SYGNATURE DISCOVERY

Successful *In Vitro* Screening Strategies For Drug Discovery

PUBLISHED BY:

FIERCE Biotech

SYGNATURE DISCOVERY

Successful *In Vitro* Screening Strategies For Drug Discovery

PUBLISHED BY:

FIERCE Biotech

Introduction

In a previous eBook, we discussed approaches to successful hit finding. Here, we discuss some considerations in designing screening cascades to progress these hits and develop drug candidates. Whichever hit finding approach or methodology for designing and synthesizing new molecules you choose to develop your drug candidate, you will need a robust and flexible system to test your hypotheses in biological systems and ensure you remain on track.

As you start to consider what an appropriate screening cascade should look like, it is important to define what key properties you are looking for in your candidate drug, as these will impact the choice of assays you may deploy in your screening cascade. For example:

- How important is selectivity for your chosen target?
- Are you targeting an orthosteric or allosteric binding site?
- Does the drug need to cross the blood-brain barrier?
- Is the target expressed in a single cell type?
- What are the potential on-target toxicity concerns?

These aspects, and others, are worth early discussion to ensure that the right assays can be planned and introduced into the screening cascade at the appropriate time.

The key purpose of a screening cascade, or design-make-test cycle, is to allow rapid decision-making. Screening cascades will develop and change as the molecules progress —a cascade that is put in place during the hit-to-lead phase can and will look very different from the cascade during the lead optimization phase. At each stage though, the cascade must quickly answer the most pressing questions facing the project team at that moment in time. The rapid decision-making can lead to faster progression of compounds, timely filing of IP on novel chemical matter and, importantly, cost savings in the long run. Knowing when to stop a program and rethink strategy is as important a consideration as driving forward when trying to hit key milestones.

As mentioned above, screening cascades can look very different depending on the drug discovery stage at which you are operating. Below, we will describe some key factors worth considering as you think through how to construct an effective and robust screening cascade. We will focus initially on the hit-to-lead phase, as this encompasses the types of assays required and highlights some general considerations. Following this, we will explore some of the key differences in screening cascades as the project moves into the lead optimization phase.

Hit-to-Lead Phase

The key goal of this early phase of the drug discovery journey is to identify one, or ideally more, chemical series in which you have confidence that you are engaging your target in the desired way. These series will likely still have a long way to go before they can be considered drug candidates, but they are the fundamental starting points on that journey. As such, the screening cascade at this stage should be simple and allow for a rapid data turnaround to allow the chemistry team to develop a robust structure-activity relationship (SAR) and an understanding of the drug-like nature of their initial chemical matter. These aspects are critical to building confidence in the chemical series being taken forward.

A typical cascade might comprise a primary assay, an orthogonal assay, a cell-based assay and a selection of *In Vitro* DMPK (drug metabolism and pharmacokinetics) assays designed to benchmark (Figure 1). It is important to stress that these are guiding principles, and bespoke cascade design is often required in response to the key requirements for each project.

Figure 1: *Standard Hit-to-Lead Screening Cascade*

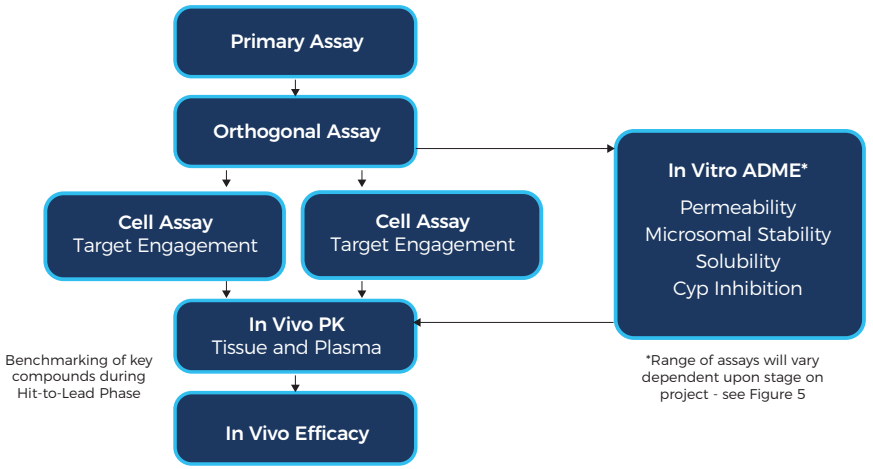
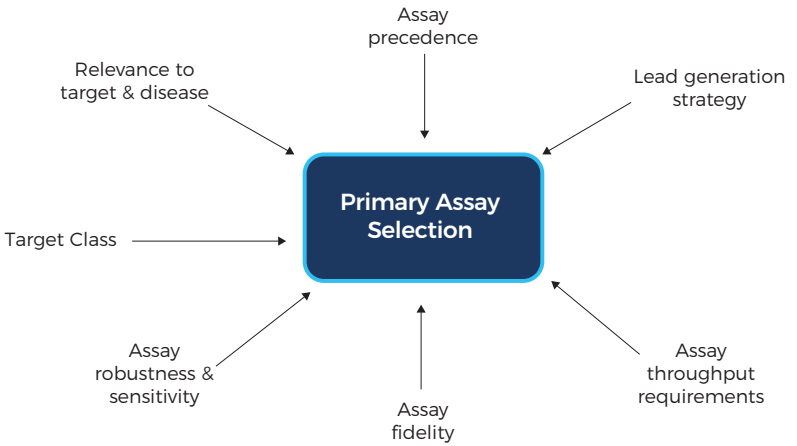


