotes
  - Numerically complete answer
  - Try fixing top or bottom as necessary

Design Key Points:

- Doses need to be spaced across a broad dose range
  - Doses may need to be in 2X steps rather than 3X (i.e., half-log)
- Preferable to use more groups but may need fewer animals per group
  - Nonlinear approaches can provide more information with fewer animals

4. Abbreviations

3PL, 3-parameter logistic

4PL, 4-parameter logistic

AMV, Assay method version

CSF, Critical success factor

DLs, Difference Limits

DRC, Dose-response curve

ED_{50}, Dose which produces 50% effect (Effective Dose 50%)

Relative ED50: Dose which produces 50% of the maximal response produced by that drug in the test system; considered to be relative measure of affinity

Absolute ED50: Dose which produces 50% of the maximal response which can be observed in the test system by a positive control; theoretically a drug could produce an effect greater than the maximum that can be measured

GM, Geometric Mean

LsA, Limits of Agreement

LsAd, Limits of Agreement on differences

MD, Mean Difference

MR, Mean Ratio

MSD, Minimum Significant Difference

MSR, Minimum Significant Ratio

RLs, Ratio Limits

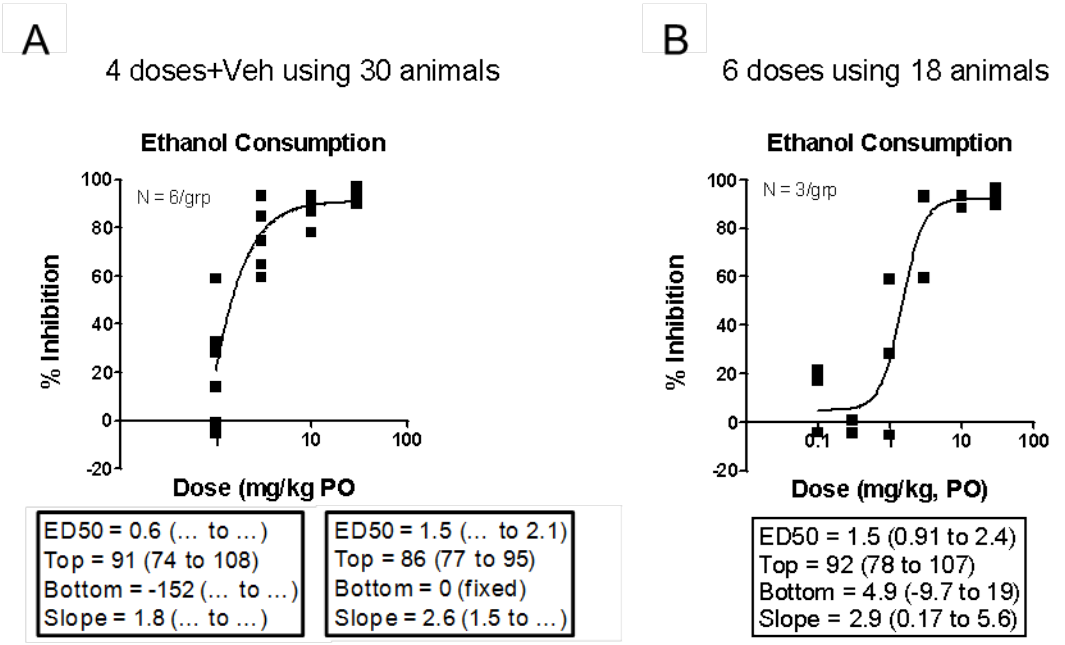


Figure 14: Comparison of (A) results with a more traditional experimental design using vehicle plus 4 doses of drug and 6 animals per dose and (B) results from a nonlinear-compatible design with 6 doses of drug and 3 animals per dose.

SAR, Structure-activity relationship

SDS, Single-dose screen

5. Suggested Reading

Landis SC, Amara SG, Asadullah K.et al. A call for transparent reporting to optimize the predictive value of preclinical research. Nature. 2012;490(7419):187–91. PubMed PMID: 23060188.

In Vivo Receptor Occupancy in Rodents by LC-MS/MS

Cynthia D. Jesudason, Ph.D,[1] Susan DuBois, BS,[1] Megan Johnson, BS,[1] Vanessa N. Barth, Ph.D,[1] and Anne B. Need, Ph.D[1]

Created: March 9, 2017.

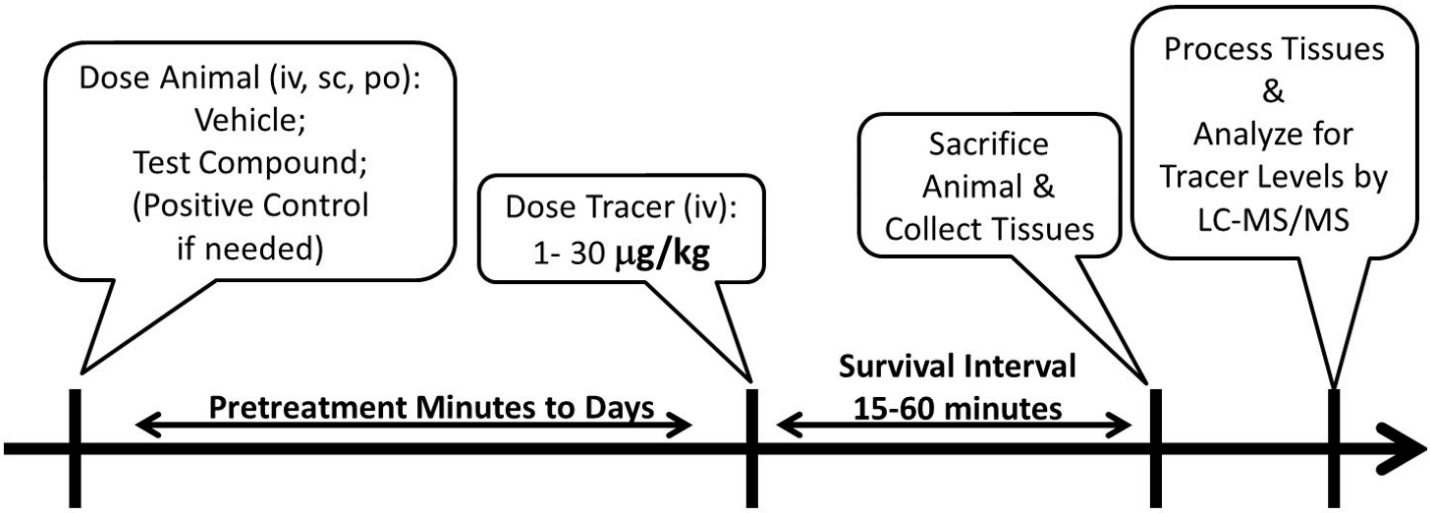
Abstract

Receptor (or enzyme) Occupancy (RO) is a quantitative measure of what percentage of the total number of available target receptors is engaged by or bound to a *particular* ligand. In drug discovery, it is important to determine the relationship of a receptor's occupancy to its final observed pharmacological efficacy and/or toxicological readout. RO is related to but distinct from a ligand's potency and maximal efficacy. Different compounds may achieve different percentages and durations of occupancy. Receptor occupancy/exposure relationships as well as the duration of occupancy have been used to project clinical doses.

The drug discovery lead optimization flow scheme can be improved by inclusion of a receptor occupancy assay that determines the level of target engagement to allow rational triage and selection of compounds to advance into more labor intensive and longer duration efficacy studies.

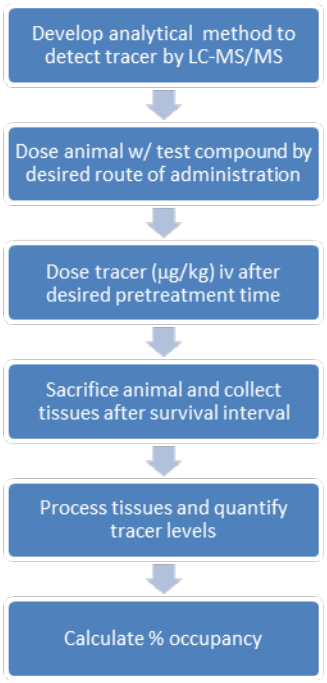
In the clinic, RO is measured using radioisotopically labeled tracer ligands whose distributions can be determined by non-invasive imaging methods, most commonly positron emission tomography (PET) and single photon emission computed tomography (SPECT) (1). Recent advances in the sensitivity of LC-MS/MS detectors have enabled determination of RO in preclinical studies without the need for a radiotracer (2, 3). This chapter will cover the details of performing an *in vivo* receptor occupancy assay using an LC-MS/MS methodology. This chapter assumes that a tracer exists for the target for which occupancy measures are being made, and will **not** detail the process of identifying a novel tracer (4, 5). A suitable LC-MS/MS tracer should be selective and have high affinity for the pharmacologic target, be readily tissue-penetrant, have low nonspecific binding, and exhibit suitable kinetics.

Experimental Design



Author Affiliation: 1 Eli Lilly & Company; Email: jesudasoncd@lilly.com; Email: dubois_susan_l@lilly.com; Email: watson_megan_a@lilly.com; Email: barth_vanessa_nicole@lilly.com; Email: need_anne_b@lilly.com.

Flowchart



Materials and Methods

Animals

Adult rats and mice are group housed in rooms with a 12-hour light/dark cycle. Room temperature is maintained at 21 ± 3°C. Food and water are available *ad libitum* until the beginning of the experimental protocol. Animals are allowed 3-5 days of acclimation to the housing environment prior to testing. In the most ideal case, the RO assay will be performed in the same strain of animals used for the efficacy/toxicology assay. The choice of animal strain will also be influenced by any differences in expression level of the receptor or enzyme between rats and mice (6). Pharmacology differences between rodents and humans due to differences in the target protein sequence can be overcome by utilizing "humanized" animals.

Test Compounds

Test compounds refer to the compounds for which receptor occupancy is to be assessed. These compounds can be administered through any desired route with any pretreatment time in a suitable vehicle. Animals may be fasted overnight if test compounds are to be administered by oral gavage. We note that the stress induced by oral gavage can be a confounding variable in many studies, so proper training in the procedure should be undertaken. In the most ideal case, the dose, vehicle and route of administration would match those used in other assays in the drug discovery flow scheme to build occupancy, efficacy/toxicology, and exposure relationships (7, 8).

Tracer

Tracer refers to the compound that is administered at low doses to characterize receptor occupancy at a specific target. The likelihood of finding a tracer is dependent on a number of factors as illustrated in Figure 1.

Compounds with low nonspecific binding will be more likely to exhibit a favorable differential distribution *in vivo* consistent with the biology. This differential distribution is directly related to the density of the target

(B_{max}) and the affinity of the compound to the target. The physiochemical properties of a good tracer molecule have been described in the literature based on analyses of existing PET ligands (9, 10).

The tracer is administered intravenously in the lateral tail vein in a dose volume of 0.5 mL/kg for rats and 5 mL/kg for mice following the desired pretreatment time of the test compound. It is typically administered in the range of 1-30 µg/kg. The survival interval between tracer administration and animal sacrifice is typically in the range of 15-60 minutes. The dose of the tracer and the survival interval will differ for each tracer, and will need to be optimized prior to setting up an RO assay to obtain the largest signal to noise window with the shortest survival interval. An example of a tracer optimization study is shown in Figure 2 in which the striatum expresses the target of interest while the cerebellum does not (i.e. null region), and is a measure of nonspecific binding of the tracer.

Analytical Method Development

The tracer of interest is diluted in 50:50 (acetonitrile:water) with 0.1% formic acid to 1 µg/kg. This is infused into the mass spectrometer using a 1 mL syringe (Hamilton Cat. No. 4023184) at 10 µL/min and the precursor and product ions are identified. The compounds are quantified after elution from the HPLC using a triple quad mass spectrometer with multiple reaction monitoring (MRM methods) and monitoring parent and product ion pairs in positive or negative mode as appropriate for the tracer. A sample LC trace and MS/MS for the tracer PBR28 is shown in Figure 3.

Live Phase

Mice or rats (n=4) receive either vehicle alone, test compound, or a positive control (PC) if no null region exists. After the desired pretreatment time the animals are administered the tracer. Following the survival interval they are euthanized by cervical dislocation followed by decapitation. Trunk blood is collected in EDTA coated 1.5 mL conical polypropylene centrifuge tubes and stored on wet ice until study completion.

Central Occupancy: Brains are removed and rinsed with sterile water. The regions of interest are dissected, placed in 1.5 mL conical polypropylene centrifuge tubes and stored on wet ice until completion of the live phase. Regions of interest should include a brain region with a high expression of the target of interest as well as a null region with little or no receptor density, if available.

Peripheral Occupancy: Peripheral tissues are dissected, rinsed with cold saline, placed in 1.5 mL conical polypropylene centrifuge tubes, minced to fine pieces using fine point scissors, weighed, and stored on wet ice until completion of live phase study.

Tissue Preparation and Analysis

Blood samples are collected at the time of sacrifice via trunk blood collection into 1.5 mL EDTA coated Eppendorf tubes and stored on wet ice until the study completion. Whole blood is centrifuged at 15,000 × g for 20 minutes to obtain plasma. Finally, 50 µL of plasma are added to Eppendorf tubes containing 200 µL of acetonitrile containing 0.1% (v/v) formic acid and vortexed to mix.

Tissue samples are collected at the time of sacrifice, rinsed with sterile water, stored in a 1.5 mL Eppendorf tube, weighed and stored on wet ice until the study completion. Stainless steel beads (NextAdvance SSB14B) are added to peripheral tissue and the tubes are placed in a Bullet Blender for 5 minutes. Acetonitrile containing 0.1% (v/v) formic acid is added to all tissues at a volume of four times the tissue sample weight in milligrams. All tissues are homogenized by using an ultrasonic dismembrator probe (Fisher Scientific Model 100, Pittsburgh, PA). Homogenized tissue and plasma are centrifuged at 15,000 × g for 20 minutes. An aliquot of supernatant containing the tracer is diluted in water to an acetonitrile content less than that of the mobile phase and 10 µL is injected by an autosampler onto an HPLC employing a suitable column, such as ZORBAX Eclipse XDB-C18

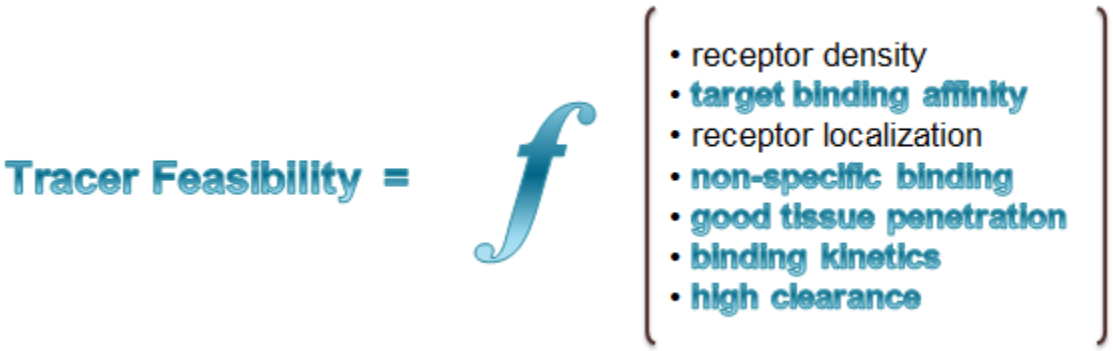


Figure 1: Tracer properties

column (part no. 971700-902, Agilent Technologies, Wilmington, DE). Separation is achieved using isocratic or gradient conditions and a mobile phase of water/acetonitrile containing 0.1% (v/v) formic acid. Individual standards for the tracer and plasma are prepared by adding known concentrations of tracer to naïve tissue or plasma and processed in the same manner as the samples. This process could potentially be automated to increase throughput but we are not aware of any group that has implemented this. A typical standard calibration curve is shown in Figure 4 (n=2, R2 >95%)

Determination of Occupancy

Receptor Occupancies can be calculated by using one of two different methods, the ratio method and the positive control method.

