



Considerations in Developing Reporter Gene Bioassays for Biologics

Jamison Grailer, Richard A. Moravec, Zhijie Jey Cheng, Manuela Grassi, Vanessa Ott, Frank Fan, and Mei Cong

Abstract

The establishment of a robust and reproducible functional bioassay that reliably measures drug potency while ascertaining its mode of action is essential in biologic drug development. Here we describe a simple bioluminescent reporter gene bioassay for assessing biologics targeting immune checkpoints without the complexity and variability of more traditional assay systems. This chapter provides an overview of key considerations in reporter gene bioassay design and optimization, as well as development of thaw-and-use cells as an assay reagent for biologic QC lot release.

Key words Biologics, Immunotherapy, Reporter bioassay, Thaw-and-use cells, Quality control, Drug potency

1 Introduction

Vaccines, blood and blood components, allergenics, tissues, and recombinant therapeutic proteins are biological products, comprising more than 50% of new drugs in development. Among the different classes of biologics, monoclonal antibodies (mAbs) represent the most broadly developed therapeutics. The continued increasing R&D investment in biologic drug development reflects their specificity and thus fewer off-target effects, as well as longer exposure compared to small-molecule-based therapeutics.

Biologics have emerged as promising new approaches to immunotherapy, which aim to utilize the patient's own immune system to combat diseases such as cancer and autoimmunity. Some of these approaches include recombinant cytokines, immune checkpoint

The original version of this chapter was revised. The correction to this chapter is available at https://doi.org/10.1007/978-1-0716-0171-6_14

Jamison Grailer, Richard A. Moravec, and Zhijie Jey Cheng contributed equally to this work.

Seng-Lai Tan (ed.), *Immuno-Oncology: Cellular and Translational Approaches*, Methods in Pharmacology and Toxicology, https://doi.org/10.1007/978-1-0716-0171-6_9, © The Author(s) 2020, corrected publication 2020

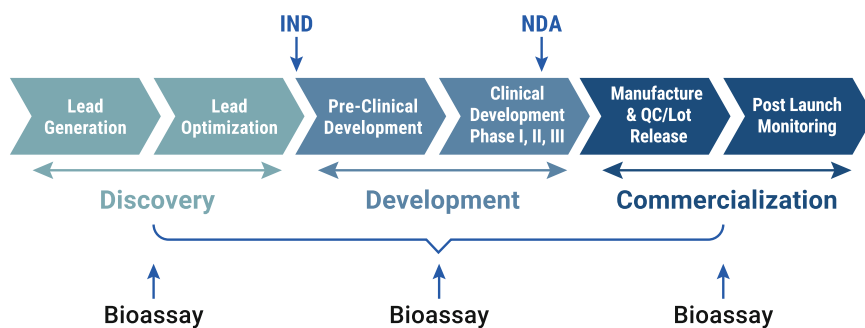


Fig. 1 Bioassays can fit into every stage of biological drug development

mAbs, bispecific molecules, antibody-drug conjugates, and chimeric antigen receptor (CAR) T cell therapies. However, the inherent complexity, functional heterogeneity, and sensitivity of biologics to external storage conditions present unique challenges in drug discovery and development, and the development and manufacture of these products must be tightly controlled to ensure consistent purity, potency, efficacy, and safety. To address these challenges, functional, mechanism of action (MOA)-based bioassays are used throughout the biologics drug discovery and development workflow to screen, characterize drug MOA, and monitor product bioactivity, potency, and stability (Fig. 1).

A bioassay uses living material (animal, plant, tissue, or cells) to measure the biological activity of a substance. In biopharmaceutical drug development, bioassays are typically cell-based assays that are used to measure the bioactivity and potency of a biologic drug. In biologic drug development, it is challenging but critical to establish a functional bioassay that meets the essential quality attributes for measuring drug potency as described by regulatory guidelines.

Potency, a measurement of the strength of biological activity, is a functional measure of the tertiary/quaternary structure of a biologic drug as it relates to its therapeutic MOA. It is assessed in a bioassay by comparing the dose–response curve of the test material with that of a reference standard in a multiwell plate-based assay format [1]. This property is a critical parameter of drug product quality release testing and is also used to monitor drug stability, demonstrate product comparability after a manufacturing process change, and to assess lot-to-lot consistency during normal manufacturing operations. It is, therefore, critical that potency bioassays developed for use in biologics manufacturing and QC lot release reflect the MOA of the drug.

Traditional approaches to developing potency bioassays have relied on animal models and primary cells. These model systems provide biological relevancy for characterizing a drug's MOA, but they are challenging to implement in a quality-controlled manufacturing environment due to variability in the sourcing of primary cells, complex assay protocols, and limited availability of qualified reagents.

In recent years, bioluminescent reporter gene bioassays have been developed and validated for use in biologic drug manufacturing and QC lot release. Reporter gene bioassays can be designed to reproduce a biologic drug MOA while providing a measure of drug potency without the assay complexity and variability of more traditional model systems. This chapter provides an overview of key considerations in reporter gene bioassay design, clone selection, assay optimization, and qualification for manufacturing and QC lot release.

2 Assay Design

2.1 *Choosing a Cell Background*

The first consideration in designing a reporter gene bioassay is to choose an appropriate cell background. The cells must express biologically relevant signaling molecules and pathways to recapitulate the *in vivo* functional response targeted by the drug. Primary cells typically meet these criteria but are challenging to implement in higher-throughput and quality-controlled environments due to their variability and complex assay protocols. Many biologic drugs are designed to recognize therapeutic targets expressed on the cell surface, and, therefore, a cell line endogenously expressing the target receptor is a good option. For example, a HER2⁺ breast cancer cell line is a good option for measuring the potency of an anti-HER2 biologic drug. However, many immortalized cell lines endogenously express multiple receptors that can potentially activate the same signaling pathway. Therefore, it is important to demonstrate assay specificity by showing that the reporter gene response is dependent upon engagement of the target receptor.

Transformed cell lines are a less variable alternative to primary cells that are easier to develop into a quantitative and reproducible bioassay. If a cell line endogenously expressing a receptor target of interest is not available, or if the cell line shows a significant non-specific response, a biologically less relevant cell line genetically engineered to express a specific target receptor can be used. Regardless of which approach is taken, a bridging study is typically required to demonstrate that the reporter gene bioassay exhibits the engagement of the target upon the binding of the biologics and an equivalent response compared to a primary cell-based assay.

2.2 *Selecting a Genetic Reporter*

The Luciferase Assay System is an extremely sensitive and rapid reagent for the quantitation of firefly luciferase. Linear results are seen over at least eight orders of magnitude of enzyme concentration, and less than 10^{-20} moles of luciferase can be measured under optimal conditions. Therefore, luciferase reporters are widely used in cell-based assays due to their large dynamic range, homogeneity, and simple add-and-read assay format.

In addition, a biologically relevant promoter that exhibits rapid and robust activation in response to the biologic product with minimal nonspecific activation should be identified. A biologic product may be able to activate multiple intracellular signaling pathways, resulting in the activation of several transcription factors and promoters. In this situation, a promoter that is most directly coupled to the target pathway is preferred. For example, if a cytokine drug is designed to promote cell proliferation, a promoter that directly contributes to cell growth would be the most appropriate choice. Criteria for evaluating candidate promoters include activation kinetics, response fold induction, and relevant EC_{50} values.

2.3 Identifying a Positive Control

In order to make comparisons within and between bioassay optimization runs, a positive control biologic must be identified and used consistently throughout bioassay development. A biologic or drug product could be used as the positive control.

The positive control biologic should be stable at the recommended storage conditions (typically stored in aliquots at -80°C), and multiple lots should be tested to ensure lot-to-lot consistency. A positive control is not the same as a reference biologic, which is manufactured according to the same processes as a biologic product and used to determine the relative potency of a biologic in later-stage development and manufacturing QC lot release. However, if a reference biologic is available, it can be used as a positive control for bioassay development and during the later stages of bioassay validation, system suitability testing, and potency determination for lot release.

3 Assay Feasibility Studies

The goal of feasibility studies is to demonstrate proof-of-concept that the bioassay will perform as expected using the cell background and genetic reporter identified during the assay design phase. Thorough characterization and demonstration of the underlying biology of a bioassay is essential to avoid unnecessary time and cost spent in the later phases of bioassay optimization and qualification.