Figure 2: *Primary Assay Selection Considerations*



Primary Assay

Selecting an appropriate primary screen is crucial for the success of a drug discovery project, as this serves as the initial filter to identify compounds with potential therapeutic activity. Some key considerations need to be made when deciding on the optimum primary screen for a project (Figure 2). For example, what is the nature of the biological target? And how much is already known about how its activity can be modulated? Enzyme target classes, for example, are often highly amenable to the development of recombinant protein-based biochemical assays, which can form highly specific and robust primary screens. Similarly, if a well-defined protein-protein interaction is to be targeted, and recombinant proteins can be expressed, then robust biochemical or biophysical assays can typically be developed.

In other cases, the development of biochemical assays may be more challenging. Membrane targets, for example, may not always be as amenable to expression as appropriately folded proteins for biochemical screening. Similarly, if a particular intracellular signaling pathway component is to be targeted, the ability to recapitulate this in a biochemical assay format may not always be feasible. In such cases, a cell-based primary screen may be the best format.

The hit finding strategy for a discovery project will also strongly influence the choice of a primary screen. For example, if a fragment-based lead generation approach is to be taken, then a biochemical or biophysical primary screen will almost certainly be required, as compounds will need to be tested at high concentrations. This necessitates solvent concentrations that are not compatible with cell-based assays.

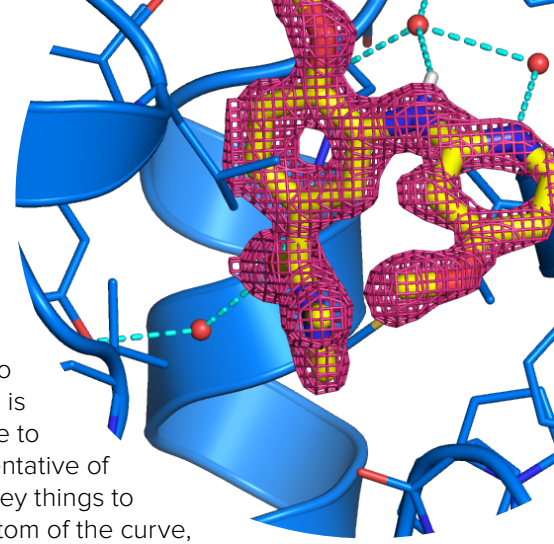
In all cases, the primary screen quality must be rigorously assessed to ensure it is suitable as an initial filter for the project. This includes an evaluation of the

assay sensitivity and specificity. A sensitive assay will be able to detect subtle changes in compound activity, increasing the likelihood of identifying hits. Assay sensitivity can often be enhanced by altering assay conditions. For example, sensitive detection of competitive inhibitors of an enzyme target will depend on substrate concentrations being set near or below the Km in the final assay conditions.

High assay specificity is also essential to maximize the proportion of identified hits that are “on-target”, reducing false positives. If the false positive hit rate is too high, this will severely impact the efficiency of downstream hit confirmation activities. Typically, biochemical and biophysical screens are less prone to off-target effects than cell assays. Still, in reality, both assay types are susceptible to compound interference mechanisms, necessitating a carefully tailored triage strategy downstream of the primary screen.