The Ratio Method

The ratio method is preferred when a null region is available in the same organ because this can correct for the potential variations in exposure of the tracer as a result of a number of variables including changes in blood flow (both increases or decreases), as well as changes in tracer efflux caused by pretreatment of the test compound. It is more difficult to identify nonspecific binding regions within a particular peripheral organ or another peripheral organ that truly represents nonspecific tracer binding. Therefore, with peripheral occupancy studies, 100% occupancy is pharmacologically defined and experimentally determined for each tracer within a single tissue using the positive control method described later.

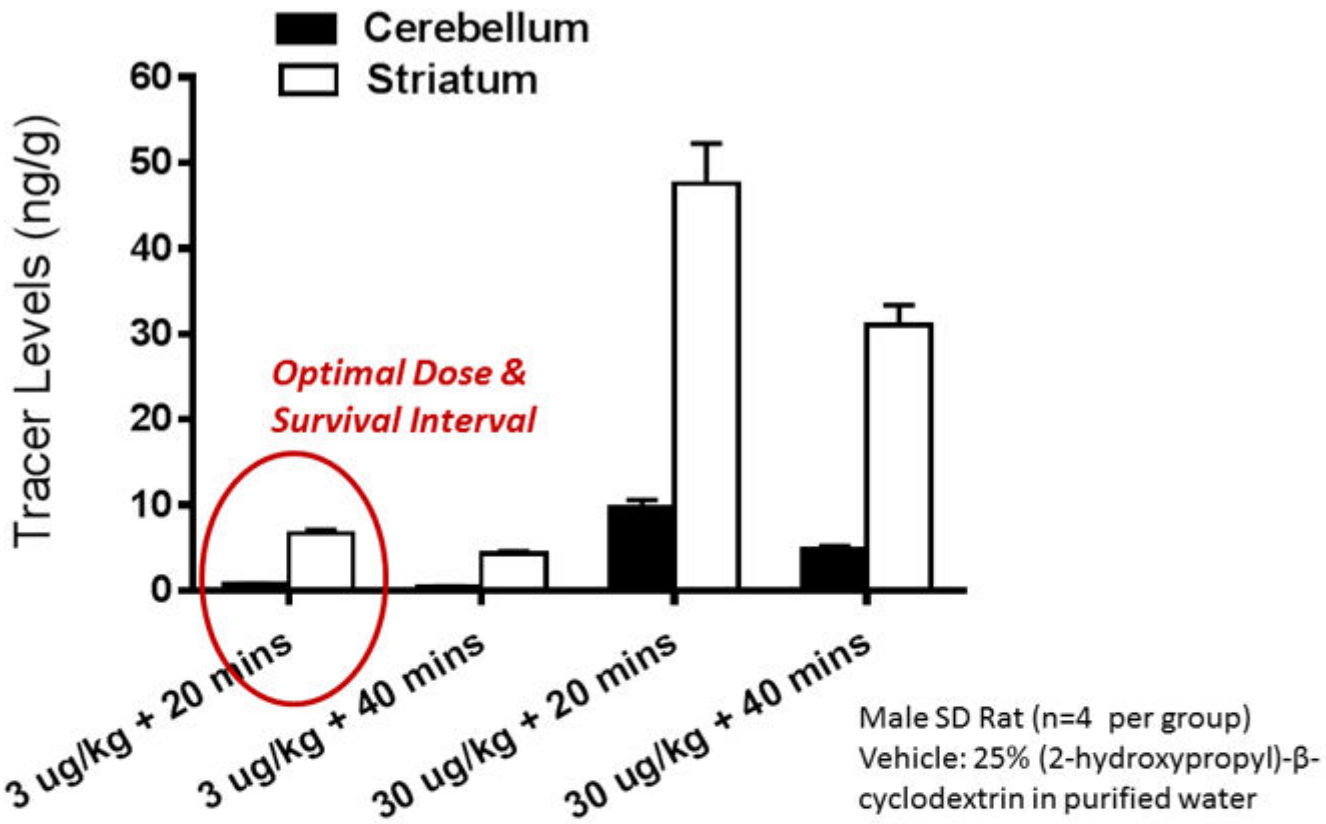
Receptor Occupancies are calculated using the well-described ratio method (7), but substituting the tracer concentrations determined by LC-MS/MS for the radiolabeled tracer levels determined with scintillation spectrometry.

The following equation is employed:

$$\%\ Occupancy = \left(1 - \left[\frac{(Ratio_t - 1)}{(Ratio_c - 1)}\right]\right) x\ 100$$

Each Ratio refers to the ratio of tracer in a brain area rich in target receptor to the tracer detected in a null region with little or no receptor density. $Ratio_t$ refers to animals treated with test compound, while $Ratio_c$ refers to the average ratio in vehicle-treated animals.

An example of a receptor occupancy dose response using this methodology is shown in Figure 5 in which the cerebellum is the null region, and the frontal cortex expresses the target.



Tracer Dose and Survival Interval

Tracer Dose, Survival Interval	BP	% SUV
3 ug/kg, 20 min	7	224
3 ug/kg, 40 min	10	146
30 ug/kg, 20 min	5	159
30 ug/kg, 40 min	6	104

Binding Potential (BP) = (Striatal levels/cerebellar levels) – 1
Standardized Uptake Value (SUV)= (Striatal tracer levels /injected tracer dose) x 100

Figure 2: Example Tracer Optimization Study

The Positive Control Method

This method is utilized when no null region is available for brain occupancy and for all peripheral occupancy assays. Average vehicle tracer levels in the target rich region represent 0% occupancy. Average tracer levels observed in this tissue from the positive control group represents 100% occupancy. The intermediate tracer levels from the test compound pretreated group are interpolated linearly between the tracer levels in the vehicle-treated animals (0% occupancy) and in the positive control animals (100% occupancy) in order to determine the percent occupancy. The positive control dose is identified as a dose that is well on the top asymptote of the receptor occupancy curve for this test compound or another previously tested compound known to bind to the target (11).

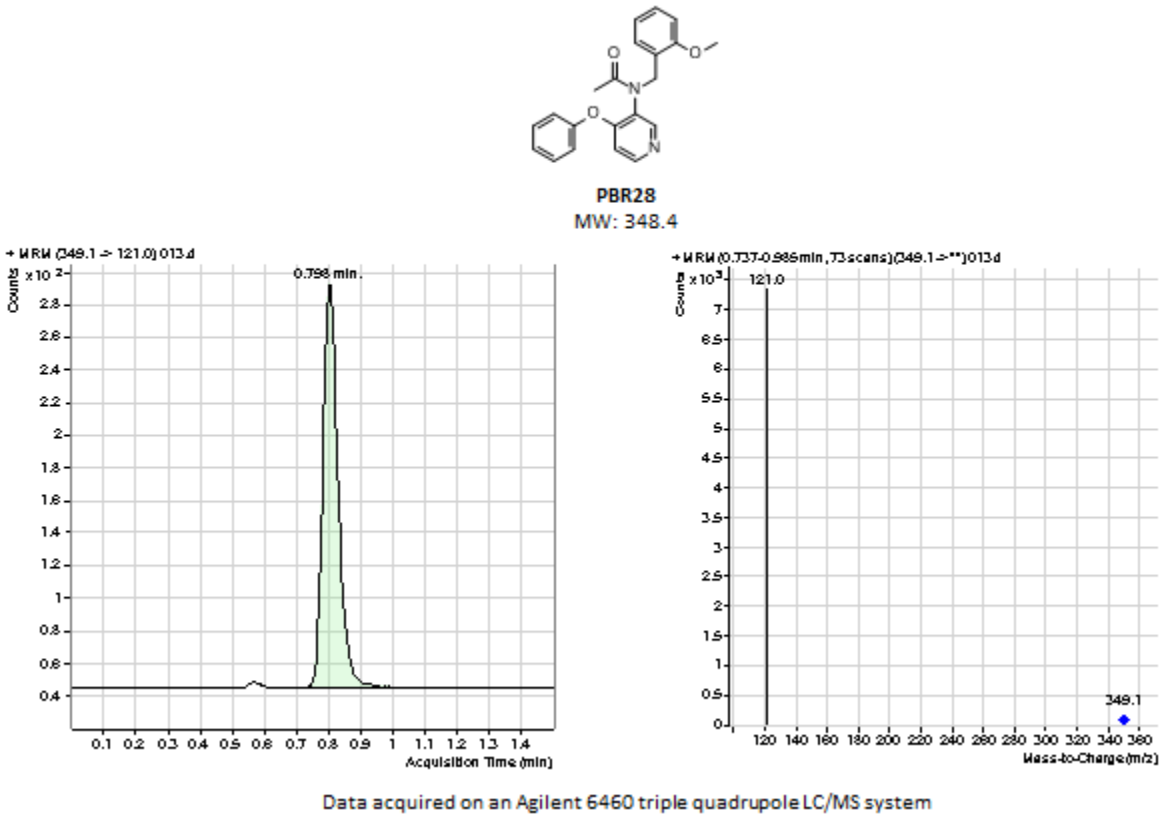


Figure 3: Example LC & MS/MS Trace

$$Intercept = Average\left(all\ vehicle\ \frac{ng\ (tracer)}{g\ (tissue)}\right)$$

$$Slope = \frac{Average\left(all\ positive\ control\ \frac{ng\ (tracer)}{g\ (tissue)}\right) - Average\left(all\ vehicle\ \frac{ng\ (tracer)}{g\ (tissue)}\right)}{100}$$

$$\%\ Occupancy = \frac{\left(\frac{ng\ (tracer)}{g\ (tissue)}\right) - intercept}{slope}$$

This method is dependent on having access to a suitable positive control compound that demonstrates 100% occupancy at the desired target, and this can often be a challenge for new targets where there are no known reference compounds available and limited chemical SAR information.

An example of a receptor occupancy dose response using this methodology is shown in Figure 6 in which both the frontal cortex and the cerebellum express the target and show dose-dependent reduction in measured tracer levels.

The duration of occupancy is also an important feature of clinical candidates, and an example of a time course peripheral occupancy study using the positive control method is shown in Figure 7.

Graphical Analysis

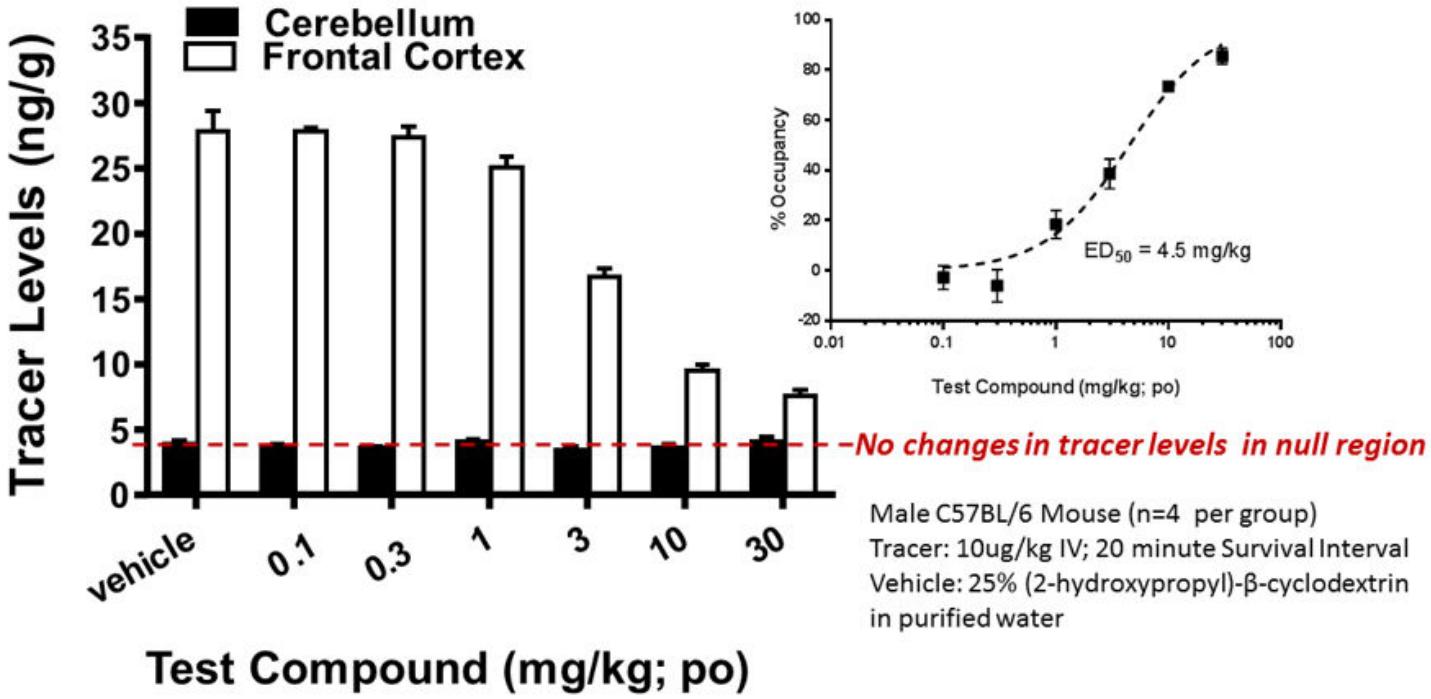
Prism (GraphPad Software Inc., San Diego, CA) software can be employed for calculations, curve fitting and graphics (Figures 5-7).