Feasibility studies using genetically engineered cell lines can be performed by transient transfection or through the creation of a stable cell pool. Transient transfection is a good option if the cells can be easily transfected with high efficiency and the target receptor is expressed at a relatively high level. Factors such as cell background, expression construct, transfection method, and underlying biology may all impact whether feasibility studies can be accomplished using transiently transfected cells. In some instances, a stable cell pool will need to be established using antibiotic selection. When using a new cell line to create a stable cell pool, it is important to generate an antibiotic kill curve to determine the optimal

concentration of antibiotic to use for selection. A concentration of antibiotic strong enough to kill the nontransfected cells but not too strong to kill all the cells should be used. Antibiotic selection typically requires 2–5 weeks to complete depending on the cell background and antibiotic selection marker.

Initial functional studies should be performed to generate a 10-point dose–response curve of a positive control biologic. At this early stage of proof-of-concept study, a perfect curve with 2–3 points at upper asymptote and 2–3 points at lower asymptote is not required. However, a lack of dose-dependent response may be the result of low receptor expression or a nonfunctional genetic reporter. The generation of a stable cell pool does not necessarily equate to enough receptor expression for bioassay development, and, therefore, receptor expression should be measured directly by flow cytometry or other methods. If receptor expression is determined to be low, the cells can be further sorted to obtain a population of higher-expressing cells. Importantly, while high target-receptor expression may be desirable for some bioassay designs (e.g., to measure the activity of a soluble ligand or antibody), other bioassays that require complex interactions between multiple cell types and receptor–ligand pairs may benefit from lower receptor expression. In these cases, it is recommended to sort multiple populations of cells with varying receptor expression levels for functional testing. Finally, if no dose–response curve is observed and relatively high target–receptor expression is demonstrated, it is possible that the genetic reporter is nonfunctional. To assess this possibility, soluble compounds that nonspecifically activate promoter elements can be tested.

As noted above, some bioassay designs that use multiple cell types and receptor–ligand pairs may benefit from lower receptor-level expression. High target-receptor expression can lead to increased basal activity and a reduced assay window due to the dynamic equilibrium between the active and inactive forms of the receptor. Therefore, while target–receptor expression can be informative to interpret feasibility data, the functional assay response must be the primary criterion for decision making.

In summary, the goal of feasibility studies is to demonstrate that the bioassay cell line can yield a dose-dependent response using a positive control biologic. Even a less than twofold positive response is typically sufficient to move toward further development of the bioassay. If a dose–response curve cannot be generated using a positive control biologic, even when the target–receptor expression and reporter function are confirmed, alternative assay design strategies (e.g., alternative cell line or promoter) should be considered.

4 Cell Line Generation and Clone Stability Testing

4.1 *Selecting Cell Clones*

The cell line developed for use as a functional bioassay should consist of a single cell clone. This ensures stable integration of genetic elements, reduces genetic drift or loss of the engineered content, and results in more reproducible assay performance over time.

Single cell clones are typically generated by limiting dilution where a suspension of cells is diluted and dispensed into 96-well (or higher-throughput) plates such that each well contains an average of at most one cell per well. The cells are then cultured and expanded resulting in a population of cells derived from a single cell clone. Limiting dilution is relatively easy to perform but does result in wells with either no cells or more than one cell. Wells with more than one cell may not be genetically identical and may result in the generation of an unstable population. Multiclonal wells of adherent cells can easily be identified by visual inspection, but suspended cells remain a challenge.

If ectopic expression of the target receptor is used, another approach used to generate single cell clones involves labeling the target of interest with fluorescently labeled antibody and using fluorescence-activated cell sorting (FACS) to sort the labeled cell population into a 96-well plate at a density of one cell per well. This technique requires a cell-labeling step and subjects the cells to high pressure during sorting, which can be overly stressful for some cell types when culturing from a single cell. Thus, not all wells containing single cells expand into clonal populations. Success rates will vary depending on the growth characteristics of the cell line and overall health of the cells prior to sorting. If antibody labeling is performed as part of the FACS protocol, it is good practice to first sort a pool of cells based on the antibody label, let the cells recover in culture, and then blind sort (e.g., without antibody labeling) single cell clones into plates. This approach limits cell handling on the day of single-cell cloning and increases the success rate of clonal cell expansion.

After limited dilution cloning (or FACS to sort single cells), the cells are cultured in multiwell plates (typically 96 wells) until the cell population has expanded to at least 20–30% confluent to allow initial functional screening. During the initial phase of cell culture and expansion, it is typical to see a wide range of cell growth rates between individual clones. Some clones may grow as well as their parental cell line while others may stop growing.

An initial screen can be performed by replicating individual wells into parallel plates and measuring the functional response with or without one single concentration of a positive control biologic. Functional responses can be categorized according to whether they show a low, medium, or high response.

Luminescence observed during the initial screen can vary dramatically among different cell clones. Clones showing high basal relative luminescent unit (RLU) signal and producing a high-fold induction in response to a positive control biologic are the best candidates for further bioassay development. Clones showing higher basal RLU will result in better assay performance as measured by percent coefficient of variations (% CV). In general, basal RLU should be at least 50- to 100-fold higher than instrument background (wells with medium only). However, if low basal RLU is generally observed in the initial screen, clones with higher-fold induction will be chosen for further assay optimization to increase the basal RLU.

When clones are further expanded, full dose–response curves can be tested on a subset of clones. When comparing clone functional responses, cells should be seeded at the same cell density so that their yields are comparable on the day of harvest for the assay. Cells should be plated at the same number of cells per well so that comparisons can be made. Full-dose titration of the ligand or biological product will give insight into the EC_{50} , maximum fold induction, shape of the response curve (hillslope and curve fit parameters), and overall luminescence intensity. Even at this early stage of bioassay development, some assays respond strongly using normal serum concentrations whereas others benefit from low percentage or alternative sera (e.g., charcoal-stripped serum).

Multiple criteria should be taken into consideration when selecting a cell clone for bioassay development. For reporter gene bioassays, luminescence fold-induction in response to a positive control biologic is a key attribute. However, other parameters such as signal background, peak signal intensity, and induction time should also be considered.

If a bioassay is developed for use with a blocking antibody to neutralize an agonist, the agonist concentration used for potency measurement of the blocking antibody and slope response of the agonist dose–response curve will impact the EC_{50} of the test blocking antibody. Higher concentrations of agonist will result in a higher EC_{50} of the blocking antibody. It is important to mimic a historically acceptable EC_{50} range when comparing clone responses.

If ectopic receptors need to be added to the parent cell, flow cytometry can be used to determine relative expression levels of different clones, correlate response to expression level, and even identify clones that most likely are not truly clonal. As shown in Fig. 2, double peaks and broad peaks with a “shoulder” of fluorescence intensity often indicate a mixed culture of clones. While this “clone” may temporarily meet the bioassay needs, it may ultimately fail when undergoing passage stability.

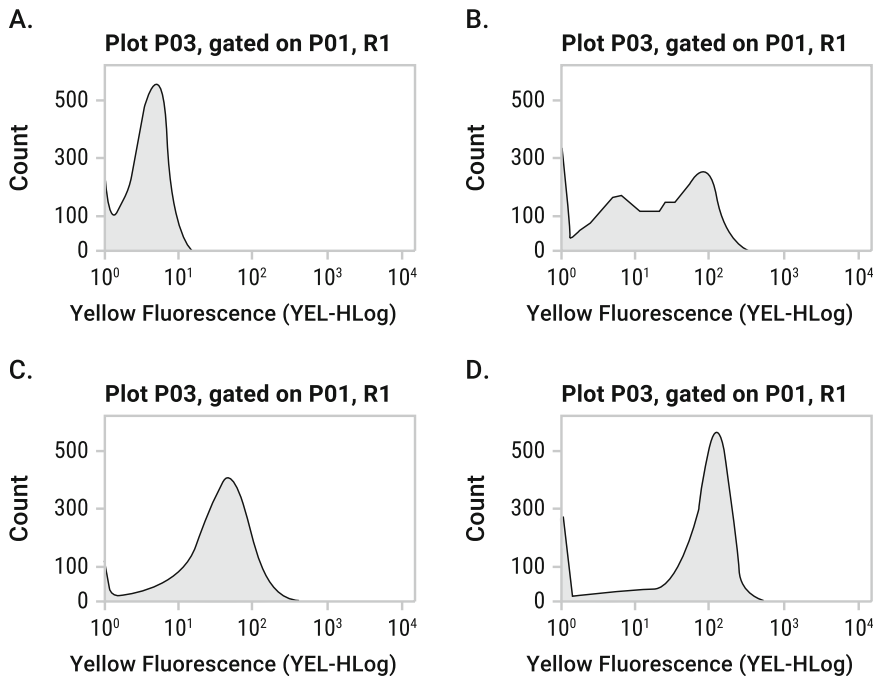


Fig. 2 Histograms of three different clones following FACS staining. (a) Isotype control. (b) Double peak indicating mixed culture. (c, d) Two clones demonstrating slightly different levels of expression