Primary assays will often be required to support high volume screening (for example, in HTS campaigns), so the design needs to be made as simple as possible. Employing homogeneous assay detection technologies is highly advantageous in these cases to minimize the number of steps required. These technologies are also highly amendable to miniaturization, which is also an essential requirement for high volume screening both in terms of efficiency and reagent costs. These technologies have enabled the use of both biochemical and cell-based high-volume primary screens, which—from experience within the Sygnature Discovery HTS group— can be equally robust as measured by retest confirmation rates.

In many cases, assays for the target of interest will be available, either from previous screening campaigns or literature precedence. In the latter case, it is good practice to carry out a significant level of assay development and optimization, as the majority of published assays have not been developed with a large-scale drug screening campaign in mind. Some of the key parameters that need to be optimized include assay signal to background and inter- and intra-plate signal uniformity, all of which impact the overall sensitivity and specificity of the final screen.



At this stage, assays are most typically run using a concentration-response-curve (CRC) to define compound potency versus the target. Although a typical assay parameter used to assess compound potency is IC_{50}, we would caution against just using these values to make decisions. Equally important at this stage is to evaluate each concentration-response-curve to ensure it fits pre-defined criteria and is representative of the mechanisms of action you are expecting. Key things to consider here include defining the top and bottom of the curve, hill slopes and compound solubility. Ignoring these parameters can run the risk of following non-specific compound activity that will impact overall project timelines and costs.

Orthogonal Assays

Orthogonal assays have two key purposes within a screening cascade that complement data from the primary assay. The first is to confirm any primary data, building confidence in the obtained results and confirming binding to the target of interest. The second is to generate data that could not conceivably be produced in the primary assay format, but will help drive the development of a potential therapeutic molecule by providing further information to characterize the binding interaction. This can include information such as stoichiometry, specificity, behavior, solubility, kinetics and more.

Primary assays are usually designed for throughput and ease of use, so the orthogonal assays of any cascade are essential to triage and enhance the data gained. This applies throughout the drug development process, from initial identification of hits to optimization and characterization of therapeutic candidates.

Orthogonal assays can take many forms but are often biophysical in nature to directly observe the binding of the interacting molecules.

Typically, these assay technologies have a lower throughput than the primary assay technology (although this is rapidly changing) but offer data

that allows a much greater understanding of the SAR and behavior of a potential therapeutic, as well as more specific data on its binding characteristics. These assays also tend be able to observe binding over a wider affinity range than primary assays, allowing characterization from millimolar binders down to picomolar, or even below. Confirmation of target engagement and generation of affinity data using these techniques are valuable to support IC_{50} data generated in a primary assay and are possible via most commonly used biophysical technologies. Some technologies allow additional data to be generated, the most prevalent of which is the determination of kinetic parameters and stoichiometry by surface plasmon resonance (SPR). The identification of kinetic parameters provides further insight into binding modes and may guide how a molecule is likely to behave in a disease setting.

For example, in most applications, a long residence time is favorable for an optimal therapeutic effect (Copeland 2015). Still, different compounds of the same affinity or IC_{50} may have significantly different kinetic properties (Figure 3). This detail would otherwise be missed by homogenous assay readouts, whether biophysical or otherwise, but is critical for identifying these properties and guiding the development of a therapeutic molecule. Stoichiometry of binding is another key factor, particularly when assessing molecules from a hit finding screen, as prioritizing binding ratios relevant to the system's biology is an important factor in identifying good hits from interfering molecules.

No assay is perfect, and this is also true of orthogonal assays, despite the data that they can generate throughout the drug discovery process. They are, in general, a highly artificial system and, despite efforts to make them as physiological as possible, are a significant step away from a natural environment. Cellular assays play a part here and, armed with the data from the primary and orthogonal assays, form the next step in developing the picture of how candidate molecules may affect the specific biology in question.