LC-MS/MS Calibration Curve for Tracer in Tissue		
Matrix =	Brain Tissue (ng/g)	
Levels	1st Run	2nd Run
0.3	0.304	0.292
1	1.08	0.985
3	3.01	2.78
10	10.3	10.2
30	31.4	29.5
60	58.7	57.5
R2 = 0.9978		

LC-MS/MS Calibration Curve for Tracer in Tissue		
Matrix =	Lung Tissue (ng/g)	
Levels	1st Run	2nd Run
0.3	0.290	0.305
1	0.962	1.09
3	2.93	3.04
10	10.3	9.96
30	29.5	28.9
R2 = 0.9976		

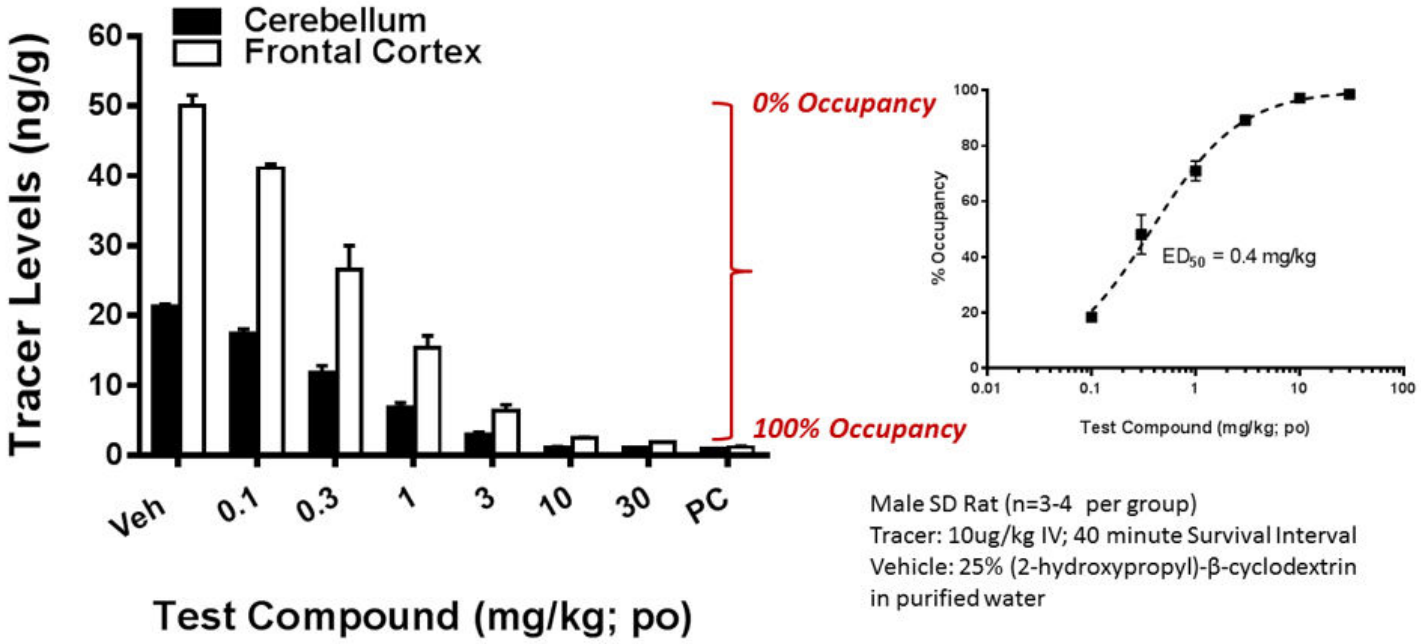
LC-MS/MS Calibration Curve for Tracer in Plasma		
Matrix =	Plasma (ng/mL)	
Levels	1st Run	2nd Run
0.3	0.322	0.275
1	0.987	1.02
3	3.17	3.04
10	10.0	9.82
30	29.3	29.0
R2 = 0.9970		

Figure 4: Example Calibration Curve in Tissues & Plasma



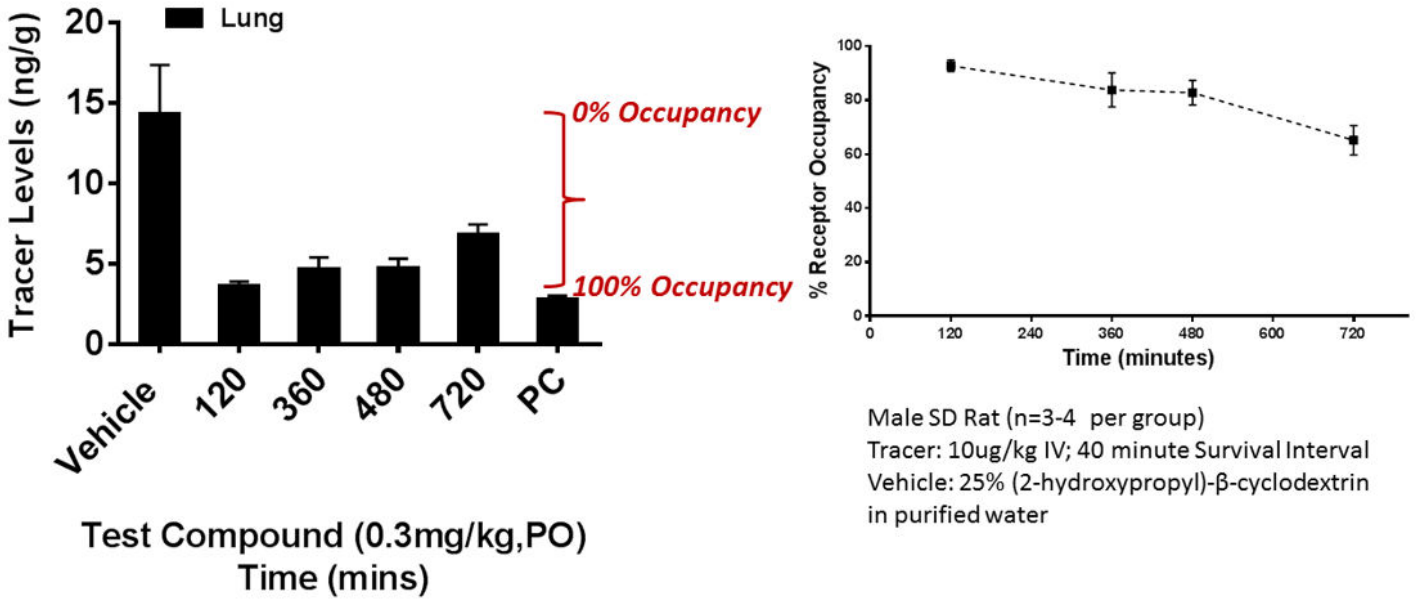
Test Compound (mg/kg) po	% Occupancy	SEM
0.1	-2.83	4.74
0.3	-6.10	6.43
1	18.4	5.67
3	38.6	5.93
10	73.6	1.51
30	85.6	3.24
Abs ED50	4.5 mg/kg po	

Figure 5: Receptor Occupancy Dose Response Curve Using the Ratio Method. Occupancy is summarized by treatment group mean % occupancy ± standard error of the mean. All dose occupancy curves were plotted using the software program GraphPad Prism. Absolute ED_{50} was estimated from the sigmoidal dose-occupancy curve analysis using the proper constraints.



Test Compound (mg/kg) po	% Occupancy	SEM
0.1	18.3	1.1
0.3	48.1	7.1
1	71	3.6
3	89.3	1.7
10	97.3	0.2
30	98.7	0.1
Abs ED50	0.4 mg/kg po	

Figure 6: Receptor Occupancy Dose Response Curve Using the Positive Control Method. Occupancy is summarized by treatment group mean % occupancy ± standard error of the mean. All dose occupancy curves were plotted using the software program GraphPad Prism. Absolute ED_{50} was estimated from the sigmoidal dose-occupancy curve analysis using the proper constraints.



Pretreatment Time	% Occupancy	SEM
120 min	92.7	2.1
360 min	83.8	6.3
480 min	82.8	4.6
720 min	65.1	5.5

Figure 7: Time Course Occupancy Using the Positive Control Method. Occupancy is summarized by treatment group mean % occupancy ± standard error of the mean.

References

1. George N., Gean E. G., Nandi A., Frolov B., Zaidi E., Lee H., Brasic J. R., Wong D. F. Advances in CNS Imaging Agents: Focus on PET and SPECT Tracers in Experimental and Clinical Use. CNS Drugs. 2015;29(4):313–330. PubMed PMID: 25948171.
2. Need A. B., McKinzie J. H., Mitch C. H., Statnick M. A., Phebus L. A. *In vivo* rat brain opioid receptor binding of LY255582 assessed with a novel method using LC/MS/MS and the administration of three tracers simultaneously. Life Sciences. 2007;81(17-18):1389–1396. PubMed PMID: 17935738.
3. Chernet E., Martin L. J., Li D., Need A. B., Barth V. N., Rash K. S., Phebus L. A. Use of LC/MS to assess brain tracer distribution in preclinical *in vivo* receptor occupancy studies: Dopamine D2, serotonin 2A and NK-1 receptors as examples. Life Sciences. 2005;78(4):340–346. PubMed PMID: 16139310.
4. Joshi E. M., Need A., Schaus J., Chen Z., Benesh D., Mitch C., Morton S., Raub T. J., Phebus L., Barth V. Efficiency Gains in Tracer Identification for Nuclear Imaging: Can *In vivo* LC-MS/MS Evaluation of Small Molecules Screen for Successful PET Tracers? ACS Chemical Neuroscience. 2014;5(12):1154–1163. PubMed PMID: 25247893.
5. Need A, Kant N, Jesudason C, Barth V. Approaches for the discovery of novel positron emission tomography radiotracers for brain imaging. Clinical and Translational Imaging. 2017.:1–10.doi: 10.1007/s40336-017-0221-3.
6. Pope G. R., Roberts E. M., Lolait S. J., O'Carroll A-M. Central and peripheral apelin receptor distribution in the mouse: Species differences with rat. Peptides. 2012;33(1):139–148. PubMed PMID: 22197493.
7. Wadenberg, M-L. G.; Kapur, S.; Soliman, A.; Jones, C.; Vaccarino, F.; Dopamine D2 receptor occupancy predicts catalepsy and the suppression of conditioned avoidance response behavior in rats. Psychopharmacology (Berlin) (2000), 150(4), 422-429. PubMed PMID: 10958084.

8. Barth V., Need A. B., Tzavara E. T., Giros B., Overshiner C., Gleason S. D., Wade M., Johansson A. M., Perry K., Nomikos G. G., Witkin J. M. In vivo occupancy of dopamine D3 receptors by antagonists produces neurochemical and behavioral effects of potential relevance to attention-deficit-hyperactivity disorder. Journal of Pharmacology and Experimental Therapeutics. 2013;344(2):501–510. PubMed PMID: 23197772.
9. Zhang L., Villalobos A., Beck E., Bocan T., Chappie T., Chen L., Grimwood S., Heck S., Helal C., Hou X., Humphrey J., Hu J., Skaddan M., McCarthy T., Verhoest P., Wager T., Zasadny K. Design and selection parameters to accelerate the discovery of novel central nervous system positron emission tomography (PET) ligands and their application in the development of a novel phosphodiesterase 2A PET ligand. Journal of Medicinal Chemistry. 2014;56(11):4568–4579. PubMed PMID: 23651455.
10. Barth V, Need A. Identifying Novel Radiotracers for PET Imaging of the Brain: Application of LC-MS/MS to Tracer Identification. ACS Chemical Neuroscience. 2014;5(12):1148–1153. PubMed PMID: 24828747.
11. Need A. B., Davis R.J., Alexander-Chacko J. T., Eastwood B., Chernet E., Phebus L. A., Sindelar D. K., Nomikos G. G. The relationship of *in vivo* central CB1 receptor occupancy to changes in cortical monoamine release and feeding elicited by CB1 receptor antagonists in rats. Psychopharmacology. 2006;184(1):26–35. PubMed PMID: 16328376.

Assay Artifacts and Interferences

Nathan P. Coussens, PhD and Jayme L. Dahlin, PhD

Interference and Artifacts in High-content Screening

Jayme L. Dahlin,[1] David A. Close,[2,3] Paul A. Johnston,[2,3] Steven A. Haney,[4] Steven A Titus,[5] Bridget K. Wagner,[1] and O. Joseph Trask[6]

Created: May 28, 2025.

Abstract

Since the late 1990s, high-content screening (HCS) approaches have contributed greatly to cell and chemical biology, drug discovery, and toxicology research. Despite the numerous advantages of multi-parameter interrogation of cell models using HCS approaches, these assay formats are also susceptible to artifacts and interference. This chapter describes 1) autofluorescence interference contributed by media, cells, and tissues which may complicate or preclude HCS assay development, 2) interference by environmental or microorganism contaminants, and 3) compound-mediated autofluorescence, fluorescence quenching, and cellular injury/cytotoxicity which may obscure whether compounds modulate the desired target or cellular phenotype or may produce false-positives or -negatives. It provides best-practice recommendations for identifying and mitigating such interference through the application of experimental design strategies, statistics, orthogonal, and counter-screens, and the use of reference compounds. Adoption of these recommendations should enhance the quality of chemical matter identified by HCS assays.

Abbreviations

DHT: dihydrotestosterone; ECM: extracellular matrix; FAD: flavin adenine dinucleotide; FRET: Förster/fluorescence resonance energy transfer; GFP: green fluorescent protein; HCS: high-content screening; HTS: high-throughput screening; IAF: image-based autofocus; LAF: laser-based autofocus; MOA: mechanism of action; NADH: nicotinamide adenine dinucleotide; RFP: red fluorescent protein; ROS: reactive oxygen species; SAR: structure-activity relationship; YFP: yellow fluorescent protein

Introduction

The selection of cellular assay format, experimental design, and components of the follow-up testing paradigm are all critical decisions that influence the ultimate success of a small-molecule high-throughput screening (HTS) campaign (1-5). Deploying a target-based or phenotypic screening approach requires the development, optimization, and validation of an assay that is biologically appropriate with a robust and reproducible signal window that has sufficient throughput and capacity to screen the compound library of interest (3). All HTS assays are subject to artifacts and interference from compounds and other sources, and high-content screening (HCS) assays are no exception (Figure 1) (1-3, 6-11). HCS assays detect perturbations in cellular targets, function, and phenotypic readouts, regardless of whether they arise from desirable or undesirable mechanisms. Additionally, since HCS imaging requires the transmission and reflectance of light for fluorescent signal generation and detection, optically active substances can alter readouts independent of a true biological effect.

HCS assays, like all cell-based assay approaches, are susceptible to interference from endogenous substances in cells, tissues, and culture media, as well as with the reagents and probes used for detection. When unaccounted for, the contributions of endogenous substances can preclude, adversely impact, or complicate the development

Author Affiliations: 1 Chemical Biology and Therapeutics Science, Broad Institute, Cambridge, MA, USA. 2 Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA, USA. 3 University of Pittsburgh Hillman Cancer Center, Pittsburgh, PA, USA. 4 Valo Health, Lexington, MA, USA. 5 Thermo FisherScientific, Eugene, OR. USA. 6 Revvity, Inc, Waltham, MA, USA; Email: highcontent@gmail.com.

Corresponding author.

of HCS assays, either by elevating fluorescent background levels sufficiently to make it challenging to detect bioactive responses, or by depressing or quenching fluorescent signals to an extent where they may be indistinguishable from background signal. Similarly, exogenous contaminants from the environment including lint, dust, plastic fragments from labware, pipette filter and lab coat fibers, and microorganisms may cause image-based aberrations such as focus blur and image saturation that can complicate downstream image analysis and impair the identification of subtle phenotypes (11). Exogenous contaminants may also perturb the underlying biology.

However, the major source of artifacts and interference in HCS assays are the test compounds themselves. Compound-dependent assay interference can be broadly divided into fluorescence detection technology-related and non-technology-related cell cytotoxicity or morphology issues, although there is considerable overlap between these two categories (Figure 1) (7).

Compounds that do not modulate the biological target or phenotype of interest but instead interfere with assay detection because of autofluorescence or fluorescence quenching may produce artifactual bioactivity readouts and/or phenotypes in HCS assays, or mask bioactivity depending on the assay construction (12-14). Similarly, compounds that alter light transmission or reflection by a variety of mechanisms, including quenchers, insoluble compounds, and colored (pigmented) compounds may also interfere with HCS assays (Figure 1).

Autofocusing methods employed by HCS instrumentation can affect the acquisition and identification of cell images, as well as impact data quality during image acquisition, segmentation, and post-processing steps. This topic is described in detail in later sections.

Compound interference or artifacts due to autofluorescence or quenching can often be identified by statistical analysis of the fluorescence intensity data because the values produced by these compounds will typically be outliers relative to the normal distribution ranges of these measurements in control wells not exposed to compounds and in wells treated with compounds that are optically transparent or inert (13-16). Compound interference flagged by statistical methods can then be confirmed by 1) manually reviewing the images, 2) implementation of orthogonal assays that utilize a fundamentally different assay detection technology, and/or 3) counter-screens for assay interferences or target selectivity.