4.2 Performing Clone Stability Testing

“Cell passage” refers to the number of cell population doublings, which accounts for variable growth rates and is not affected by the number of times the cells are passaged in a week. It is important to establish the length of passage stability and acceptable level of loss or change, which will depend on the individual requirements of a clone. Long-term clonal stability ensures consistent cellular functional responses throughout passages in terms of EC_{50} response, fold induction, overall luminescence, cell growth, and receptor expression during a defined amount of time in culture. Functional instability can manifest as a decrease in luminescence while maintaining fold induction or as a decrease in fold induction (Fig. 3). At a minimum, cell clones must be stable enough to sequentially prepare seed stocks, master cell bank, and working cell bank.

Stability studies can begin after initial functional screening and selection of a limited number of clones that exhibit assay specificity and good assay response. Cells are maintained under full antibiotic selection pressure as indicated during preliminary kill curve tests with the parental cell line. Cell culture medium with freshly supplemented antibiotics from reliable suppliers will help individual clones maintain their original characteristics and minimize genetic drift. Sufficient banks should be made for each testing clones at passages as early as possible to serve as a source for subsequent seed stock preparation or “backup” cells.

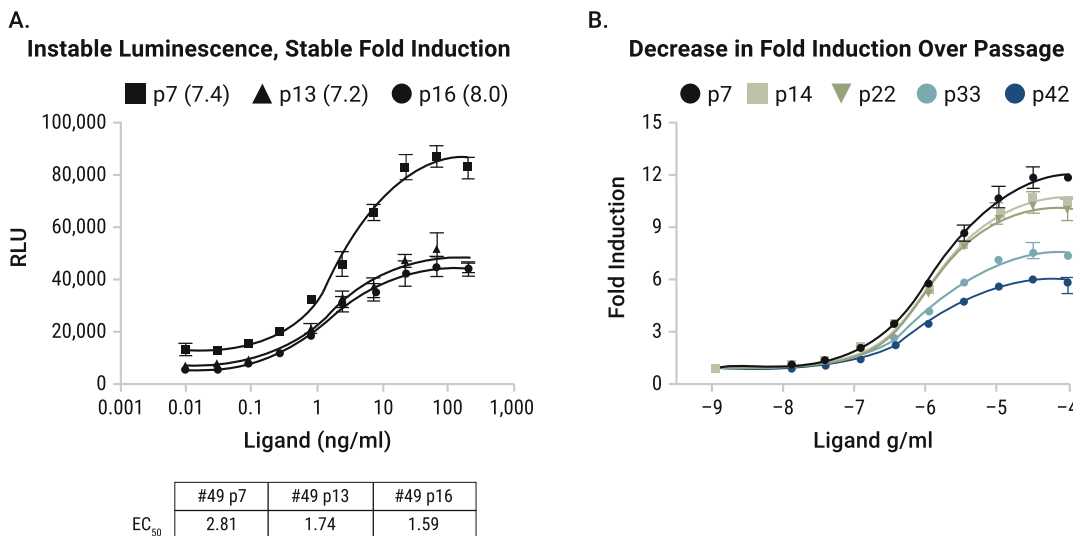


Fig. 3 Examples of clonal instability. **(a)** Decrease in luminescence across passages 7–16 (stable fold induction noted). **(b)** Steady fold induction decrease across passages 7–42

Several strategies can be used to prepare a candidate cell clone for stability testing. Cells for each clone representing one or more cell passages can be maintained in parallel over time as active cell cultures. Cells will be harvested and frozen every 5–6 passages to generate a series of staggered passages. Cells frozen at different passages will then be thawed and grown in parallel for at least 2 weeks, followed by side-by-side functional measurement of basal luminescence (measure of reporter gene stability), signal fold induction (measure of receptor expression), and EC₅₀. Comparing functional response data from different days is not a preferred approach as many variables in assay and cell culture conditions will contribute to the assay results and complicate data interpretation. To confirm functional results, additional tests such as flow cytometry analysis for receptor expression are recommended using the same staggered passage cultures. Flow cytometry and functional response can be used in tandem to demonstrate a stable clone (Fig. 4). Other cell culture characteristics, such as cell population doubling time and cell morphology, can also be noted during this expansion duration and contributed to final clone selection.

5 Assay Optimization

Optimization of bioassays is important to ensure the best possible sensitivity, signal-to-noise ratio, and assay window. The bioassay must also be reproducible, ensuring that small day-to-day variations do not significantly impact the results. Several critical factors during bioassay optimization are considered below.

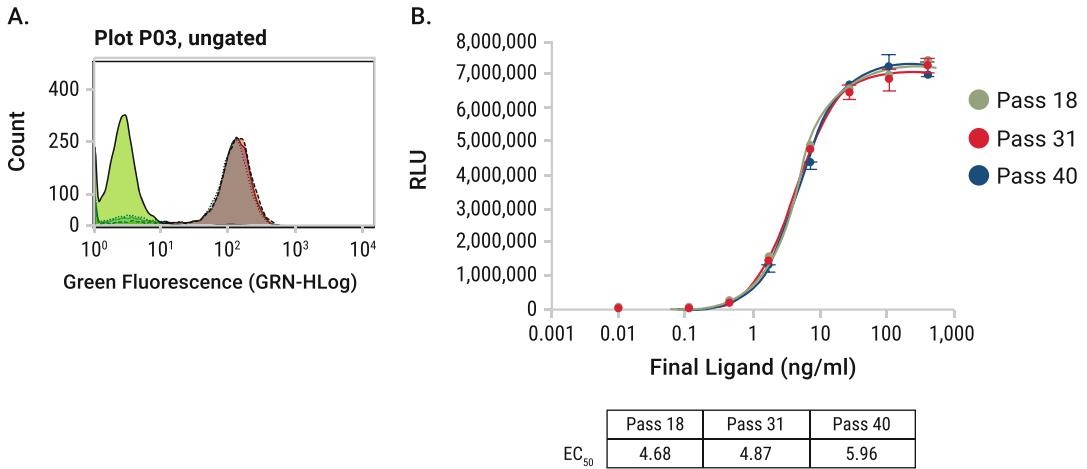


Fig. 4 Stable functional response confirmed with surface expression. **(a)** Cells from passages 18 to 40 were analyzed for receptor expression by flow cytometry; isotype control included. **(b)** Bioassay using cells from passages 18 to 40

5.1 Standardizing Assay Reagents

Quality cell culture and assay reagents are critical for consistent bioassay performance. It is prudent to identify reliable suppliers of quality sera, media, and cell culture supplements and establish assay media formulation during bioassay development. Sera, in particular, can have significant impact on cell growth and assay performance depending on grade, region of origin, and supplier. Always anticipate supply constraints and test multiple media and sera suppliers to ensure consistent performance of the bioassay.

5.2 Culturing Cells

Passaging of cells according to a defined schedule is integral to consistent assay performance. For most bioassay cell lines, a Monday–Wednesday–Friday–Monday schedule is recommended. Optimization of culture conditions often results in consistent growth rate (doubling time) with high cell viability at the time of harvest.