Figure 3: *Surface Plasmon Resonance (SPR) Characterisation of Compound Binding*

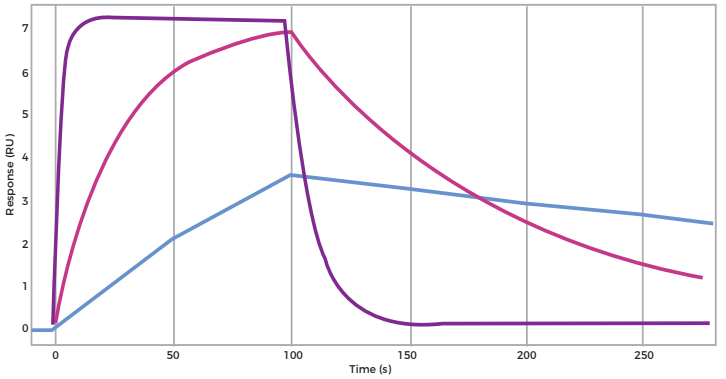


Figure 3: Simulated sensorgrams of three compounds with 100 nM affinities (produced using Biacore Simul8 software: https://apps.cytivsalifesciences.com/spr/). All three compounds have significantly different kinetic properties despite having the same equilibrium K_D, and may be better suited to different therapeutic modalities (Blue: k_a=2 x10^4, k_d=2 x10^{-3}, Red: k_a=1 x10^5, k_d=1 x10^{-2}, Purple: K_a=1 x10^6, k_d=1 x10^{-1}).

Cell-based Assay

Having some indication of cellular activity at the end of the hit-to-lead phase is desirable, and so developing a cell model that allows compounds to be evaluated in this context is an important component of the early drug discovery cascade. The choice of cell model and the endpoint to be measured is driven by the target biology. For example, programs that are looking to develop an anti-cancer compound will likely be looking for a cell-killing phenotype, whereas programs developing a drug to be used in an inflammatory disease setting are more likely to be interested in the modulation of a cytokine profile rather than cell death as an endpoint. Initial target validation studies will have provided some indication of the cellular phenotype to be expected upon target inhibition, either through the use of genetic or pharmacological inhibition (ideally both) of the target, and so these studies can provide a starting point for choosing the right assay. Irrespective of the amount of pre-existing data

available, it is always advisable to spend time building confidence in the literature data and ensuring that, in your own hands, the validity of the target holds up to scrutiny.

As mentioned earlier, one of the key tenets of the early drug discovery process is the speed of decision-making. This principle holds true for the cellular assay, as well as the primary and orthogonal assays. It is, therefore, worth considering the complexity of the cell assay you deploy at this stage. While the disease biology may employ multiple cell types and interactions, these can be very challenging to replicate in an *In Vitro* setting. Breaking the biology down to a simple system that can provide information on cellular activity on a weekly or bi-weekly basis will enable those key decisions to be made in a timely fashion.

Cell activity of compounds can be measured either as a phenotypic endpoint, such as cell proliferation, or using a mechanistic assay that confirms engagement of the target within the confines of the cell. Ideally, both assays would be in place in the screening cascade as they answer slightly different questions. The mechanistic or target engagement assay provides a direct indication that the compounds are binding and inhibiting the intended target. Examples of such assays include:

- The measurement of a marker validated to be downstream of the intended target, such as a phosphorylation modification on a protein
- The measurement of the thermal stability of the target upon compound binding (CETSA©)
- An engineered system that uses FRET-based technology to monitor compound binding in the cell.

The phenotypic assay provides evidence that the compound is impacting the biology in the desired manner.

A successful candidate will need to balance the biological activity and pharmacological effects with adequate human pharmacokinetics to deliver sufficient drugs to the site of action for clinical efficacy.

Typically, representative examples from each series being considered for progression are assessed across a broad range of *In Vitro* assays to understand the properties that will need to be optimized during the lead optimization phase. The aim should be to identify the series that is most likely to be able to address these in parallel with the pharmacological profile. Figure 4 illustrates a range of typical assays used in the hit-to-lead phase and how that data can be used.

Figure 4: *Key ADME Assays used in the Hit-Lead Phase*

Assay	Interpretation of Data
Kinetic Solubility	Risk of insufficient solubility to generate robust *in vitro* data.
Log D7.4	High LogD is often associated with poor metabolic stability and nonspecific binding. Low LogD often leads to poor permeability, and/or transporter mediated pharmacokinetics. Allows calcification of ligand, and in efficiency, metrics such as LipE as well as correlation with DMPK properties such as intrinsic clearance.
Metabolic Stability	Poor metabolic stability will lead to low unbound exposure for a given dose - this needs to be balanced with potency and the likely PK-PD relationship.
Permeability	Low permeability can lead to incomplete oral absorption and/or a dependence on transporter mediated pharmacokinetics. CNS targets should ideally avoid a high degree of efflux.
P450 Inhibition	Inhibition of metabolizing enzymes should be minimized to mitigate the risk of clinically relevant drug-drug interactions.