Non-technology-related compound interference occurs when compounds alter or modulate the biology without directly interfering with the HCS fluorescence detection technology. For morphological profiling, these effects could be considered desirable, as they produce quantifiable effects on cells, but there are many examples where unique phenotypes result from undesirable and artifactual properties (17). The effects of such compounds are most frequently manifested as cellular injury/loss ("cytotoxicity") or dramatic changes in cell morphology including cell shape, spreading and/or attachment (Figures 2, 3). Observations of cytotoxic dead cells may include loss of signal but more often as dead cells round-up, the fluorescence from some probes, e.g., nucleic acid probe concentrates fluorescence that may saturate camera or detector dynamic range directly comprising image-based autofocusing methods. Laser-based autofocusing methods are impacted by cytotoxic dead cell fluorescence. Selecting an optimal cell seeding density is one of the critical components of the assay development and optimization process for cell based HCS imaging assays (3,14,18,19). During HCS assay development, there are many variables that can impact the assay performance, including cell seeding density, media, poly-D-lysine (PDL), and extracellular matrix (ECM) or other microplate coatings, the objective magnification (4x, 10x, 20x, 40x, etc.), and the number of fields of view captured per well will determine the total number of cells that will be included in the image analysis. For more details on HCS assay development, please see Assay Development Guidelines for Image-Based High Content Screening, High Content Analysis and High Content Imaging (20). Compound-mediated cytotoxicity may obscure small-molecule activity at the target of interest or can be scored as false positives or negatives depending upon whether the assay is designed to identify inhibitors or activators (1,2,4-8,10). Similarly, compounds that substantially reduce or disrupt the adhesion of cells to assay plate surfaces can also produce significant cell loss which may invalidate the image analysis algorithm (3,13,17,18).

For HCS assays, the robustness and reproducibility of the multiparameter data extracted by the image analysis algorithm typically depends upon the number of cells captured and analyzed. In most HCS assays there is a threshold number of cells below which the coefficients of variation (CVs) of the multiparameter data will increase dramatically and the corresponding Z-factor coefficients for the assay signal window separation between maximum and minimum control populations declines precipitously (3,13,17,18). An adaptive image acquisition process where multiple fields of view are acquired until a preset threshold number of cells have been imaged is one strategy that can mitigate the impact of compound mediated cell loss. However adaptive image acquisition may prolong the image acquisition process considerably to a point where it becomes prohibitive, especially if the screening library contains a high percentage of compounds that are cytotoxic or induce extensive cell loss due to altered adhesion properties. Additionally, adaptive image acquisition may not prove to be effective in wells where the compound mediated cell loss is substantial. Substantial reductions in the number of cells captured and analyzed may reduce or impair the statistical significance of the image analysis parameters being measured, and dramatic alterations to cellular morphology may eliminate or disrupt the ability of the image analysis algorithm to accurately segregate objects and/or regions of interest. Substantial loss of cells due to compound mediated cytotoxicity or loss of adherence can readily be identified by statistical analysis of the nuclear counts and nuclear stain fluorescence intensity data because the values of such compounds will be outliers relative to the others.

Examples of undesirable compound mediated mechanisms of action / mode of action (MOAs) include nonspecific chemical reactivity, colloidal aggregation, redox-cycling, chelation, or denaturation mediated by surfactants or detergents (1,2,4,5,7,8,10). Other examples related to specific organelles or molecular pathways are lysosomotropic agents (cationic amphiphilic drugs, 'CADs'), cytoskeletal toxins (tubulin poisons), mitochondrial toxins (electron transport chain poisons), and genotoxins (DNA intercalators, alkylating agents) (7). Of course, for the discovery of mechanism-based antineoplastics or antimicrobials, cellular injury may be the appropriate and intended phenotypic readout. Furthermore, reaction- or chelation-based MOAs with a high degree of specificity could represent viable chemical probes/drugs for some screens. Biological targets and systems can exhibit enhanced sensitivities to certain undesirable MOAs. For example, the actives identified in screening assays for targets with catalytic cysteine residues or metal cofactors may be enriched for electrophilic and chelating agents, respectively. In cellular models with redox-sensitive pathways, screening active lists may be enriched for oxidants or redox-active compounds if the pathway is sufficiently dominant in terms of phenotypic contribution (1). For any HCS imaging (or HTS) campaign, a testing paradigm of appropriate counter screens and orthogonal assays can be deployed to ascertain whether active compounds can be confirmed as hits or whether they should be eliminated or "flagged" from further consideration because they are cytotoxic and/or alter cell morphology.

This chapter describes compound-mediated interference that impacts the imaging detection approaches and biological/cellular integrity of HCS assays. Guidelines and strategies to identify and mitigate the impact of compound interference in HCS assays are presented together with tips for experimental design, helpful statistical methods for flagging potential interference, the use of reference interference compounds, and sample orthogonal and counter-screen protocols.

This chapter is organized into sections by the HCS imaging workflow process:

- Cells and media
- Seeding and growing cells
- Labels and probe choices
- Image acquisition considerations
- Analysis
- Observations and insights from compound treatments
- Confirmation and follow-up

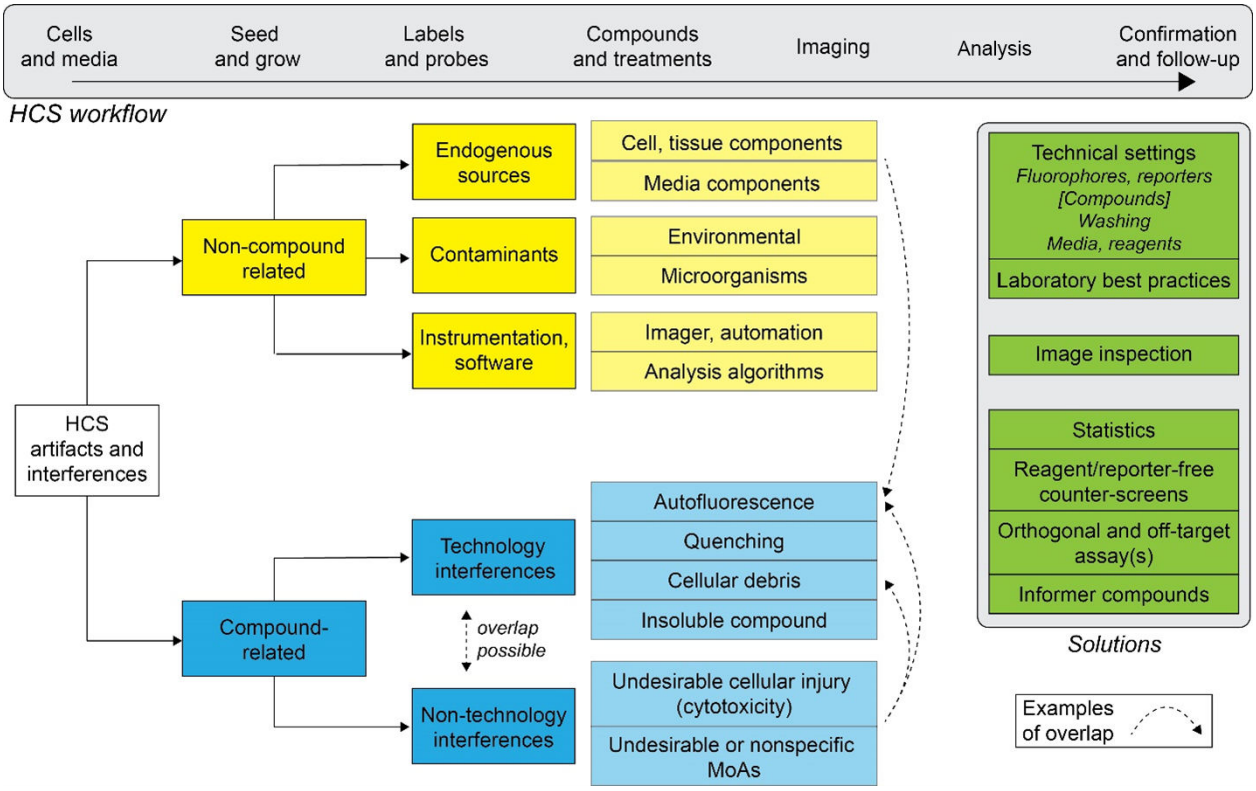


Figure 1. Summary of HCS artifacts and interference. Interferences can be broadly divided into compound-related and non-compound-related categories. Non-compound-related interferences include endogenous materials and contaminants. Compound-dependent interferences are further divided into technology-related and non-technology-related sources. Technology interference produces artifactual phenotypes, while non-technology interferences produce "real" phenotypes but through undesired (usually nonspecific) mechanisms of action.

Cells and media

There are several potential cell model sources of HCS assay interference that are independent of compound exposure. A variety of endogenous substances in culture media, cells or tissues, probes, and vessels or microplates can interfere with HCS fluorescent readouts that will need to be addressed during assay development (Figure 4). This section reviews examples of these endogenous sources of interference and recommends best practices to mitigate their impact.

Media components. Although tissue culture media is generally not a major source of artifacts, some components of culture media have fluorescent properties that can potentially interfere with image acquisition, analysis, and data interpretation (Figure 4A). Riboflavins (co-factors for flavoprotein enzyme reactions) fluoresce in the ultraviolet through green fluorescent protein (GFP) variant spectral ranges (ex. 375-500 nm and em. 500-650 nm) and within these ranges may elevate the fluorescent backgrounds acquired in live-cell imaging applications (Figure 4B, 4C). Typically, fluorescence interference contributed by media components is only a problem when the HCS assay signal of interest is relatively weak and falls within that spectral range. Media with lower riboflavin concentrations (e.g., F-12K, EBM) can be an alternative if the biology is not substantially perturbed (Figure 4D) (21). For short-term live-cell imaging, solutions lacking riboflavin such as HEPES-buffered HBSS or other commercially available products may be substituted. Protein supplements in media to maintain cell health and growth from calf, bovine, or other animal-derived serum contain rich mixtures of albumin, amino acids, and other growth factors and components that may increase background in fluorescent imaging experiments. Reducing the concentration of sera used in media can circumvent the risk of background fluorescence exposure.

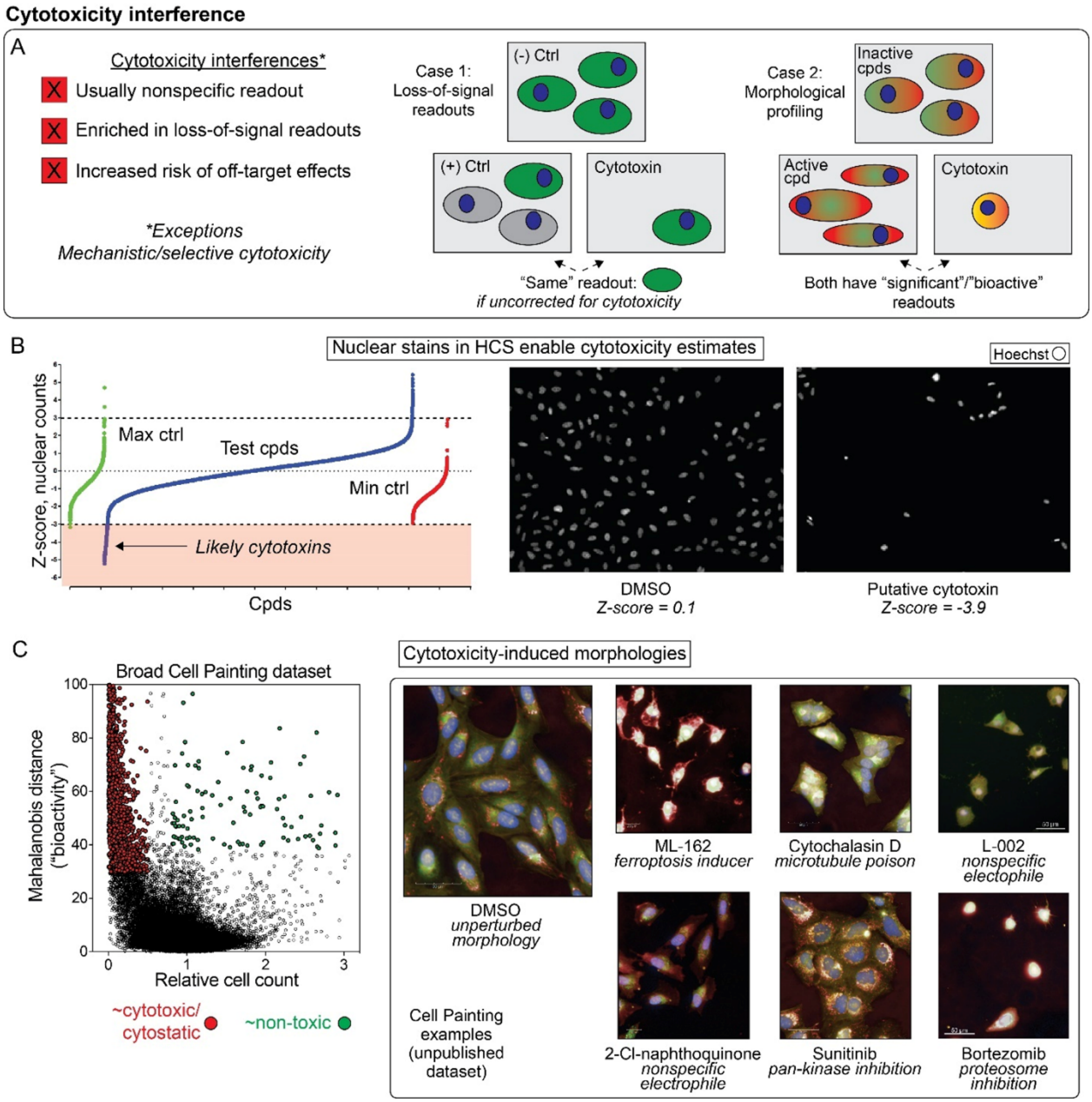


Figure 2. Cytotoxicity compound interferences in HCS assays. (A) Schematic of interference in HCS assays by cytotoxic compounds. (B) Example of cytotoxicity encountered in an AR-RFP/TIF2-GFP biosensor HCS assay. Note that cells can be identified by including a nuclear stain (Hoechst) to estimate cytotoxicity in the primary HCS assay. More details can be found in the appendix (13). (C) In profiling assays, cytotoxins can produce "significant" morphologies. The utility and specificity of these morphologies depends on many factors. Note that dying cells display a characteristic shrunken, round morphology with intense fluorescent staining in all channels. Similar morphologies are often encountered across a variety of cell lines and assay protocols. Panel C, left: adapted with permission (7).

Cells and tissues with high fluorescent background. Similarly, intrinsic autofluorescence of biological substances contained in normal and diseased tissues can sometimes interfere with fluorescence microscopy analysis of these samples. Cell types that are rich in pigment (e.g., retinal, melanocytes, etc.) can exhibit a high degree of light absorbance, light scatter, and/or autofluorescence that may confound HCS imaging approaches. Autofluorescence is often more apparent in highly metabolic cells such as hepatocytes, adipocytes, and connective tissues (22,23). Primary isolated hepatocytes are highly metabolic and intrinsically exhibit autofluorescence, leading to an increased fluorescent background in the blue-green part of the visible spectrum. This is particularly evident with excitation light sources between 360-488 nm and emission in the 500-550 nm

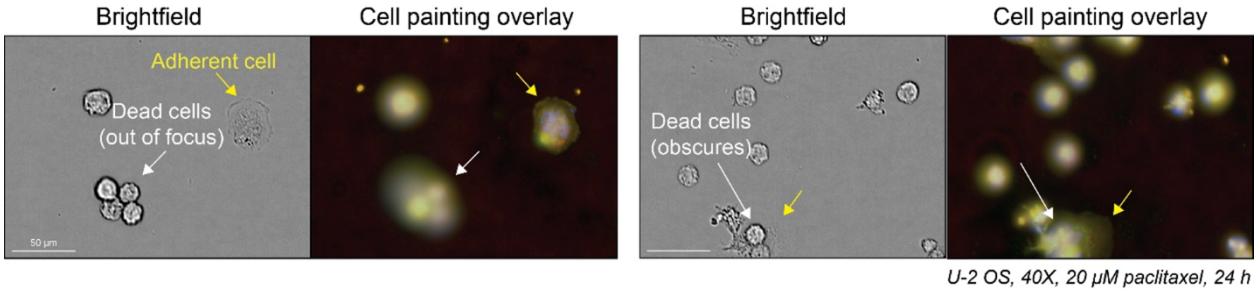


Figure 3. Example of dead cell interference in HCS. Non-viable cells are typically round and on a higher Z-plane relative to most adherent cells. They are often out of focus, and consequently, they can produce "significant" morphologies. They can interfere with imaging nearby cells. Including washing steps prior to fixation can remove many of these weakly adherent cells.

range from metabolic redox cofactors NADH (nicotinamide adenine dinucleotide) and FAD (flavin adenine dinucleotide). Both can mask the detection of low-intensity fluorescent probe signals within these spectral ranges (24,25). The impact of cellular autofluorescence will depend on its relative intensity and distribution compared to the probe signal intensities and patterns being measured in the HCS assay. Cells/tissues with high autofluorescence coupled with relatively weak object staining intensities (e.g., low protein expression) may substantially reduce the effective signal window and sensitivity of the assay. To determine the extent and potential impact of endogenous autofluorescence in a particular sample, it is recommended that background images of "unstained and untreated" control samples without any reporters, probes, or dyes, together with secondary detection antibody-only controls, be acquired in the relevant HCS channels during assay development. Again, the degree of interference contributed by cell and tissue autofluorescence only constitutes a problem for HCS assay development when the signal of interest is relatively weak and falls within the same spectral range.