When expanding cells, seed them into flasks and measure the seeding density. Seed suspension cells at a defined cell density (number of cells per mL) in a set volume of medium per flask size. For adherent cells, use cell numbers per cm² for calculating cell seeding density in each flask. Volume differences between flasks can affect gas exchange and cell performance. Therefore, a standardized media volume per flask (e.g., 20 mL per T75 flask) is recommended. Cell seeding and harvesting densities are best tested empirically for effects on cell performance. Seeding cells too densely results in overgrown cultures that typically are detrimental to assay performance. On the other hand, seeding cells too sparsely can result in increased costs. Test a range of seeding densities during a standardized 2- or 3-day passage to determine the best range so that variations in cell seeding density (e.g., inaccurate cell

count at seeding) have as little impact on reproducibility of the bioassay as feasible. Record cell harvesting density, doubling rate, and viability at the time of harvest as integral data for the bioassay.

5.3 Plating Cells

While optimizing bioassay conditions, it is important to test a range of cell numbers per well in order to determine how signals are detected. Since instrument detection sensitivity can vary, it is important to consider the signal-to-background ratio to determine the level of signal that can be obtained using a positive control above the background noise of the machine. Background noise is indicative of the lowest sensitivity that can be determined for the bioassay. If the bioassay includes a coculture of multiple cell types (e.g., effector and target cells), cell–cell interaction and cell ratio can have a significant impact on the assay window and sensitivity. Test a range of cell ratios to ensure optimal performance of a bioassay with cell cocultures. For adherent–suspension cell interactions, adherent cells can be plated overnight to reach 80–100% confluency the next day and interact with 50,000 suspension cells per well in a 96-well plate. For suspension–suspension cell interactions, a 1:1 ratio (e.g., 50,000 cells for each cell type) in each well of a 96-well plate will be a good starting point.

5.4 Induction Time

Assay incubation time can also have a significant impact on reporter-based bioassay performance. Induction time should be optimized by performing a time course experiment at 4, 5, and 6 h, or overnight (18–24 h) to determine the maximal performance and assay robustness. When harvesting and staging the bioassay with adherent cells, significant response differences may be observed when using traditional trypsin or weaker enzymatic alternatives as surface proteins and receptors can be temporarily damaged. Certain bioassay cells may need an overnight “recovery” after plating to achieve an optimal and consistent response. A solid understanding of the nuances of a cell clone will greatly improve the odds of a successful bioassay or its transfer to another facility.

5.5 Assay Buffer and Media

For cell culture media, a common practice is to use a general, well-established medium composition according to the cell origin document and recommendations from commercial cell line providers (ATCC, DSMZ) and supplement it with fetal bovine serum (FBS) (e.g., normal, heat-inactivated, or gamma-irradiated) and together with nutrients that support cell growth. Since FBS is a natural product without clearly defined composition, test sera from different commercial sources, and even different lots from the same commercial source, are used to confirm that they do not cause significant assay variation. The optimal percentage of FBS is typically between 0.5% and 10%. In some situations, components of FBS can interfere with the bioassay, in which case the use of dialyzed FBS should be considered. Being able to lower FBS

concentration in assay buffer without impacting cell health, luminescence signal, and assay robustness will help to minimize variations introduced by different FBS lots.

5.6 Plate Edge Effects

Plate edge effect is an inconvenient phenomenon caused by more media evaporation along the edges of the microplate during incubation. The edge effect can cause many problems, including varying volumes and concentrations, which can alter cell viability and assays results. In general, only the inner 60 wells of a 96-well plate are used for most bioassay applications to avoid plate edge effects on data. If a bioassay has a short induction time (e.g., less than 5–6 h) plate edge effects are of less concern. To test for edge effects, generate a plate “heat map” where each well gets the same treatment (typically the EC_{50} of the test biologic). Luminescence across the plate will indicate whether there are significant position-dependent effects.

5.7 Hook Effect

Occasionally, upon careful inspection of the fitted data, a “hook effect” may be noticed at the highest concentration(s) tested that may impact the curve fit. To troubleshoot, confirm the stock biologic does not contain a toxic component (such as sodium azide, detergent, or stabilizer) which may carry over to the first dilution (highest concentration) sample tested, thereby impacting the cell health and decreasing luminescence. Sometimes antibodies are prone to hook effects when used at high concentrations. Preparing a series of tightly spaced data points across a narrow concentration range can shed insight as to where along the concentration range the hook appears. Altering the starting dilution concentration, with anticipation of the highest potency desired, will typically resolve the problem.

5.8 Design of Experiments (DoE)

Assay optimization can be performed by assessing one parameter at a time. However, interplay between assay parameters can significantly impact the robustness and repeatability of a bioassay. Therefore, multifactorial analysis performed by varying multiple experimental parameters in a single experiment is recommended for assay optimization. Design of experiment (DoE) is a commonly applied method of multifactorial analysis that is used to define the relationships between assay parameters that impact assay output. It is extensively used for the implementation of Quality by Design in both research and industrial settings. DoE can be simple or complex, depending on the application and number of critical assay parameters being evaluated. Figure 5 shows an example of a simple DoE experiment used to demonstrate the robustness of a bioassay previously optimized one parameter at a time.

In this example, five assay parameters (two conditions per parameter, one being the final optimized condition) were tested in a single experiment (Fig. 5a). Statistical analysis software (JMP)

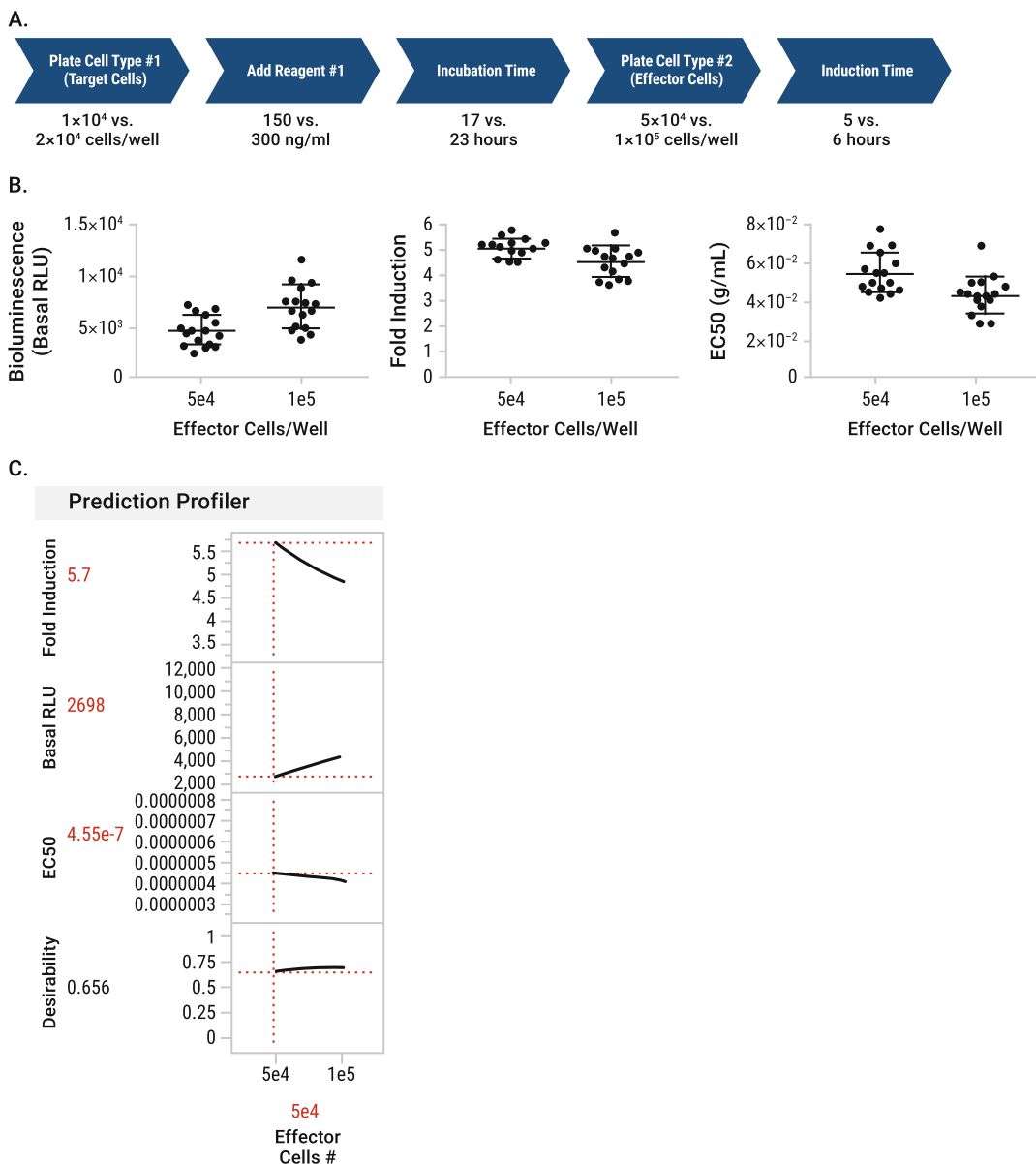


Fig. 5 Example DoE design and data analysis. (a) Five critical assay parameters were tested in a single DoE experiment. (b) Aggregate data of differences in relative luminescence (RLU), assay window (fold induction), and EC₅₀ of a control biologic when the number of effector cells is varied. Bars show mean ± standard deviation. (c) Example of data analysis using JMP software. Shown are the effects of effector cell number on fold induction, RLU, and EC₅₀. The desirability index is shown where fold induction and RLU are equally weighted