Ultimately, this initial profiling exercise should inform the screening cascade in the lead optimization phase and should focus on properties where significant risk has been identified. For example, routine screening for permeability might not be required if initial hits have high permeability and the chemistry plan is likely to operate in a similar physicochemical space. Conversely, a CNS project will likely need to understand the potential for efflux early in the cascade in order to maximize the unbound exposure in the brain relative to plasma, whereas pro-drugs may require early assessment of plasma stability.

Lead Optimization (LO) Phase

The transition into the LO phase will occur when the project team feels they have established one or more chemical series that they believe can be progressed towards a drug candidate. At this stage, there will be an understanding of the liabilities associated with each series, so the screening cascade can be modified to ensure that these become the focus going forward.

The main assays developed within the hit-to-lead phase are still important and form the core of the screening cascade, but often, additional assays can be introduced. These could include assays to access selectivity versus close family members, cross-species evaluation (important for *in vivo* studies and toxicology) and more complex cell-based assays systems.

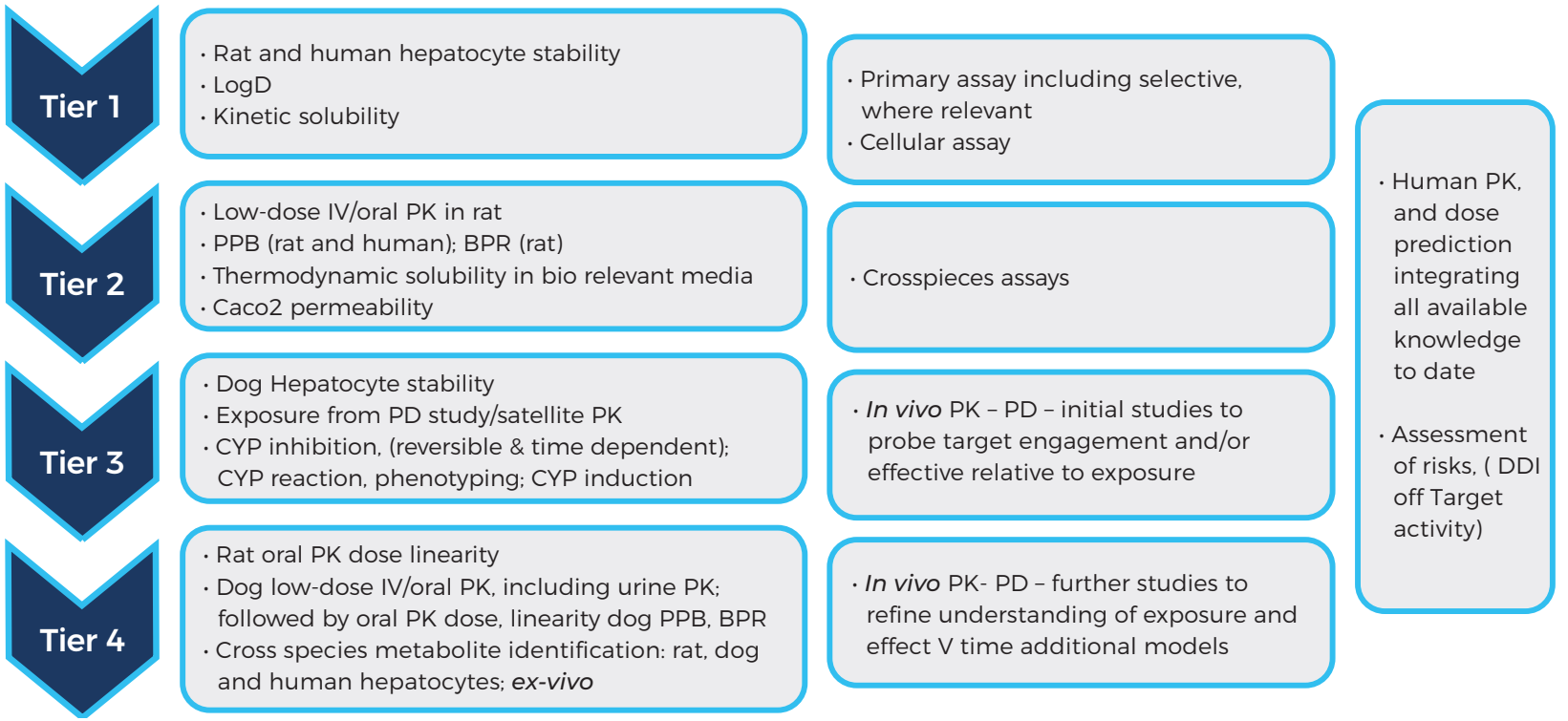
Depending on how important to the overall profile of the final drug candidate these are believed to be, these assays may become part of the core cascade or be used in a more ad-hoc manner to ensure that the project remains on track to deliver the desired profile. During this phase, the focus on the lead chemical series may change, and each series will have its own unique challenges. It is important, therefore, to continually re-evaluate the screening cascade and ensure that the correct assays are being prioritized to deliver the key data which will drive the decision-making for the project. Another difference from the hit-to-lead phase will be an increased emphasis on the overall profile of the molecules, not just potency, to ensure they have the right properties to be viable drug candidates. This is where expert DMPK input really comes to the fore.