Cellular autofluorescence can be altered by cellular metabolism and perturbations that change the concentrations of intrinsically fluorescent molecules such as nicotinamide adenine dinucleotide (NADH), or that modify cellular proteins and metabolites such as lipofuscins (24, 25). Intrinsic fluorescent cellular molecules often exhibit heterogeneous expression levels and generally have broad excitation and emission spectra (Figure 4A). The amount of endogenous autofluorescence in cells is dependent upon cell type, their metabolic state, and the nature of the culture model. The measurement of cellular autofluorescence can sometimes be exploited to diagnose diseased tissues (26-30), and changes in autofluorescence as a function of disease state or treatment can provide "label free" opportunities to quantify treatment-induced effects. For example, dual monitoring of NADH and flavin adenine dinucleotide (FAD) fluorescence lifetimes in patient-derived breast cancer organoids generated from the primary cells of various breast cancer subtypes exposed to therapeutic treatments exhibited differential effects in proliferating cells at the periphery of the organoid versus the lower metabolic activity of cells in their cores (30). FAD and NADH fluorescence correlated with DNA damage after radiation treatment in primary cancer cells and immortalized cancer cell lines (31), while increased FAD and NADH green fluorescence were observed in bacteria, yeast, and human cells exposed to cytotoxic or chemo static treatments (33). An increase in cell or tissue autofluorescence produced by agents that induce DNA damage, cell death, or generate reactive metabolites could be misinterpreted as an increase in fluorescent signal from a reporter or probe being monitored in the same spectral range when the HCS assay signal is relatively weak.

Seeding and growing cells

Other potential sources of HCS interference include exogenous substances ("contaminants") and microplate vessel materials or coatings (Figure 5A). This section reviews examples of potential contaminants and microplate effects, and recommended practices for mitigating their impact.

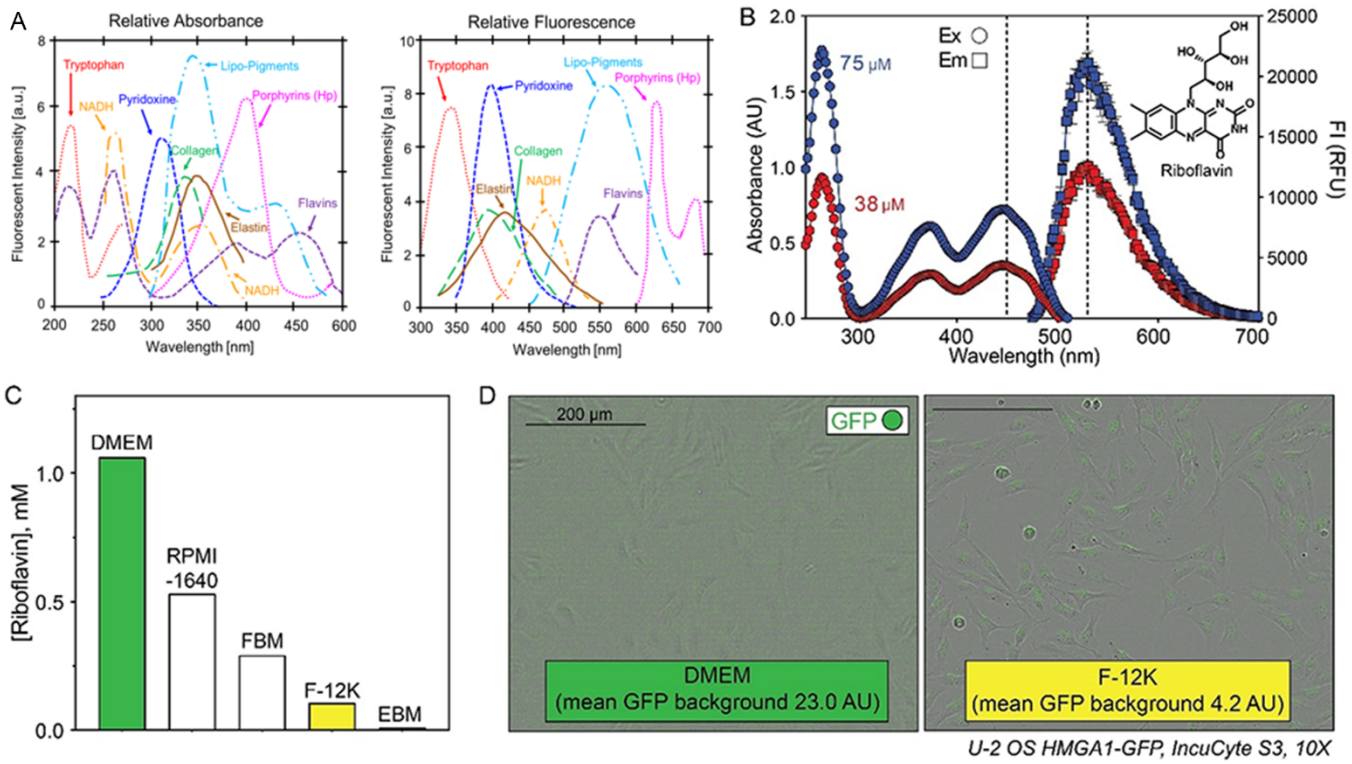


Figure 4. Endogenous sources of fluorescence. (A) Cells, tissues, and culture media contain a variety of endogenous substances that fluoresce in common fluorophore channels, often with relatively broad excitation and emission peaks. Adapted with permission (75,76). (B) Riboflavin absorbance and fluorescence in PBS (75). (C) Summary of riboflavin content in common cell culture media formulations (76). (D) Comparison of live-cell imaging of U-2 OS HMGA1-GFP cells using media containing low (F-12K) and high (DMEM) concentrations of riboflavin. Note the high GFP channel background signal from riboflavin in DMEM prevents recognition of GFP-labeled cells, unlike F-12K media.

The selection of microplates and their composition is an important decision for HCS assay development. Black-walled microplates are generally preferred for fluorescence microscopy assays because of low autofluorescence and decreased well-to-well crosstalk from light scatter (34). The composition of different microplate bottom surfaces such as polymers (polystyrene, cyclic olefin copolymer, cyclic olefin polymer) can exhibit varying levels of autofluorescence and may have profound effects on cell attachment and spreading. Most microplate materials specialized for imaging applications are designed and fabricated to increase optical light penetration for transmittance as well as reduce autofluorescence and light scatter to improve image resolution.

Hydrogels and extracellular matrix reagents, such as Matrigel, are typically gelatinous substances containing proteins, growth factors, and other constituents that are commonly used as a substrate overlay for primary cells as a matrix for 2D cells or as a scaffold to support 3D cell models. Some Matrigel components naturally fluoresce but overall, they tend to exhibit low autofluorescence for most applications (35,36). Furthermore, the increase in overall thickness of Matrigel coating of cells or the resulting embedding of 3D cells can interfere with or impede light penetration, scatter, and absorbance that may reduce image quality (37). Synthetic hydrogels such as polyethylene glycol ("PEG"), and natural hydrogels such as hyaluronic acid and plant-derived alginates have also been reported to exhibit autofluorescence (38). Therefore, it is recommended that microplate supplemental coatings and matrices should always be assessed for optical interference and/or autofluorescence. In practice, this requires obtaining quality imaging microplates with desired coatings (PDL, collagen, fibronectin, laminin, hydrogel, etc.) and testing each cell model for autofluorescence and adherence under reference control conditions. In addition to cell adherence, coated surfaces can alter cellular characteristics (size, shape, morphology, texture), which should be thoroughly evaluated during assay development before screening. For

additional details on microplates, types and choices for HCS imaging assays, consult the AGM chapter, Microplate Selection and Recommended Practices in High-throughput Screening and Quantitative Biology (34).

Environmental contaminants

Another source of artifacts captured by HCS imaging are exogenous contaminants from the laboratory environment including fibers from laboratory coats and paper consumables (Figure 5B), and pieces of plastic from pipette tips, bags, and lids. Inert contaminants may also be introduced at the time of microplate fabrication. In practice, these artifacts cannot be fully eliminated and may be introduced during the multiple processing steps in most HCS assays. To help reduce introduction of inert environmental artifacts, there are some standard HTS practices that can be adopted for HCS. This includes using laminar-flow hoods or contamination-free cabinets to seed, treat, and label cells; using microplates pre-barcoded from the manufacturer (rather than removing lids to place barcodes on the sides of microplates); thoroughly cleaning automation devices (e.g., plate washers, centrifuges, pre-cleaning plates with an ionized air gun), filtering solutions when possible if it does not impede assay performance (e.g., media, buffers), and even utilizing non-cotton-based laboratory coats (39). Lint contamination of cell models can create havoc with interpretation of the data as fluorescent debris light scatter travels considerable distances. This can create objects that can be scored and flagged for hit analysis review in 2D images and in reconstructed 3D image stacks (Figure 5C), highlighting the importance of such pre-screen activities.

Microorganism contaminants

Microorganisms are environmental contaminants that can produce a variety of interferences in HCS assays, whether by direct observation of the organisms themselves in images or by secondary effects on the assayed biology (Figure 5D). Just as with other assay formats, HCS practitioners should always exclude experiments with clear evidence of microbial contamination from further consideration. *Mycoplasma* contamination is frequently a source of poor reproducibility which often goes undetected in cell cultures only to be revealed in the images of cells that have been stained with bright DNA stains such as Hoechst. It is recommended as good laboratory practices that immortalized cell lines and cell types being used in the experiment undergo cell line authentication to verify cell type, and the potential contamination from other cells, i.e., HeLa cells and testing of *Mycoplasma*. To reduce potential contamination in HCS cell-based assays, best practices consistent with good sterile technique and media preparation should be applied. However, depending on the assay workflow and instrumentation setup, certain experimental steps may not occur in completely sterile environments (e.g., compound addition by robotic transfer head, pin tool, or acoustic ejection).

Tips: addressing bacterial and fungal contamination:

- Prepare all reagents for cellular experiments using aseptic technique.
- Including broad-spectrum antibiotics (penicillin-streptomycin or gentamycin) in culture media is common for many tissue culture applications including HTS and HCS. However, antibiotics can potentially alter the measured biology, and even mask poor tissue culturing technique. Therefore, the decision to include antibiotics into an HCS workflow should be carefully considered. Note: 1) eliminating antibiotics is essential for many genetic transfection methods; 2) some bacteria, particularly *Mycoplasma*, are growth-suppressed but not eliminated by antimicrobial agents. 3) Routine testing of cellular cultures for *Mycoplasma* or yeast contamination and elimination of contaminated stocks (there are commercially available *Mycoplasma* test reagents and yeast can be identified under HCS microscopy prior to scale up for large screen).

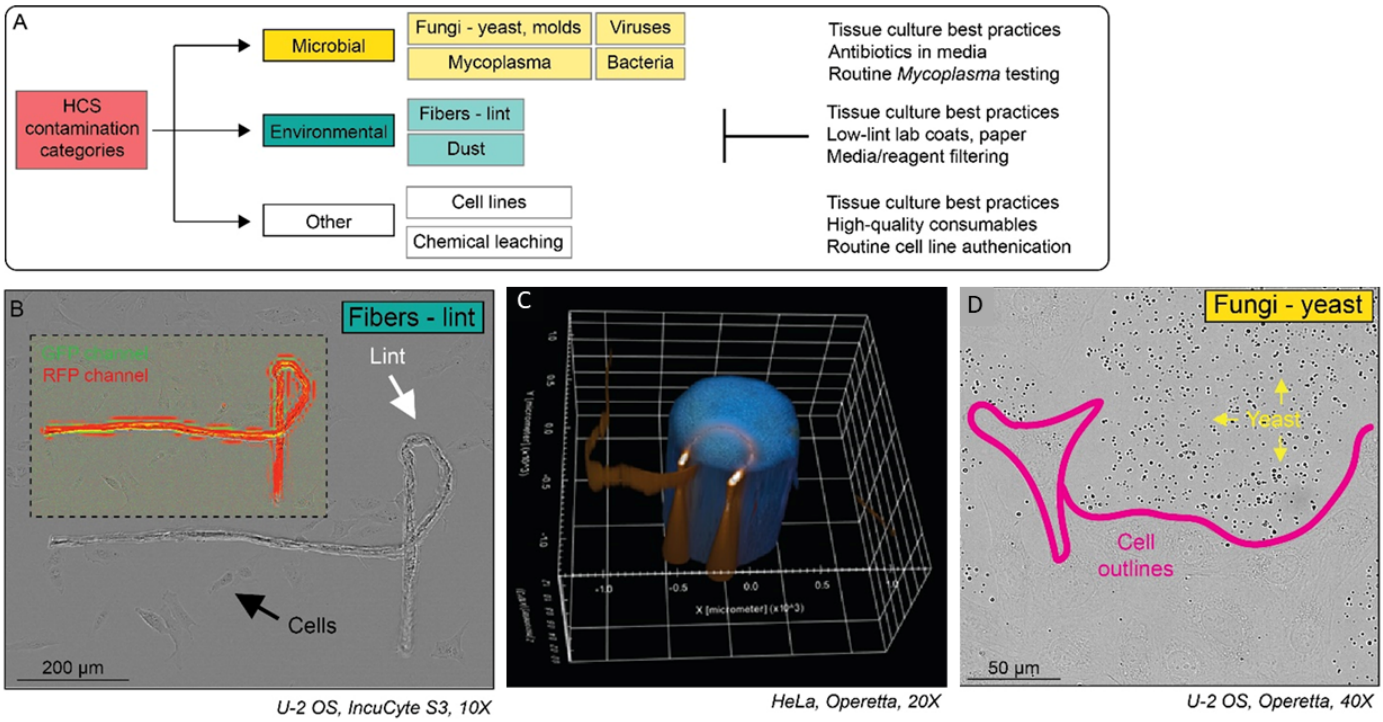


Figure 5. Contamination interferences in high-content screening assays. (A) Summary of microbial, environmental and other contaminants. (B) Environmental contaminants in HCS. Note that lint fibers from laboratory coats or paper towels produce characteristic large (approximately one-half millimeter scale) strand-shaped artifacts that can show up even in fluorescence imaging channels. (C) Lint (brown) light-scatter through a 3D cell model in HeLa cells labeled with Hoechst 33342 (blue); 20xW lens, 120 µm captured from 60 Z-slices at 2 µm intervals, 3D reconstruction and rendering, Operetta CLS (Revvity, Waltham, MA). (D) Yeast contamination (yellow arrows, note characteristic 1-10 µm diameters and circular shapes) in U-2 OS cells (magenta), likely from a contaminated compound stock solution.