was used for the assay setup design and data analysis (Fig. 5c). The assay was successful as it demonstrated a positive response to the biologic in all cases (Fig. 5b), but some assay parameters, such as cell number, had bigger effects on the assay readout than others

(e.g., plating time). Again, this is only a simple DoE used for a specific purpose. A more detailed explanation of how to perform DoE is outside of the scope of this chapter. Consult a biostatistician to assist in DoE use for bioassay optimization.

In summary, optimization of bioassay conditions includes many critical steps to achieve consistent bioassay performance. The list presented in this chapter is not exhaustive, and other important factors may need to be addressed and optimized depending on the bioassay system.

6 Development of Thaw-and-Use Cells

While many bioassays use a Master Cell Bank (MCB) and Working Cell Bank (WCB) for continuous production, thaw-and-use (T&U) cells (also called ready-to-use or assay-ready cells) can be used for a bioassay straight out of the vial [2, 3]. These cells offer convenience and reproducibility and are rapidly being adopted for cell-based bioassays. Through manufacturing process development and optimization, T&U cells can be produced in large quantities with each batch harvested from the same working cell bank at defined scale-up conditions. Assay variations caused by daily cell culture are eliminated through a controlled process. Bulk production and storage of cell banks save on labor and time, reduce long-term development costs, and facilitate assay transfer between different laboratory sites. The simple and homogeneous nature of the reporter bioassay format goes hand-in-hand with more easily adoptable T&U cells.

This section will discuss general considerations in producing T&U cells for reporter bioassays. In particular, critical factors affecting cell culture, cell freezing, and long-term storage in the context of functional testing parameters will be examined.

Cell culture conditions prior to harvesting for T&U cells are critical for bioassay performance. Different culture vessels can impact cell quality and assay performance by affecting cell growth characteristics (cell growth rate and viability) and protein expression level (Fig. 6). T-flasks are standard vessels during the early stage of bioassay development; for large-scale cell production, many choices of culture vessels are available to produce several liters of culture. Testing a variety of culture vessels, such as flasks, spinners, roller bottles or bioreactors, is important for developing a process for scale-up. These vessels should be tested and chosen depending on the cell types (adherent or suspension), desired batch sizes, and other practical factors such as ease of handling, yield, and production cost. During cell expansion, cell seeding density, culture volume per vessel, and cell passing schedule need to be standardized to ensure batch-to-batch consistency. The cell doubling rate and viability should be recorded at each passing to

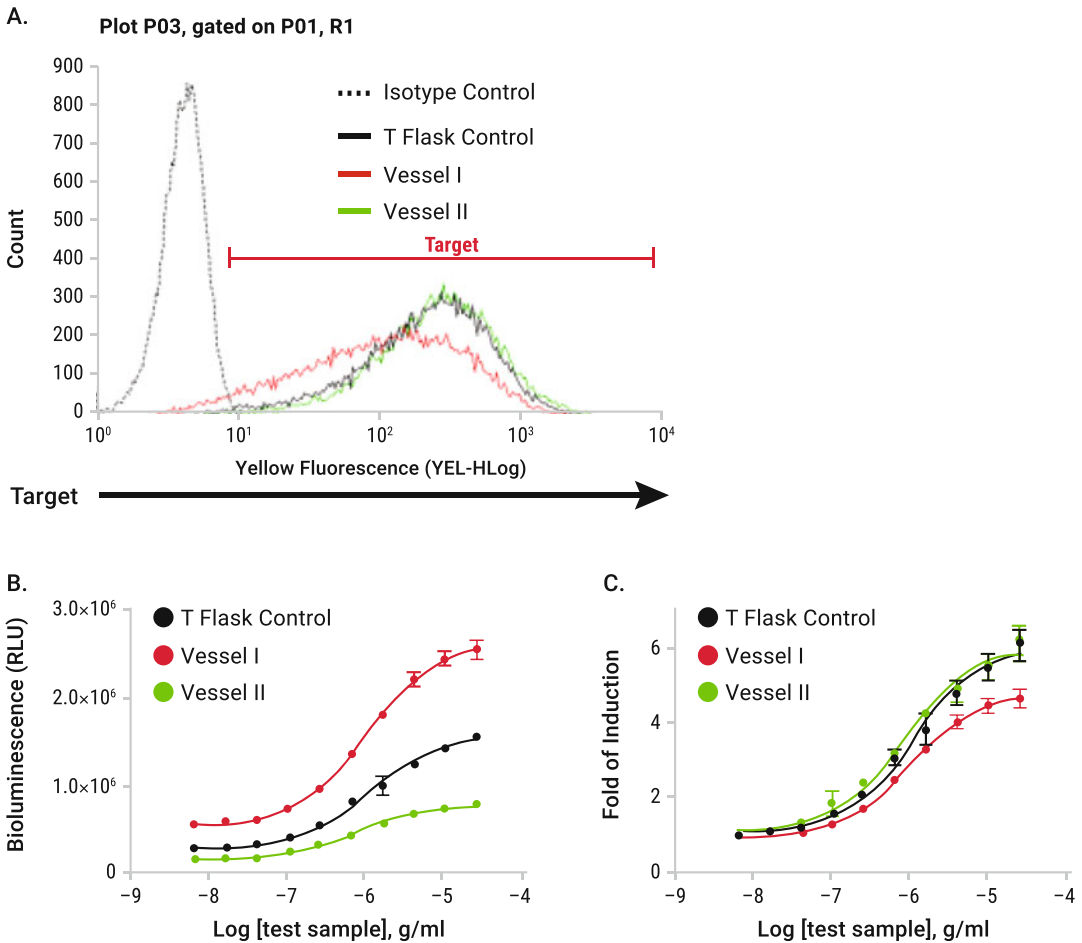


Fig. 6 Example of the selection of culture vessels impacting protein expression and assay performance. (a) FACS analysis showing different expression level for a cell surface target of a biologic product, from the cells produced by T flask, culture vessel I, and vessel II. (b, c) Functional bioassay showing different assay performance from the cells produced by T flask, vessel I, and vessel II. The cells produced from culture vessel I showed lower target expression level (a), higher basal luciferase activity and RLU in the absence of drug stimulation, (b) and lower drug-induced assay window (c) when compared with the cells from T flask control and vessel II

monitor cell growth and help troubleshoot in the event of unexpectedly low production yield and undesired assay performance.

Cell-freezing conditions are another critical factor impacting T&U cell performance. Freezing conditions for T&U cells can be very different from WCB and MCB because T&U cells will be directly seeded into microplates for biological potency testing. Cell-freezing media can be optimized for the concentrations of FBS and cryoprotectant or choices of cryoprotectants. Many serum-free freezing medium products are commercially available but need to be carefully evaluated to make sure they do not impact cell performance. At the time of cell harvest, freezing medium

should be precooled to 4 °C before adding to the cell pellet, and the cell suspension should be kept on ice while cells are dispensed into cryovials. The cell-dispensing time course, or DMSO tolerance time, should be carefully studied to minimize the loss of cell viability at the end of the dispensing. For the cell-freezing process, a controlled, programmable electronic freezing unit (e.g., CryoMed Freezer) with the ability to rigorously maintain the rate of cooling is highly recommended to ensure consistent viability and assay performance for the whole batch production, which may contain hundreds to a thousand vials. Manufacturers of controlled freezers typically recommend cell type-specific freezing programs though further modifications may still be necessary after the cell lines are engineered with exogenous protein expression to improve the consistency of cell quality and functionality. Unlike the majority of tumor cell lines, which are very sensitive to freezing conditions, certain common cell lines such as CHO-K1 and HEK293 are not overly sensitive to the cell-freezing protocol. Other freezing chambers (Mr. Frosty, Styrofoam box, or CoolCell) can also be used to produce satisfactory results. However, they are not recommended in a manufacturing setting for producing large quantities of T&U Cells due to the possibility of introducing large intrabatch variation.