A typical LO DMPK screening cascade is shown alongside the *In Vitro* and *in vivo* pharmacology assays in Figure 5. Tier 1 is focused on identifying those compounds with the right balance of metabolic stability and potency while keeping an eye on physicochemical properties. Tier 1 should also incorporate key liabilities identified in the hit-to-lead phase.

Compounds with a suitable profile are then progressed into Tier 2 where, for the first time, we obtain *in vivo* data to understand the pharmacokinetics in rodents and to assess the potential for progression into pharmacology studies.

A good correlation with the *In Vitro* DMPK data is a key aim of these studies, along with an assessment of whether sufficient (unbound) exposure to explore the PK-PD relationship in pre-clinical models is likely to be achieved. If this isn't the case, then the reasons for this should be explored, and the screening cascade should be adjusted to reflect this. For example, metabolic stability in hepatocytes is unlikely to be relevant for compounds predominantly eliminated by renal clearance or hepatic uptake.

Figure 5: *Typical LO Screening Cascade*



Tiers 3 and 4 aim to build both the PK-PD relationship as well as enhance our understanding of pharmacokinetics across species.

For example, by integrating the PK data in Tier 2 with the *In Vitro* pharmacology, we can design PK-PD studies to probe the relationship between (unbound) exposure and effect *in vivo*. Non-rodent PK allows us to confirm if our understanding of PK is robust across species while helping to assess suitable species for toxicology.

Together, this data allows us to continuously refine the human PK and dose prediction by integrating all the available data generated. This prediction also allows us to assess DDI risk and likely safety margins to support candidate selection.

Perfect is the enemy of the good' (Voltaire) and Multi-parameter Optimization

While most projects will start with generic target criteria for a candidate, these tend to be based on a notion of perfection to which the team should aspire. In reality, a project can often compromise in one or more properties, providing there is an advantage in others. For example, exceptionally potent compounds that require very low unbound plasma exposure for efficacy may tolerate a much higher intrinsic clearance to achieve sufficient exposure from an acceptable dose than those that are less potent. Such compounds may also present a low risk of clinical DDI even with relatively low CYP inhibition IC_{50}. Further work to address this may be of no benefit.

Rather than applying arbitrary cut-offs, a flexible approach that integrates the knowledge from the bioscience, DMPK and *In Vitro* pharmacology assays into the screening cascade will often be required. At its heart, medicinal chemistry is the science of optimizing multiple parameters while understanding their interrelationship.

Beyond the rule of 5 space: degraders, bifunctional molecules and beyond (one size does not fit all)

The discussion up to this point has focused on developing screening cascades for small molecule drug discovery. As discussed, these are guidelines, and each screening strategy requires careful thought. This is particularly true for new modalities, such as bifunctional molecules, where the challenges with achieving orally bioavailable drug candidates become more acute.

Targeted protein degradation using bifunctional degraders, which connect a targeted protein binder and E3 ligase via a linker, has emerged as a powerful drug discovery tool for seemingly undruggable targets. These compounds have molecular weights exceeding 500 and can contain more than 5 H-bond donors and 10 H-bond acceptors. These molecular properties are required to deliver the essential compound interactions at both the protein of interest and the E3 ligase binding sites. This functionality results in these compounds having a much higher topological polar surface area (TPSA) than traditional small molecules. However, the high polar surface area can be detrimental to the ability of compounds to cross hydrophobic membranes. Absorption from the gastrointestinal tract can, therefore, be limited due to poor permeability. On the other hand, high levels of polarity are beneficial to compound properties such as aqueous solubility.