Other contaminants

Cross-contamination or misidentification of cell lines is a major challenge for cell-based assays including HCS assays (40-42). It is recommended that cell lines be regularly authenticated using molecular methods such as short tandem repeat profiling (40).

Tips: best practices for avoiding cell line contamination or misidentification:

- Obtaining cell lines only from trusted vendors or reliable sources.
- Performing initial cell line authentication and periodic re-authentication.
- Follow tissue culture best practices to maintain sterility and avoid cross-contamination.
- Minimizing simultaneous handling of multiple cell lines in culture hoods.

Labels and probe choices

The selection of fluorescent labels and probes is a critical component of HCS assay development that can also have a profound impact on the type and pervasiveness of interference in an HCS campaign. This section reviews potential sources of artifacts and interference related to the selection of HCS labels and probes, and recommended practices for mitigating their impact.

The number of compounds that show substantive autofluorescence is generally lower at red-shifted excitation and emission wavelengths (6,10). Selecting fluorescent reporters or probes with red-shifted excitation and emission wavelengths is one tactic that can reduce the incidence of compound-dependent autofluorescence interference in screening assays (9,43). Examples include mCherry and RFP derivatives (relative to GFP) and newer near-infrared (NIR) synthetic dyes. Red nuclear dyes such as DRAQ5 or Sytox Deep Red are accepted alternatives to UV-excited blue nuclear stains like DAPI and the Hoechst dyes.

Image acquisition considerations

The autofocusing method used during the automated image acquisition can impact image quality and data generation. In brief, all HCS imaging systems are equipped with autofocusing hardware and software components to rapidly locate and identify objects of interest in the correct focal plane. The two most common autofocus methods are laser-based autofocus (LAF) or image-based autofocus (IAF). It will be important for HCS Practitioners to match microplate types with AF routines for proper focal plane of cells, otherwise, out-of-focus images may be captured affecting data generation. For additional details on image autofocusing, consult the Assay Guidance Manual chapter Assay Development Guidelines for Image-Based High Content Screening, High Content Analysis and High Content Imaging (20).

The detection and compensation of certain types of interference may be specific to one image acquisition process but not to another. Compounds that exhibit autofluorescence or quenching in any of the acquired fluorescent channels may impact the image analysis algorithm from accurately segmenting the images in those channels to define objects and/or regions of interest and to extract the desired quantitative data. The impact of compound autofluorescence or quenching on 'normal' image acquisition results from various factors, including out-of-focus images, segmentation algorithm failures, and obscuration of small-molecule activity at the target of interest. Depending on the desired activity profile (activation or inhibition), this can lead to false positives or negative scoring. This is observed more often in IAF acquisition processes when an identifiable artifact is in one or more focal planes, creating problems previously mentioned.

Artifacts and autofluorescence from compounds will rarely directly affect LAF image acquisition. However, technical challenges arise from interferences of the microplates or vessels used to host cell models. If these devices are too thick or thin, or constructed with components that impede NIR light scatter, then these can create unwanted out-of-focus images. Selection of a suitable high-quality microplate is therefore a critical component of the assay development process. Additionally, the use of different LAF thresholding strategies (i.e., single, or double peak) may be useful to identify the correct focal plane containing the cells of interest. And in some cases, a combination of LAF and IAF can be used. The use of transmission brightfield imaging on the HCS imager or conventional microscope can augment the fluorescent image acquisition approach for morphological evaluation and can be very useful to help determine if the source of the artifact is lint, microbes, compound precipitant, or other debris. This label-free approach has many advantages including no interference from fluorescent molecules or compounds, and with the advances in artificial intelligence (AI) and machine learning (ML) with improved segmentation of a brightfield channel, offers value-added information from cell images.

The following are some best practices to enable autofocusing strategies for imaging:

- Use high quality validated microplates or vessels intended for HCS imaging experiments.
- Determine the focal plane (i.e., offset) of the cell model to optimize the image autofocus routine.
- When necessary, such as natural product or other suspect libraries, collect transmission brightfield images to augment and confirm potential sources of artifacts.

Special considerations for live-cell imaging

Compound autofluorescence can present unique challenges for live-cell imaging, mainly because cells are continuously exposed to compounds throughout the experiment (Figure 6A).

- Poorly soluble compounds can precipitate and form crystals or fine precipitants that can adhere to microplate surfaces (Figure 6B, C).
- Large, circular fluorescent patterns can form at the compound delivery site (Figure 6C). These artifacts can sometimes be addressed by adjusting the mechanics of the acoustic or pin tool dispensing system including xy centering or offset and Z-height, including adding a gentle mixing step, or optimizing compound concentration/solvent/delivery volume.
- More prominently, compounds that fluoresce in the fluorophore channels can increase background fluorescence. Even with various background subtraction and segmentation smoothing filter methods, fluorescent compounds can still produce localized signal within cellular compartments or entire cells (Figure 6D).
- Cellular debris from damaged cells can also interfere with certain automated confluence measurements (Figure 6E). This phenomenon is associated with surfactants, such as saponins, or protein precipitants, such as tannins. Cellular debris interference from cytotoxic compounds is usually less of an issue in assays that include washing steps prior to fixation.

In live-cell imaging, continuous or pulse exposure of fluorescent light to cells, compounds, or other constituents in the well may inadvertently result in cytotoxicity from photon exposure (phototoxicity) or photo-bleaching. To reduce potential issues from excitation light during live-cell imaging, reduce the number of times the same field of view is imaged. Assay development parameters should be established, typically for fluorescent proteins, lower excitation power with longer exposure times results in lower photobleaching than high energy excitation power and shorter exposure times.

Analysis

The sheer volume of feature data generated from HCS necessitates automated image and data analyses. In addition to helping identify bioactive compounds, well-designed analyses are crucial for flagging potential artifacts and interferences amongst the large number of images generated by HCS. These approaches are critical for prioritizing the highest quality compounds for follow-up and the design of appropriate orthogonal experiments. This section describes guidance for image analysis, recommended statistical methods and HCS readouts for outlier analyses, and cheminformatic tools for flagging potential interferences.

Image analysis

Artifacts can be flagged by analyzing HCS readouts for outliers. While it may not be practical to visualize all images captured by HCS campaigns, it is good practice to view a subset of the outliers. In the simplest form, image analysis algorithms are used to identify and segment cells to determine the number, size, shape, or intensity of individual level objects summarized as cell-level, field-level, or well-level data. Some direct image analysis approaches and strategies combine multiple HCS feature data that use formulae, Boolean logic, or weighting techniques to identify subpopulations of cells (44). These can either be an informative biological subpopulation or a cluster of artifactual events. Other image analysis approaches use supervised machine learning techniques to measure texture and morphology to classify cells, some of which *may* fall under an artifact classification if identifiable and segmented in the image. This approach is enabled by teaching the image analysis algorithms using an informer set or other known artifactual fluorescent compounds.

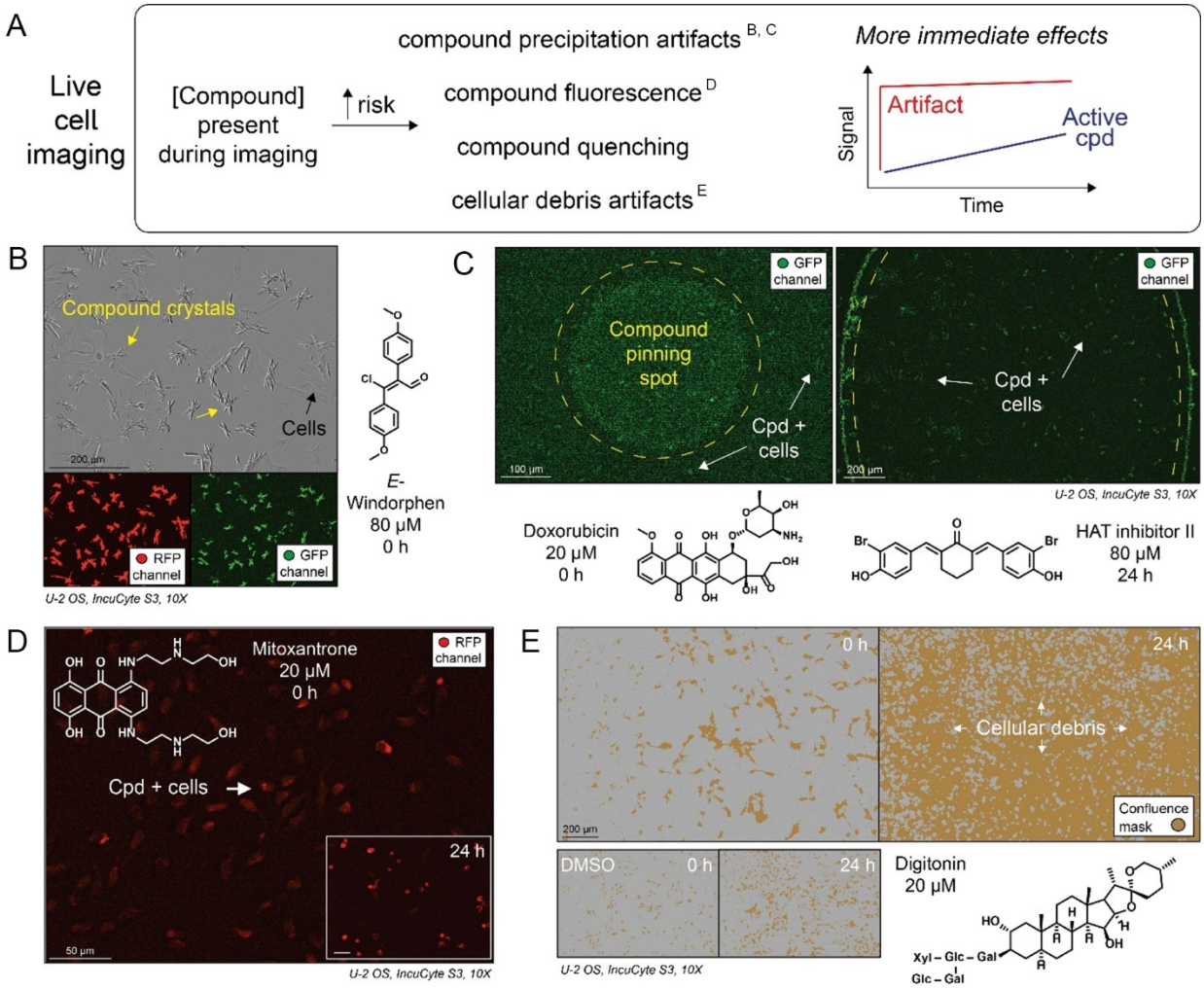


Figure 6. Compound-mediated interferences in live-cell imaging assays. Since compounds are not removed during the assay, live-cell assays are more susceptible to certain types of interference. A major clue of compound-mediated light interferences are positive readouts within hours of compound treatment. (A) An overview of live cell imaging considerations includes potential compound interferences such as compound precipitation artifacts, compound fluorescence, compound quenching (not shown), and cellular debris artifacts. Peak fluorescent signals from these artifacts are generally observed immediately during imaging, whereas active compound fluorescence may develop over time (B) Compound precipitants (crystals) can produce intense, shaped fluorescent readouts in live-cell imagers. (C, D) Compounds can exhibit autofluorescence in the GFP and RFP channels. (C) Note high local concentrations of compound from compound dispensing can produce halo-like artifacts. (E) Interference by cellular debris. Surfactants often induce rapid cell permeabilization. Noticeable debris appears within 24 h. Note that cellular debris from lysed cells interferes with automated confluence measurements over time (compared to DMSO controls).

Data analysis

The interpretation of outliers will depend on the experimental design and assay construction. In gain-of-signal assays, fluorescent and quenching compounds could produce false-positive and false-negative readouts, respectively. Conversely, in loss-of-signal assays, quenchers may lead to false-positive readouts. There are several options for calculating or choosing outlier cut-offs, each with specific pros and cons.

Identification of active compound 'hits' in typical HCS assays based on sample distribution is usually defined by the statistical threshold of three standard deviations (SD). The sample set for calculating means and sample distributions can include active, inactive controls, or other test article or training set compounds. An advantage of distribution-based analyses is that they incorporate assay imprecision. A potential disadvantage is that the sample distribution may not adequately reflect outlier behavior[1]. Converting HCS readouts to Z-scores means

compound measurements are rescaled relative to within-plate variation. Z-score analysis of multiparameter HCS data also provides a relatively straightforward method to identify compounds that are cytotoxic or disrupt cell adhesion to substantially reduce cell number, compounds that are autofluorescent in any of the acquired channels, and wells with out of focus fields or large contaminants. A detailed review of Z-score is defined in Appendix 1.

- Example 1: Flagging a compound that exhibits mean object fluorescent intensities ≥3 x SD above the mean of all non-control compounds on a microplate, Z-score > 3.
- Example 2: Flagging a compound that produces mean cell counts ≥3 x SD below the mean of all non-control compounds on a microplate, Z-score <-3.

Preset cut-offs or thresholds are sometimes not based on sample distributions. This approach can be arbitrary, and such cut-offs are usually selected based on expected readouts for the given biology and/or assay technology. Depending on factors such as the assay, size of library screened, capacity to follow up, and hit frequency, the actual threshold for hit calling may be more stringent than a Z-score = 3 (which is often used as a conventional starting point for hit calling). Other compound selection criteria may also be a factor. When using these approaches for determining outliers, it is usually best to set such cut-offs prior to the HCS.

- Example 1: Flagging of a compound that reduces cell counts below 50% of vehicle controls (e.g., pre-determined project criteria for defining cytotoxicity).
- Example 2: Flagging a compound that produces mean object fluorescent intensities above a preset threshold of 1000 units when the typical values fall within the 0-to-100-unit range.

Alternatively, when conducting High-Content Screening (HCS) in quantitative high-throughput screening (qHTS) mode (45), compound activity evaluation relies on the pharmacological response obtained from a compound titration spanning 5-7 orders of magnitude. In this context, concentration-response curve (CRC) classification criteria have been developed to analyze the structure-activity relationships (SAR) of chemotypes based on the CRCs and library chemotype substructures (46, 47).

Types of readouts

Assay endpoints can be simple (object number, target intensity such as for a kinase site-sensitive antibody), algebraic (nuclear/cytoplasmic ratio, cell cycle distribution) or complex (high-dimensional algorithms that include texture measurements). Fluorescence interference can affect all of these, as well as how such measurements are presented. In terms of increasing subtlety:

- *Object counts.* Fluorescent artifacts such as compound precipitates may be identified as a large number of fluorescent objects compared to controls in a given imaging channel if they fall within the object identification parameters (typically pixel area but can include shape parameters as well).
- *Outliers among scaled measurements.* Autofluorescent and quenching artifacts may produce unintended outliers from high background fluorescence intensity measurements. For example, the mean and integrated object fluorescent intensities in each HCS channel can be calculated, and compounds that result in higher Z-scores (e.g., > +3.0) would be flagged as likely autofluorescent compounds. Conversely, compounds that result in lower Z-scores (e.g., < -3.0) would be flagged as likely light absorbing/fluorescent quenching compounds (Figure 7).
- *Well-level readouts versus object-level readouts versus image-level analyses.* Analyzing readouts at the object level (e.g., per nucleus, per cell, per colony or cluster) can help to reduce the impact of background signal from fluorescent extracellular compounds. Measuring the fluorescence per well typically calculates

[1] The Z-score method assumes that in a randomly distributed compound library, active compounds are relatively rare, and that most of the compounds on a plate are inactive and can serve as controls. *Therefore, in a plate enriched with bioactive compounds or artifacts, this assumption may not hold, and the Z-score may not be the best choice.*

summary statistics (mean, median, etc.) for the objects in a well and shares the same advantages for eliminating aggregate and contamination-based fluorescent artifacts. Image-level analysis (object- or segmentation-free analysis of the image), such as deep-learning algorithms, will include fluorescent signal from extracellular compounds or other fluorescent substances. Also note that for object-based assays, post illumination flatfield correction to remove vignetting effect, algorithms that use background correction may inappropriately offset the intensity measurement if autofluorescent debris is outside the defined object area. Careful assessment of reference control wells to established background intensity parameters should be verified.