Once the T&U cell batches are made, they should be transferred and stored in the vapor phase of liquid nitrogen or a –140 °C freezer. Functional performance was compared in multiple cell lines stored at –80 °C or in liquid nitrogen. The data showed that some cell lines, such as CHO-K1, can maintain good cell viability after 2–3 weeks in –80 °C freezers, although the viability might still change during longer time periods. Other cell lines, such as Jurkat T cells, lost 20% cell viability during the same time duration (Table 1). In general, we highly recommend storing the T&U cells at liquid nitrogen temperature to maintain consistent assay performance for the long term.

Table 1
Change of cell viability for thaw-and-use cells upon short-term storage at –80 °C

	Days	Postthaw cell viability			
		Jurkat, %	Raji, %	CHO, %	HEK293, %
Stored at –80 °C	0	86	91	99	82
	3	85	91	97	89
	7	77	92	98	89
	10	72	91	95	87
	14	69	91	94	87
	17	65	91	93	85

T&U cells produced with Jurkat, Raji, CHO-K1, or HEK293 cell lines were transferred from liquid nitrogen storage to –80 °C freezers and stored for the time period as indicated, up to 17 days. The cells were transferred back to liquid nitrogen storage after each time point until all test conditions were complete. The cells were then thawed, and postthaw viability was measured by trypan blue staining

7 Bioassay Prequalification

At this point, a bioassay must be further qualified for potency testing of drug products by demonstrating the accuracy, precision, linearity, and range following a series of repeated assays. Classic prequalification outcomes and metrics include establishing that the reporter bioassays can be performed using cells from continuous cell culture or in cryopreserved T&U format. If cells are used from continuous cell culture, their staging is critical for a consistent assay response. The previous section on assay optimization details some of the factors affecting cell functional responses.

For relative potency testing, T&U cells provide an important advantage over continuous culture cells as they repeatedly yield a consistent response across the batch of cells. This consistent response makes prequalification development far simpler as it eliminates the variability often associated with staging continuous culture cells. Large numbers of T&U cell vials can be prepared and stored below $-140\text{ }^{\circ}\text{C}$, ensuring consistent responses for many years.

Accuracy and precision among replicates are important for any bioassay, especially for successful potency testing. Replicate precision will allow for better discernment of small potency differences and overall absolute reproducibility (e.g., EC_{50} and response to a standard concentration range during the test). Poor replicates with $\%CV > 5\%$ can be an indication of some protocol or assay variable that was not fully optimized. Following good pipetting practices, including the choice and condition of equipment, can play a major role in bioassay success. It is undeniable that, for cell dispensing and serial dilution purposes, quality data depend on pipettes in good operating shape and without any biases they may introduce during dispensing. Choose the proper pipette for the working volume needed and ensure all channels of a multichannel pipette are in good working order without internal piston leakage. For dispensing purposes, electronic pipettes provide rapid and accurate dispensing, and facilitate larger and complicated experiments.

Begin by establishing and understanding the ligand or biologic product response as previously discussed in the section on assay optimization. Protocol optimization to generate a consistent EC_{50} is important as this impacts the EC_{80} and EC_{20} results. Agonist EC_{20} and EC_{80} are especially critical if used in blocking bioassays. The appropriate agonist concentration at either EC_{80} or EC_{20} depends on how big the assay window is. An agonist EC_{80} will be chosen if the assay has a fold induction less than a few hundred. When fold induction reaches over a thousand for certain bioassays, the assay will suffer from “over-sensitivity” and the potency assay will tend to fail during statistical analysis due to the large variation in RLU at the upper asymptote. In this case, the EC_{20} will be

selected instead to reduce assay variation while still maintaining assay robustness.

Establishing the dilution series of the ligand or effector molecule is an important first step for potency assay development and prequalification. Identifying any potential biases or location effects across the entire plate, including rows, edges, and corners, will dictate which wells can be reliably used for samples. Often, the peripheral outside rows and columns of a 96-well plate can be prone to bias, but this issue should be empirically tested and not assumed. With 96-well plates, using the inner 60 wells for samples is a safe starting point until the outer rows and columns can be demonstrated to lack bias. Often, plate location bias can be remedied by changes in the protocol, such as preplating cells for a minimum of 1 h prior to sample addition. With the identification of which wells can be used for samples, the number of points for each dilution series can be established.

Sample placement and location throughout the plate is dependent on the number of series and data points for each series and any additional controls that might be used as part of assay acceptance. Place samples throughout the useable area of the plate in a non-clustered fashion to minimize any unintended response bias, including any potential luminometer plate reader bias. Alternating rows or columns is usually sufficient.

Potency assays typically encompass full dose responses of the test, reference, and control biologic samples. It is imperative to choose a dilution range such that a full response is created. The responses are then fitted using curve-fitting software, often as a four-parameter curve. Any potency difference is noted as a left or right shift across the x -axis if the curves are determined to be parallel either by an F -test or equivalency test. Each fitted curve should have an adequate number of data points to accurately establish the upper and lower asymptotes and as many points as possible on the linear range of the response curve containing the inflection point. Generally, the starting concentration of the test sample is serially diluted (e.g., three- to fivefold) before adding it to the cells, but other dilution schemes are possible. As shown in Fig. 7, the dilution series chosen will impact where the data points end up being located across the bioassay response. The dilution factor, starting concentration of sample, and number of data points in the series can all be manipulated to achieve a full response for the samples. Ultimately, this sample concentration range will need to accommodate samples with a potential change in potency with a response shift relative to the reference. For the purposes of assay prequalification, potency samples can be prepared by intentional dilution to create mock potency samples representing 50%, 70%, 140%, and 200% of the reference. After a series of repeated tests, this potency range and the resulting recoveries are used to establish the linearity response of the potency assay.

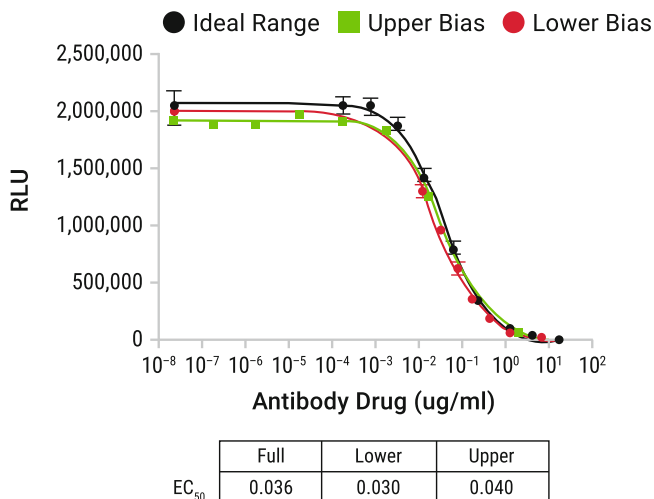


Fig. 7 Dilution series of an agonist-blocking antibody can impact curve fit. Agonist blocking by an antibody drug was demonstrated, starting at 20 $\mu\text{g/ml}$, using a series of data points across the entire response range (1:4.25-fold in black), one series with a bias of points at the lower asymptote (1:2.5-fold in red), and a third series (1:10-fold in green) with a bias at the upper asymptote

Special considerations should be noted for designing blocking bioassays, where a single concentration of ligand, typically the EC₈₀ response, is inhibited by a titration of blocking antibody. It is important to understand if the antibody reacts to a surface receptor on the reporter cell or the ligand itself. Protocol adjustment should reflect the antibody’s target: for a surface receptor; the antibody sample may need to be added first and preincubated with cells prior to addition of the ligand. If the antibody targets the ligand itself, the antibody and ligand should be coincubated for some amount of time before addition to the reporter cell as a sample. To create a robust protocol, this preincubation time should be empirically determined with a time course experiment as shown in Fig. 8.