Due to the flexibility of bifunctional protein degraders and their potential to form intramolecular hydrogen bonds, some of the polar surfaces can be buried or shielded, depending upon their environment, thereby increasing permeability. This property, known as molecular chameleonicity, is facilitated by the inclusion of multiple rotatable bonds in bifunctional degraders. The remaining exposed, or effective, polar surface area (EPSA) then determines the cellular permeability. Therefore, incorporating features in bifunctional degraders that reduce the EPSA, yet maintain acceptable lipophilicity, can result in compounds with improved gastrointestinal permeability.

The use of traditional *In Vitro* permeability assays employed for small molecule characterization, e.g. Caco-2 or MDCK assays, can be challenging for these

types of compounds. Therefore, Tier 1 assays include the use of supercritical fluid chromatography to determine the EPSA, and a chromatographic LogD assay to rapidly measure lipophilicity. By combining these two measurements, it is possible to rapidly characterize the physico-chemical properties of novel compounds and identify the ones with the best chance of acceptable intestinal permeability.

Once identified, these compounds can be prioritized for further profiling, including an assessment of their metabolic stability to avoid the impact of high levels of hepatic metabolism. Additionally, solubility in biorelevant media should be measured to assess their absorption potential. Measurement of fraction unbound in plasma can be problematic for bifunctional degraders, resulting in the need for assay adaptations including the use of extended incubation times, plasma dilution and/or pre-saturation methods. The stability of these compounds in plasma should also be assessed, particularly where poor recovery is observed in binding assays. Promising bifunctional degraders identified in these *In Vitro* assays can then move on to *in vivo* PK studies to determine their oral bioavailability and potential as future human therapeutics.

Data Analysis

With the high volume of data being generated during the hit-to-lead phase, it is important to have mechanisms in place to analyze the data as a whole and ensure that it all makes sense. This is most simply done using correlation analysis, ensuring that the activity of the chemical series is tracked consistently across the different assay platforms. These correlations will indicate whether compound activity in the primary and orthogonal assays is consistent with the predicted SAR. In addition, they also have the benefit of highlighting when a potency wall is encountered in one of the assays, indicating that a re-evaluation of the assay parameters may be required to increase the sensitivity.

It is very important to track the correlation between the primary assay and cell-based readouts to ensure that the inhibition of the chosen target is driving the desired biological effect. Often, there may be some frameshift in potency measurements from biochemical to cell assay (depending, for example, on the coupling of target inhibition to the cell assay readout or due competition from endogenous ligands). However, where there is a significant deviation from a linear correlation, further investigation is needed. A lack of cell activity may be due to poor physiochemical properties preventing access to an intracellular target. Alternatively, it may indicate that the method of inhibiting the target is not translating into the desired cellular effect, and so highlights a need to re-evaluate the project strategy. In other cases where unexpected cell activity is being observed, this may indicate off-target activity of compounds and should trigger wider selectivity assessments to better understand these cellular effects.

Bringing it all together

Having a dynamic, fit-for-purpose and robust *in vitro* screening cascade, both for biology and DMPK, will improve the chances of discovering a new drug in a timely manner. Efficient assays carried out on co-located sites drive the design-make-test-analyze process, giving the scientists the data they need to make decisions. Drug discovery is a collaborative process – scientists who work in isolation will fail alone.

To discuss your screening strategies and to learn how a strategic, holistic approach contributes to successful project outcomes, contact the Sygnature Discovery team via our website.

SYGNATURE DISCOVERY

Sygnature Discovery is a world-leading integrated drug discovery Contract Research Organisation based in the UK and Canada with its headquarters in Nottingham and additional sites in Alderley Park, Macclesfield, Glasgow, Montreal and Quebec City. Its staff of over 1,000 employees, which includes more than 900 scientists, partners with global pharma, biotech and NFP organisations.

Since 2011, Sygnature Discovery has delivered over 40 novel pre-clinical and 22 clinical compounds, with its scientists named on over 170 patents. Therapeutic areas of expertise include oncology, inflammation and immunology, neuroscience, metabolic diseases, infectious diseases, fibrotic diseases and more.