Observations and insights from compound treatments

Compounds are the major source of HCS artifacts and interferences. The selection of test compounds, details of how they will be tested, and the inclusion of compound controls can all have significant downstream consequences in HCS not just in the general context of a screening campaign, but with regards to how fluorescence interference is detected and addressed. This section describes the two main categories of compound dependent HCS interference: compound autofluorescence and quenching, and nonspecific cytotoxicity or reduced cell adherence. Also described are reference compounds that can be used to identify or flag potential sources of compound interference in HCS assays.

Interference magnitude

When viewing HCS images and analyzing for autofluorescence interference, it is critical to consider the magnitude of fluorescence interference (Figure 8A). For compounds with low intrinsic fluorescence, the potential for interference increases if the corresponding dye is weak (Figure 8B). The dynamic range of fluorescent proteins (particularly in the green range) is extremely wide, so robust options are typically available (48). Weak staining also occurs when only suboptimal primary antibodies are available for a target protein, and for cases where expression of a labeled target protein is deliberately kept to the lowest level achievable (for example keeping the expression of a GFP-fusion protein low to minimize effects in the cell that may include results from a subpopulation biological response). HCS assays with large signal intensities should be inherently more resistant to fluorescence interferences. It is also important to note that some compounds (or their associated impurities) are strongly fluorescent, with intensities far above those of the fluorescent dyes. These can be flagged as significantly higher well-level and object-level intensities (e.g., by Z-score) and most likely affecting other features (such as texture) if these are part of the assay metric.

Direct compound-mediated fluorescent artifacts

Compound autofluorescence can result in image sets that are out of focus, create image analysis failures, may obscure small-molecule activity at the target of interest, and can be scored as false positives or negatives depending upon the desired activity profile, activation or inhibition. This is a significant issue because many HCS assay formats measure an increase in fluorescent signal intensity, while spectroscopic profiling of small-molecule HTS libraries have shown that about 10% of compounds are fluorescent (8,10). As a practical example, the EPA ToxCast compound library was screened in a 3 channels (blue, green, & red) biosensor assay for agonist and/or antagonist activity against DHT-induced androgen receptor protein-protein interactions with transcription intermediary factor 2 (TIF2/SRC2) a p160 coactivator (13,14) and was evaluated for fluorescent artifacts as part of the hit evaluation process (19). A detailed review of this process is included as Appendix 2. A key lesson learned from this study is that fluorescent anomalies resulting from cytotoxicity are frequent and their discovery can be made through reviewing wells with increased fluorescence in each channel and a correspondingly high Z-score.

In some situations, it may be possible to identify compound autofluorescence by incorporating a fluorescent pre-read of the compound library in a plate reader (or low magnification whole well or wells imaging) at the defined

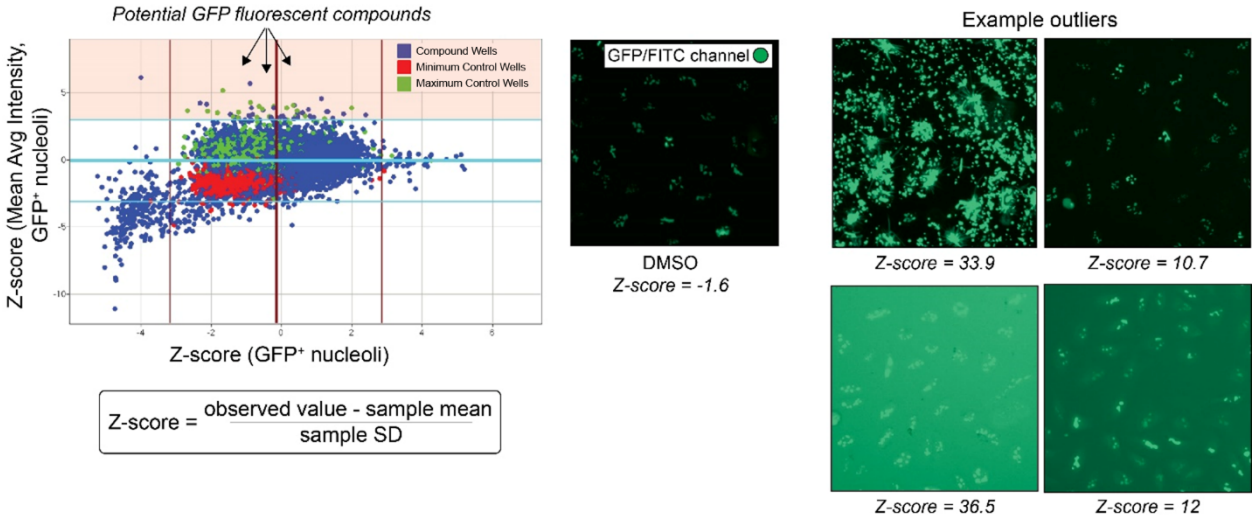


Figure 7. Example of using Z-score to identify compound interferences. Z-scores can be calculated for HCS features/readouts that may be informative for particular interferences (fluorescence, brightness; cytotoxicity, nuclei counts; quenching; dimness). (Left) In the example shown for data acquired in an AR-RFP/TIF2-GFP biosensor HCS assay, the Z-scores for the mean average fluorescent intensity of TIF2-GFP positive nucleoli for compound-treated cells in the GFP/FITC channel are plotted on the Y-axis and the corresponding Z-scores for the number of TIF2-GFP nucleoli are plotted on the X-axis. Compound wells blue circles, DHT treated maximum controls red circles, and DMSO treated minimum controls green circles. The thick blue and red lines mark the average (Z-score = 0) fluorescent intensity and number of TIF2-GFP positive nucleoli respectively for the compound treated wells, while the thin red and blue line lines indicate Z-scores of -3 and 3 respectively. Cytotoxic (number of TIF2-GFP nucleoli Z-scores <-3) and fluorescent outlier (mean average fluorescent intensity Z-scores >3) can be readily identified. (Right) Pseudo colored FITC channel TIF2-GFP images of outlier compounds – those with Z-scores >3.0 flagged as potential autofluorescent compounds. For comparison, a pseudo colored FITC channel TIF2-GFP image from a DMSO control well is shown.

excitation and emission wavelengths for the channels being acquired prior to compound transfer into the HCS assay plate. However, there are numerous examples where autofluorescence only becomes detectable and/or apparent after compounds enter cells and bind to proteins, nucleic acids, and/or lipid macromolecules. The use of counter-screens to identify fluorescent artifacts usually involves cells treated with compounds but omitting the stains, probes and/or other reporters, this may include a counter-screen to identify cellular fluorescent artifacts by using a control cell line that lacks the biology or reporter of interest. It follows that compounds that fluoresce in the HCS imaging channels in the absence of stains/reagents would do so by an autofluorescent mechanism (Figure 9A). This would not be practical for an HCS campaign of compound libraries greater than a few hundred compounds but could be used to confirm autofluorescence identified by statistical methods. Compounds that fluoresce in cells can constitute between 2.5% (blue) to 0.5% (red) of a screening library, depending on the wavelength (6). In general, the incidence of fluorescent artifacts decreases with red-shifted (higher wavelength) channels. The exact impact of autofluorescent compounds ultimately depends on the assay format. In gain-of-signal assays, fluorescent compounds can present as false-positives (Figure 9B). Fluorescence will usually exhibit concentration-dependence, so in general decreasing compound concentrations will help mitigate this interference, though this must be balanced with the desire to identify weakly potent actives. As such, primary HTS campaigns are frequently run at compound concentrations of 10 - 20 µM (Figure 9C), oftentimes as a matter of policy and experience by the lab, and by the hit follow-up/confirmation plan post-HTS. The distribution of fluorescent compounds in cells varies between compounds, from nuclear patterns (often intercalators) to cytoplasmic to punctate (Figure 9D) and is difficult to predict reliably.

Some factors affecting the incidence of autofluorescent compound interference include:

- The concentration and spectral properties of a given compound.
- The spectral properties of assay reagents and image acquisition settings.

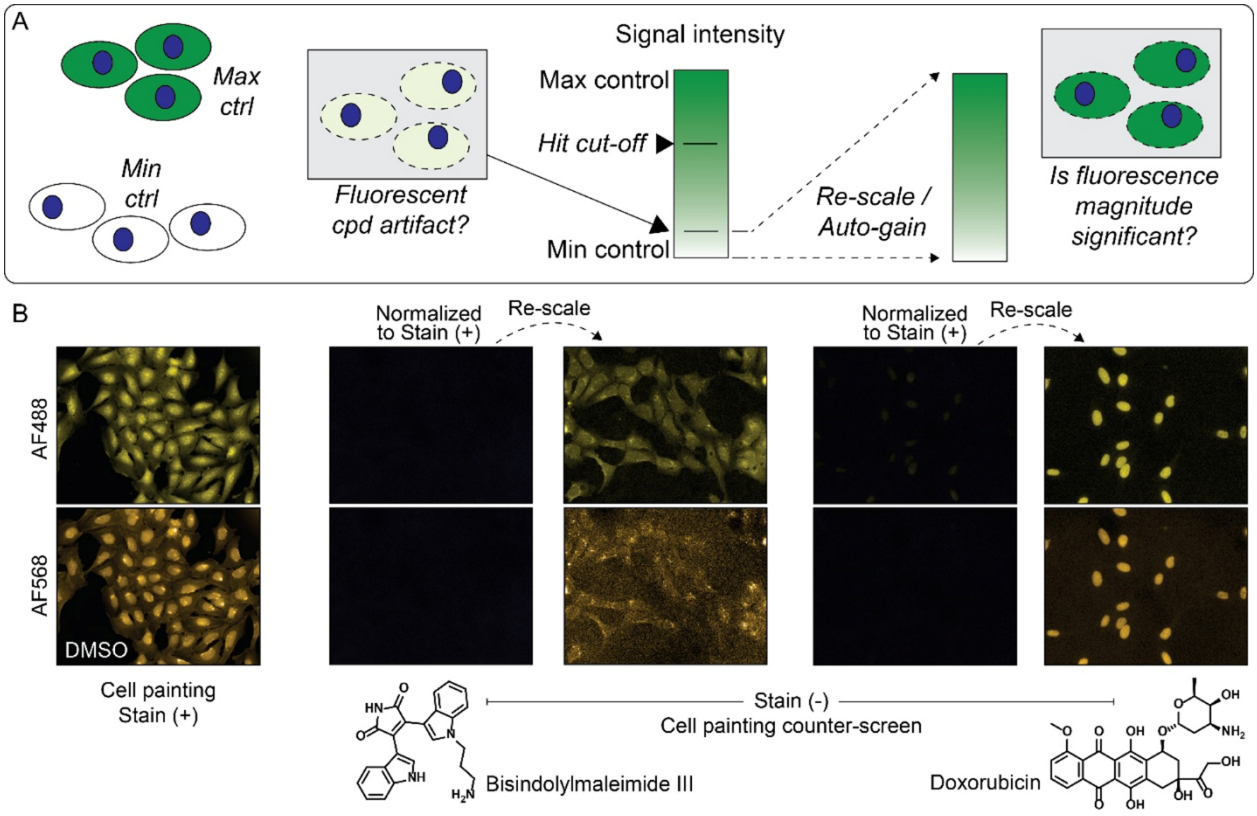


Figure 8. Impact of compound fluorescence depends on magnitude of interference. (A) Summary of fluorescent compound interference interpretation. Due to the high sensitivity of modern imagers, the magnitude of compound fluorescence should be interpreted with respect to bioactive controls. Right: viewing re-scaled images can suggest interference that may not be functionally significant. (B) Examples of weakly fluorescent test compounds revealed by a Cell Painting fluorescence interference counter-screen. Unstained compound-treated cells were imaged under identical conditions to the conventional Cell Painting method. Note when the intensities are normalized to stained cells, the magnitude of compound autofluorescence can be negligible. Increasing signal gain reveals weak compound autofluorescence.

- The intensity of the assay readout relative to the magnitude of compound fluorescence.
- The potential for the interfering agent to interact with one or more cellular components, increasing the fluorescence intensity or altering the spectra of the fluorescence.

In an HCS campaign to identify inhibitors of miR-21 biogenesis (8), all 1,130 primary hits from a screen of 315 K compounds were found to optically interfere with the enhanced green fluorescent protein (EGFP) reporter. The fluorescence interference was determined by counter-screening with the primary cell line minus the EGFP reporter. Methods for predicting compound fluorescence based on chemical structure are imperfect, especially in cellular contexts.

In addition to direct autofluorescence, some compounds that undergo degradation or are metabolized *in situ* can become fluorophores. This challenge is growing for image-based screening because morphological changes in assays like Cell Painting rely on defined pixel measurements of fluorescent regions within cells (49). Interfering fluorescent compounds can obscure these regions of interest, making it difficult to distinguish them from cell compartment bioprobes. As an example, the pro-drug pentamidine (DB289), a therapeutic for African sleeping sickness, has been reported to produce a cascade of five metabolites, including the product (DB75) which is highly fluorescent in the green spectrum (50,51).

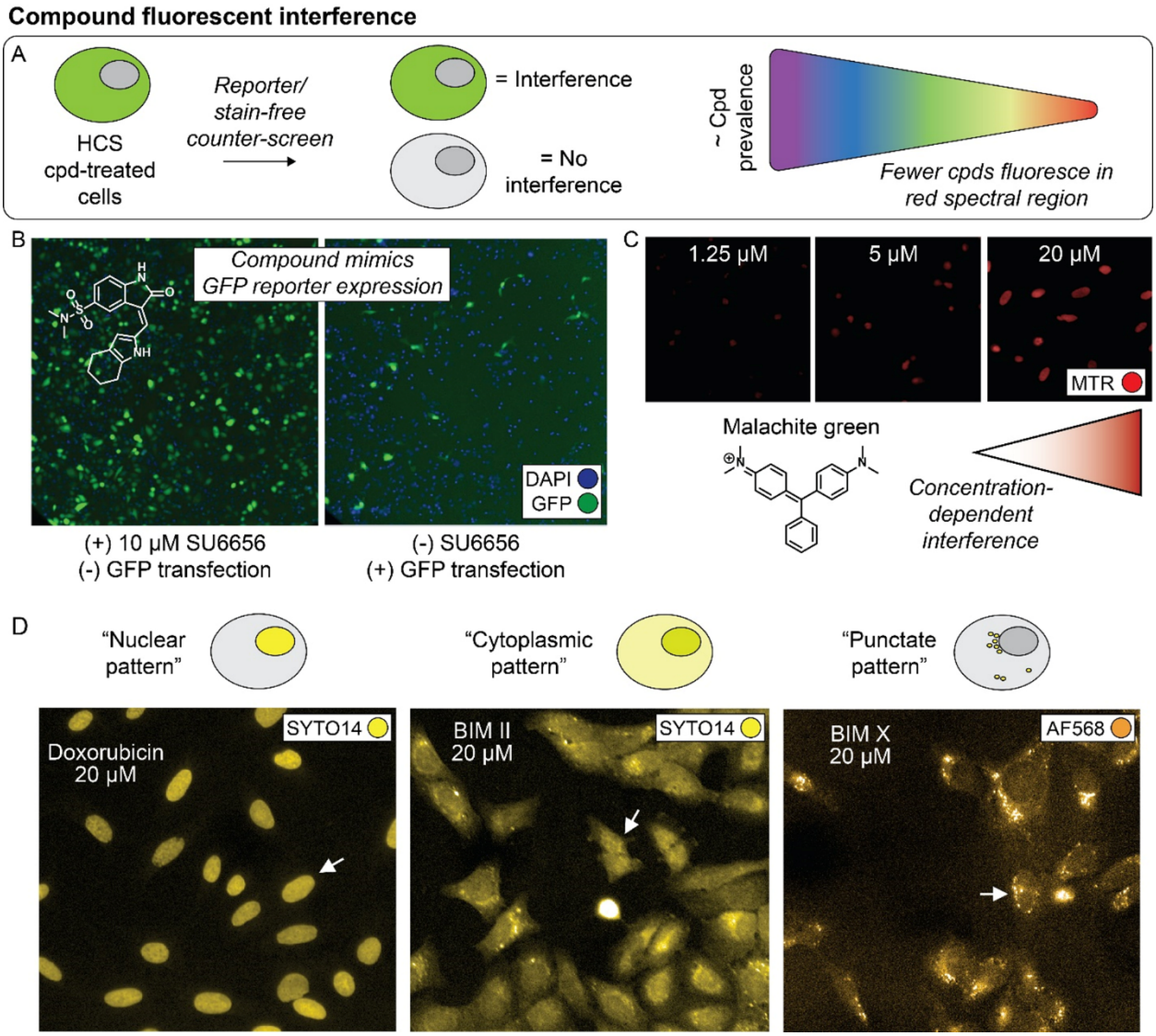


Figure 9. Fluorescent compound interferences in HCS assays. (A) Schematic of interference in HCS assays by fluorescent compounds. Fluorescent compounds will produce "active" readouts in the absence of stains/reporters, and are usually more prevalent in the blue and green fluorescent spectral regions. (B) Fluorescent compounds can interfere with the interpretation of cellular GFP reporters. Note that most SU6656-treated HeLa cells appear GFP-positive, while HeLa cells transfected with a plasmid harboring an eGFP reporter show 5-10% efficiency. (C) Fluorescent compounds usually display some concentration dependence. (D) Fluorescent compounds can produce a variety of patterns such as nuclear, cytoplasmic, and punctate patterns.