8 Qualifying Potency Bioassays

Once the assay optimization and prequalification phase is completed, the assay moves into the qualification phase. During this phase, the assay design is confirmed to be capable of generating reproducible results for the specified purpose. Therefore, for a reporter gene bioassay, the assay qualification can be defined as a set of experiments performed under defined assay conditions, aimed to demonstrate that the method is capable of reliably measuring the relative potency (RP) of the drug under investigation. For best practices, the assay qualification should be conducted following a preapproved test protocol generated during the

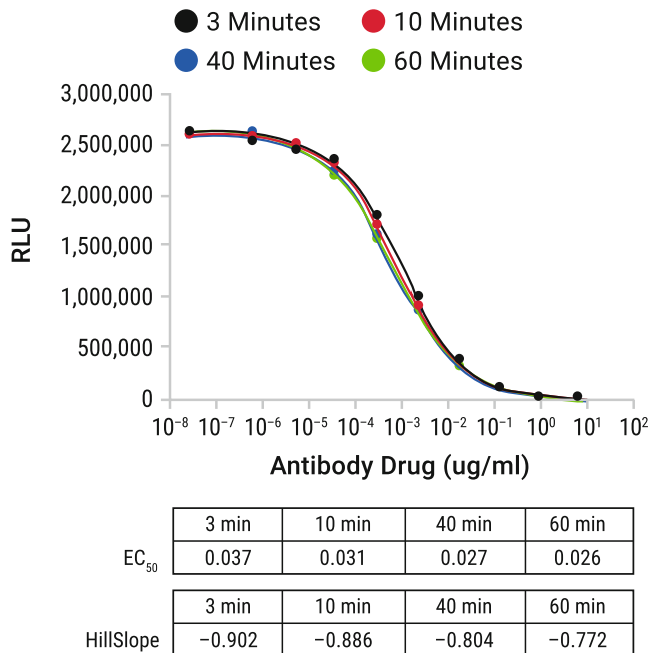


Fig. 8 Time course of antibody and ligand. A time course for the preincubation of an agonist with antibody drug demonstrates a response bias at short incubations where equilibrium has not been reached between the antibody and its agonist

prequalification phase that outlines the assay conditions, plate layout, number of sample replicates, and experimental design for each parameter investigated.

The following parameters are typically assessed during a reporter bioassay qualification: sample and assay suitability criteria, specificity, repeatability, intermediate precision, accuracy, dilution linearity, and range.

At the completion of the assay qualification, a report is generated that summarizes results, analysis, and conclusions on the assay achieving its intended purpose.

8.1 Bioassay Method and System Suitability Acceptance Criteria

A defined assay procedure is generated at the conclusion of the assay optimization/prequalification and prior to the assay qualification. A written method that includes the drug serial dilution, plate layout, number of samples and standard replicates on each plate, reagent concentrations, and incubation time is generated and made available to all the scientists performing the assay qualification. Typically, for a reporter gene bioassay, the plate layout will include multiple replicates of the drug sample and the drug standard, which are then used to calculate the relative potency of the drug sample. The relative potency of the drug is typically calculated as the EC₅₀ ratio of the sample and standard curves [4, 5].

In addition to the assay conditions, the method should include a series of system suitability and sample acceptance criteria. Each plate should be evaluated against these criteria prior to proceeding with evaluating the qualification parameters. Doing so will ensure that only valid assays are included in the qualification analysis. The system and sample acceptance criteria should be derived from the assay prequalification data and usually will include a measure of how well each standard and sample curve fits, a measure of the replicates agreement, a minimum signal-to-noise ratio (or A/D ratio) for each sample and standard curve, and a measure of parallelism between the standard and the sample curve. Parallelism between the sample and standard curve can be assessed by several statistical analysis models [4, 5].

During assay qualification, the relative potency of the sample is calculated and reported for each plate that meets all the assay and sample acceptance criteria. For each plate that does not meet the assay and sample acceptance criteria, the plate is invalidated and repeated. If a high incidence of failed assay/samples acceptance criteria is observed during the assay qualification, steps should be taken to investigate the failed results and address the cause. If necessary, additional assay optimization or prequalification experiments may be conducted to ensure the consistency of the assay performance across different days and as performed by different scientists.

8.1.1 Qualification Parameters

The following parameters are typically evaluated for a reporter bioassay:

1. Specificity.
2. Precision.
 - (a) Repeatability.
 - (b) Intermediate precision.
3. Relative accuracy.
4. Dilutional linearity.
5. Range.

A definition of each parameter for general analytical procedure can be found in the ICH Harmonized Tripartite Guideline [6, 7].

Specificity is the ability of the bioassay method to specifically detect the potency of a drug. The bioassay should be specific to the receptor or signal pathway of the drug under investigation. Specificity can be tested using formulation buffer prepared as a sample to ensure the noninterference of the buffer with the bioassay. Additionally, the specificity can be tested using a sample drug not specific to the targeted pathway. In both cases, no dose–response curve should be observed when compared to the standard curve tested on the same plate.

Precision of a method expresses the closeness of agreement between a series of measurements obtained from multiple testing of the same sample. Precision of the bioassay is considered at two levels: repeatability and intermediate precision. Repeatability (also defined as intraassay precision) expresses the precision under the same operating conditions over a short interval of time. The precision of the method is expressed as the %CV of the series of measurements.

In a reporter gene bioassay, repeatability can be tested over one assay setup by one analyst. Sample and standard preparations are loaded into multiple assays plates in one experimental setup according to the method. If available, the drug product can be used as sample; alternatively, reference standard material can be used to prepare a “mock” sample. The sample and standard dilutions should be prepared independently. The number of plates included in the single setup will depend on the complexity of the method (4–6 plates are recommended). Each assay plate will generate a single reportable percent relative potency value and the intraassay precision will be reported as the %CV or % geometric coefficient of variation (%GCV) of all the reportable percent relative potency values.

Intermediate precision (also defined as interassay precision) expresses within-laboratory variations such as different days, different analysts, different equipment, and different lots of bioassay reagents. In a reporter bioassay, the intermediate precision experiments are performed by at least two analysts in multiple independent assay setups and on different days. Similar to repeatability, the drug product or reference standard can be used as a sample. Each analyst will prepare all assay reagents and samples independently and will use a different luminometer to analyze the plates. A different lot of thawed cells can be used by Analyst 2 in this portion of the qualification. Each assay plate (8–12 plates total is recommended) will generate a single reportable percent relative potency value and the interassay precision will be reported as the %CV or %GCV of all the reportable percent relative potency values.

The accuracy of an analytical procedure expresses the closeness of agreement between an accepted reference value and the value found. The accuracy of the reporter bioassay can be measured using a reference standard to prepare samples at different concentrations relative to the nominal drug concentration. Typically, samples at 50%, 75%, 100%, 125%, and 150% are prepared and tested against the reference standard material prepared at 100% of the nominal drug concentration. It is also not uncommon to dilute samples at 50%, 70%, 140%, and 200%. The individual and mean relative potency values for each sample concentration are reported. The individual and mean % biases (from 100% nominal value) are also calculated and reported.

The linearity of an analytical procedure is its ability (within a given range) to obtain test results that are directly proportional to the concentration of the sample. For the reporter bioassay, the linearity is calculated by plotting the experimental log relative potency values for each concentration versus the theoretical log relative potency values on a linear scale and by performing a linear regression analysis. The coefficient of determination (R^2), slope, and y -intercept from the linear regression analysis are calculated and reported.

The range of an analytical procedure is the interval between the upper and lower sample concentration for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy, and linearity. The analysis and conclusions derived from the assessment of the repeatability, intermediate precision, relative accuracy, and dilutional linearity are used to establish the bioassay range over which results can be reliably reported.

8.2 Qualification Report

Subsequent to the execution of the protocol, all qualification results, data analysis, and conclusions are summarized in a report. It is recommended that information about critical equipment and reagents used (such as FBS and T&U cells) is also included.

The report should also contain a statement summarizing the qualification status of the bioassay method.

9 Challenges for IO Bioassay Development and Conclusion

Developing a cell-based functional assay to reflect the MOA of a drug that will be accessible at early stages of drug discovery will provide a smooth transition to support product lot release. Current methods to measure the potency of drugs for immunotherapy targets rely substantially on in vitro binding assays, primary T cell-based cytokine release assays, and in vivo model systems. Although in vitro binding assays satisfy high-throughput needs, lack of correlation with cellular functional response makes this method unreliable to screen out functional antibody candidates. Although antibody candidates can display high binding/blocking affinities, they may not display any functional response at the cellular level. Primary cell-based assays and in vivo model systems better reflect MOA but operate at lower throughput. High reliance on primary cells without continuing culture ability, as well as donor-to-donor variations, limit their ability for antibody screening in early drug discovery and lot release analysis. There is an urgent need for a simple, robust, plate-based functional bioassay to measure the potency of a drug candidate. Such an assay requires high sensitivity with appropriate specificity, precision, and accuracy for drug screening and characterization in early drug discovery, lot release, and stability studies. Luciferase reporter bioassays are designed based

on cellular signal cascades responding to drug treatment. Activation of the corresponding pathway triggers luciferase gene transcription, and a luminescent signal can be read out using a luminometer. Due to its inherent sensitivity, large signal dynamics, and simplicity to set up, this reporter assay platform has been widely used in high-throughput screening for decades. The assay specificity is even more reliable for biologic development due to the cell surface nature of all targets. Some concerns for small-molecule screens, such as false positives or signaling events distal from receptor activation, are less applicable for large molecules. During bioassay development, cells are considered critical reagents and reproducibility is extremely important to qualify an assay for drug lot release. Further development of reporter-based functional assays into a T&U format eliminates the burdens of daily cell culture and variables introduced by cell health and cell preparation from other factors that could directly contribute to assay variability. Proper functional QC and optimized criteria for the number of cells per vial, cell viability, and mycoplasma and bacteria contamination testing will ensure assay consistency and ease of bioassay transfer from one location to another.

Designing a successful cell-based bioassay reflecting a drug's MOA requires a clear cellular mechanism, and many designs rely on publications demonstrating target validation *in vivo* with a well-studied antibody as positive control. Accessibility of the control antibodies is critical to validate bioassays. In many cases in immunoncology, the targets of interest have been recently identified and antibodies are either proprietary or inaccessible from any commercial sources. Collaborations with pharma–biotech companies or reputable academic labs are critical to validate the assay design. Even when the antibodies in publications can be accessed from commercial sources, most of those antibodies are purified for research use only. The formulation of the products, especially the presence of the preservative sodium azide, often produces a hook effect at high antibody concentration, creating major challenges for assay optimization. In addition, these research-grade antibodies are mostly qualified for cytokine release, flow cytometry, and Western blotting. Therefore, it can prove challenging to apply these antibodies to T&U cells to demonstrate the suitability of a bioassay for testing biological potency.

Many immunotherapy drug targets show promising responses in mouse models without a clear understanding of their cellular mechanism of action, which creates a hurdle for designing cell-based assays for antibody screening and assay validation. Some targets are clinically relevant but a clear understanding of their ligand or corresponding receptors is unavailable. B7-H4, for example, was discovered as a B7 family member molecule that is responsible for T cell immunity [8]. Ligation of T cells with B7-H4 has a profound inhibitory effect on T cell growth, cytokine secretion, and

development of cytotoxicity. However, the T cell receptor responding to B7-H4 is still unknown as well as how the inhibition is mediated. In other cases, some targets are reported to have multiple ligands, and the clinical significance of each target–ligand interaction is largely still under investigation. The V-domain immunoglobulin suppressor of T cell activation (VISTA) is a negative immune-checkpoint protein that controls a broad spectrum of innate and adaptive immune responses [9, 10]. However, the ligand-or-receptor paradigm of VISTA in regulating T cell activation is unclear. VSIG3, VSIG8, and VISTA all interact with VISTA [11–13]. The immunoinhibitory molecule Lymphocyte-Activation Gene 3 (LAG-3, CD223) synergistically regulates T cell function with PD-1 to promote tumoral immune escape [14, 15]. Major histocompatibility complex Class II (MHC-II) is the canonical ligand for LAG-3, but it remains controversial whether MHC-II is solely responsible for the inhibitory function of LAG-3. It was reported recently that a newly identified ligand Fibrinogen-like Protein 1 (FGL1), a liver-secreted protein, is a major LAG-3 functional ligand independent from MHC-II [16]. FGL1 inhibits antigen-specific T cell activation, and ablation of FGL1 in mice promotes T cell immunity. Poor clinical outcomes from several MHC-II blocking anti-LAG-3 mAbs evaluated in clinical trials for the treatment of advanced human cancer may suggest that these antibodies do not block the clinically relevant ligand. Moreover, identifying and determining the availability of a biologically relevant cell line background to study the cellular pathway of a target of interest is critical and sometimes a limiting factor in designing a valid cell-based assay.

Immunotherapy is a novel, rapidly evolving cancer treatment with exciting benefits, but it also presents unique challenges for validating targets and determining their clinical roles in cancer treatment. With many immunotherapy trials in clinics, what is known today is very likely to change tomorrow. Developing cell-based functional bioassays that reflect the true MOA in this dynamic area, and embracing the challenges, will guide future research in this rapidly growing field.

References

1. Robinson CJ, Sadick M, Deming S et al (2014) Assay acceptance criteria for multiwell-plate-based biological potency assays. Biopharmaceutical Best Practices Association, Citrus Heights, CA
2. Gazzano-Santoro H, Chan LG, Ballard MS et al (2014) Ready-to-use cryopreserved primary cells: a novel solution for QC lot release potency assays. *BioProcess Intl* 12:28–39
3. Cheng ZJ, Garvin D, Paguio A et al (2014) Development of a robust reporter-based ADCC assay with frozen, thaw-and-use cells to measure Fc effector function of therapeutic antibodies. *J Immunol Methods* 414:69–81
4. USP 35-NF 30 (2012) USP Chapter <1032> Design and development of biological assays. USP Pharmacopeial Convention

5. USP 35-NF 30 (2012) USP Chapter <1034> Analysis of biological assays. USP Pharmacopeial Convention
6. Guideline, ICH harmonised tripartite. Validation of analytical procedure: text and methodology Q2 (R1)
7. USP 35-NF 30 (2012) USP Chapter <1033> Biological assay validation. USP Pharmacopeial Convention
8. Sica GL, Choi IH, Zhu G et al (2003) B7-H4, a molecule of the B7 family, negatively regulates T cell immunity. *Immunity* 18:849–861
9. Yapa EH, Roschea T, Almob S et al (2014) Functional clustering of immunoglobulin superfamily proteins with protein-protein interaction information calibrated Hidden Markov model sequence-profiles. *J Mol Biol* 426(4):945–961
10. Flies DB, Han X, Higuchi T et al (2014) Coinhibitory receptor PD-1H preferentially suppresses CD4+ T cell-mediated immunity. *J Clin Investig* 124(5):1966–1975
11. Xu W, Hiếu TM, Malarkannan S et al (2018) The structure, expression, and multifaceted role of immune-checkpoint protein VISTA as a critical regulator of anti-tumor immunity, autoimmunity, and inflammation. *Cell Mol Immunol* 15:438–446
12. Wang J, Wu G, Manick B et al (2019) VSIG-3 as a ligand of VISTA inhibits human T-cell function. *Immunology* 156(1):74–85
13. Wang J, Manick B, Renelt M et al (2018) Poster: VSIG-8 is a co-inhibitory ligand and an immune checkpoint molecule for human T cells. *J Immunol* 47(4)
14. Woo SR, Turnis M, Goldberg M et al (2012) Regulate T-cell function to promote tumoral immune escape. *Cancer Res* 72:917–927
15. Nguyen LT, Ohashi PS (2015) Clinical blockade of PD1 and LAG3—potential mechanisms of action. *Nat Rev Immunol* 15:45–56
16. Wang J, Sanmamed M, Datar I et al (2019) Fibrogen-like protein 1 is a major immune inhibitory ligand of LAG-3. *Cell* 176:334–347

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