Quenchers

Broadly defined, fluorescent quenchers are compounds that attenuate fluorescence intensity of a given substance, and this phenomenon can occur by a variety of fundamental mechanisms such as chemical and physical quenching (Figure 10A), or by direct competition at the binding site for the fluorescent probe in use. In FRET formats, quenching can occur at one or more steps of the signal cascade between donor fluorophore excitation to acceptor fluorophore emission. The exact mechanisms of FRET fluorescence interferences are usually not determined during HCS triage.

In HCS assays, quencher detection may already be "built into" the assay and compounds that reduce the intensity of nuclear stains such as DAPI, DRAQ5, or Hoechst are flagged as potential quenchers (Figure 10B). Counter-screens for quencher artifacts usually involve brief compound treatment times to minimize the chance of biological perturbation. It follows that compounds that lead to reductions in fluorescent intensities would do

so by quenching. Like fluorescence, quenching will usually exhibit concentration-dependence, so in general decreasing compound concentrations will help mitigate this interference.

Some factors affecting the incidence of quenching interference include:

- The concentration and associated fluorescent spectral properties of a given compound quencher.
- The spectral properties of assay reagents and image acquisition settings.
- The intensity of the assay readout relative to the magnitude of compound quenching.

Quenchers have been used as tools for many cell-based assays including live-cell imaging for decades as they can be used to remove or reduce non-cell-based fluorescent signal. Many quenchers used in live-cell imaging are cell impermeant and therefore only block extracellular fluorescent signal, thereby effectively increasing the contrast of cell-specific signaling. While these quenchers can have great utility in live-cell imaging, small-molecules with quenching activity could prevent detection of a fluorescent signal and be considered a false positive or false negative depending on the nature of the HCS assay.

For additional details on compound-mediated light interferences, consult the Assay Guidance Manual chapters Interference with Fluorescence and Absorbance (52) and Compound-Mediated Assay Interferences in Homogenous Proximity Assays (53).

Compound informer sets

A practical approach to investigate phenotypes of interest for possible nuisance behaviors is the use of compound informer libraries. In many cases, such compound collections are composed of compounds with well-characterized activities to explore signaling pathways ('MOA-box') and previous clinical use (repurposing libraries) (54). In the context of assay artifacts and interference, such an informer set is a small collection of prototypical interference compounds (Figure 11) (7). Employing this approach from the onset helps to estimate the likelihood of specific interferences for a particular assay and characterizes the interference phenotypes which are difficult to predict *a priori*. In this way, screening compounds that produce similar phenotypes as informer nuisance compounds are clustered and can be triaged or investigated by additional counter-screens. Many industrial screening operations employ nuisance informer sets (55). The following are additional tips for using nuisance compound informer sets:

- Informer sets can include technology-related light generated interference compounds (fluorescent compounds, quenchers, light scattering compounds) and non-technology interference compounds (nonspecific electrophiles, aggregators, redox cyclers, surfactants, and various cytotoxins), among other classes.
- An example informer set has been reported (see *Supplemental Information* of reference (7)).
- Test compounds at multiple concentrations, as interferences are often concentration-dependent and can vary from assay-to-assay.
- Include multiple compounds (both analogs and distinct chemotypes) from each interference class, and inactive analogs when possible.
- Interference informer sets can be used during the assay development/validation phase and/or during the primary HCS.
- We recommend including high-quality chemical probes and reference compounds in parallel, since nuisance and "on-target" phenotypes can potentially overlap.

An informer collection focused on known fluorescent compounds, particularly across spectral wavelengths and mechanisms (such as those that change fluorescence as a function of their local environment) could alert a group on the sensitivity of an assay to fluorescence interference and which types of interference are most difficult to deconvolute.

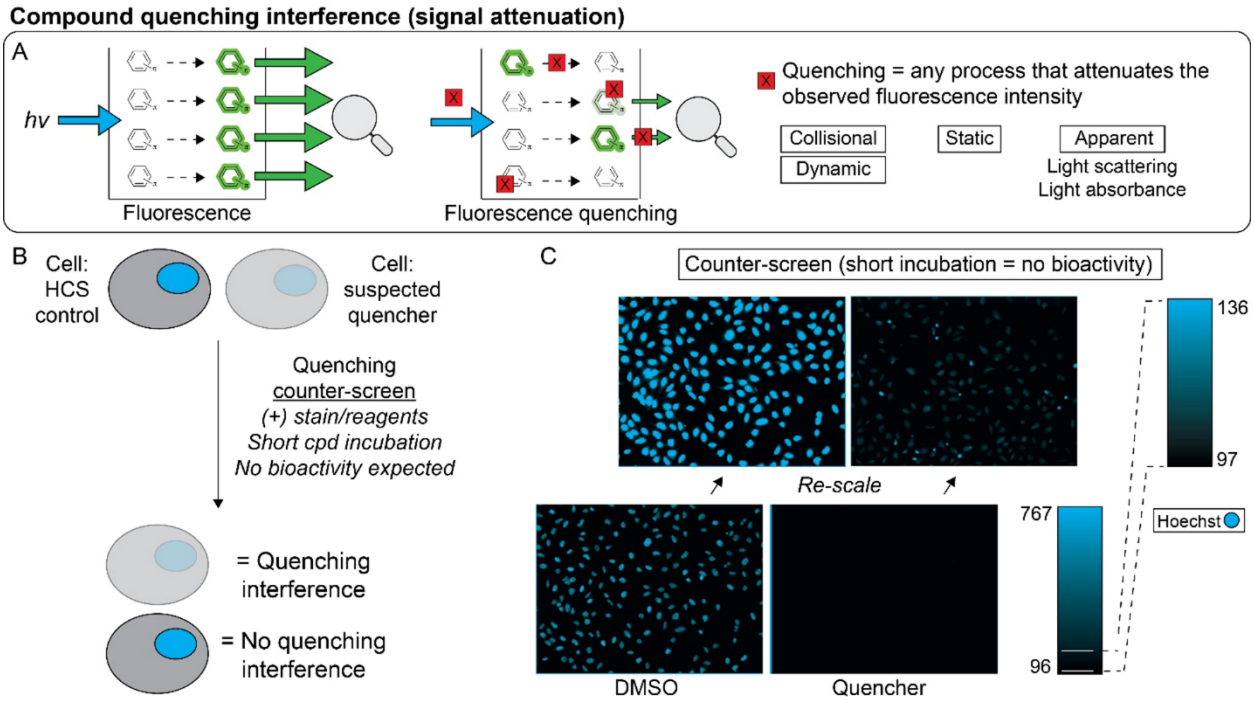


Figure 10. Quenching compound interferences in HCS assays. (A) Schematic of interference in HCS assays by quenching compounds. Quenching compounds will produce "active" readouts in loss-of-signal assays by attenuating the expected fluorescence intensity without modulating the bioactivity. This can occur by several chemical mechanisms. (B) Suspected quenchers can be identified by counter-screens that treat cells with compounds for a limited time to minimize any bioactivity. (C) Example of a fluorescence quencher (Hoechst channel) identified by a quenching counter-screen for an AR-RFP/TIF2-GFP biosensor HCS. Note that cells can be identified by including a nuclear stain and re-scaling the intensity.

Structure-based and cheminformatics-based fluorescence predictions

While general rules for predicting intrinsic compound fluorescence exist, the prediction of compound fluorescence based on chemical structure is still a nontrivial endeavor and algorithms tend to score with 70-80% accuracy, too low to rely on (6,56). In general, compounds with conjugated electronic systems ('aromatics') can better absorb and fluoresce UV and visible light, but chemical substituents and other structural features can have complex effects on optical properties. Related to this, the property of how intrinsic fluorescence changes as a function of the environment is also difficult to predict. Compounds can change excitation/emission wavelengths in a less polar environment, such as in membranes or when complexed to proteins. The prediction of compound fluorescence quenching is even less straightforward, as this can occur by several mechanisms: static quenching (complex formation), collisional quenching, FRET, and others. Cheminformatics can assist with fluorescent compound prediction. These methods generally require HTS-scale counter-screens to profile compounds for specific interferences. An example open-source tool from the National Toxicology Program is InterPred which uses random forest machine learning to predict interference of active versus inactive chemicals (6,56). While not an approach to define fluorescent compounds explicitly, looking at the distribution of structural motifs in compounds to be screened can alert one to the potential for interference and trigger a mitigation plan for dealing with the rare one-off, or a broader approach. For significant screening campaigns, complementing this *in silico* approach with the development of a corresponding informer set of compounds (discussed above) can build robustness into the effort.

Another approach involves flagging compounds with increased hit rates in fluorescence-based assays. While standard practice in established screening laboratories, such institutional data may be sparse in laboratories starting out. Because of the potential for false-negative and false-positive predictions, it is necessary to confirm

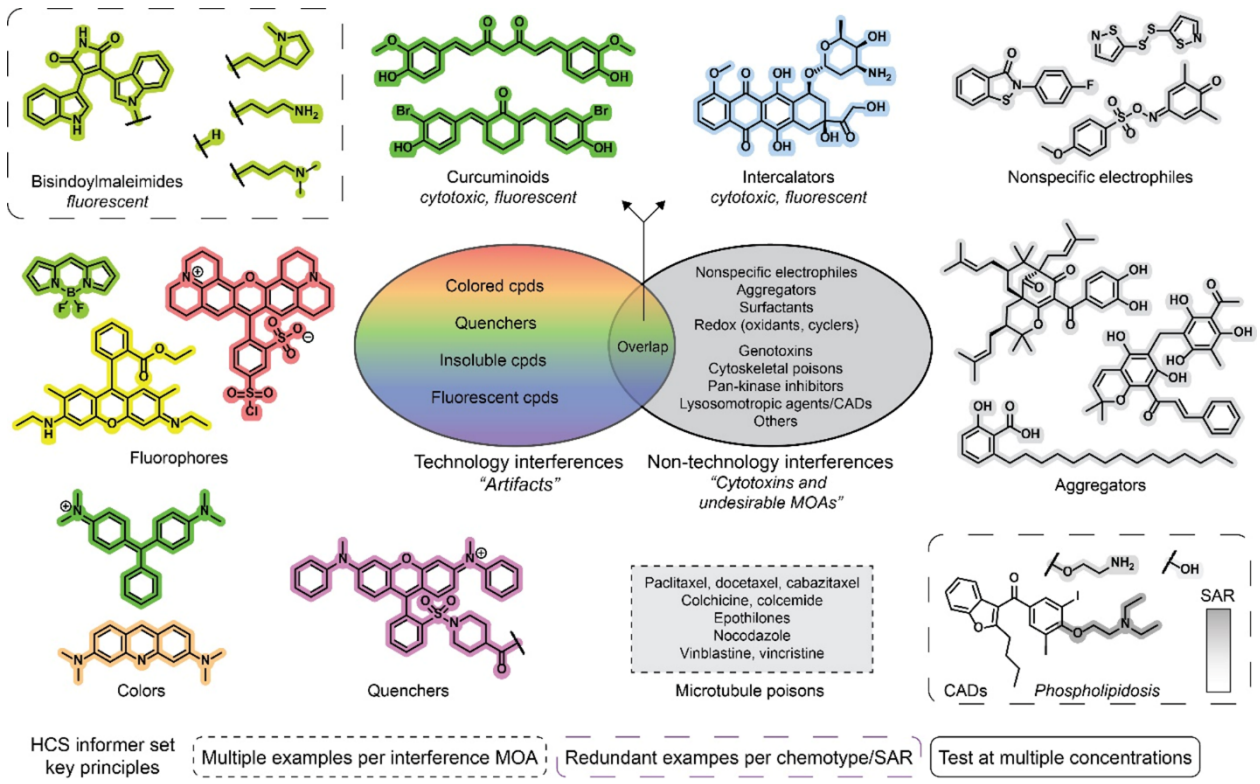


Figure 11. Schematic for the design of an HCS nuisance compound informer set. It is recommended to use multiple chemotypes per interference MOA, multiple analogs per chemotype, and testing at several concentrations. Note that compounds can both interfere with the assay technology but also produce phenotypes by undesirable MOA. Nuisance informer sets can be used for: 1) compounds that produce HCS readouts similar to known nuisance compounds should undergo further scrutiny, 2) HCS assays can be modified during development if they are highly susceptible to certain interferences, and 3) efficient screening cascades can be designed based on nuisance MOA sensitivities. See also Supplemental Information of reference (7) for an example informer set.

by a wet-lab experiment the cheminformatic predictions on high-priority compounds. For example, samples can contain highly fluorescent impurities that may only be revealed by experimental counter-screening or looking for activity in previous fluorescent assays (10).

Confirmation and follow-up

After identifying compounds of interest by appropriate data analyses, it is crucial to perform a series of experiments on HCS-active compounds to dually select for desirable compounds and triage undesirable compounds. This screening cascade should assess for artifacts, irreproducible hits, and undesirable MOAs, and triaging as indicated. It should include orthogonal assay formats and counter screens designed and selected to confirm activity at the target or phenotype of interest, evaluate selectivity and/or specificity, and evaluate cytotoxicity. This section details experimental counter-screens for compound fluorescence and quenching in HCS assays and summarizes other high-yield secondary assays in the screening cascade for these purposes.

Optical interference counter-screens

A recommended counter-screen for fluorescent compound interference involves testing active compounds in the absence of stains, labels, or reporters (Figure 12A). For example, in a fluorescent protein assay (i.e., GFP), one can use a parental cell line without transfecting the GFP reporter (19,48). A recommended counter-screen for quenching compound interference involves testing active compounds in the presence of stains or reporters, but with short incubation times to preclude actual bioactivity (Figure 12B). Both significant fluorescent and quenching compounds should generally show similar readouts as the primary HCS readout.

An overview of these process steps including technical information is highlighted in Table 1.