Zebrafish-Based Small Molecule Discovery

Calum A. MacRae and Randall T. Peterson*
Developmental Biology Laboratory and Cardiovascular Research Center
Massachusetts General Hospital
Charlestown, Massachusetts 02129

The earliest examples of small molecule discovery involved serendipitous phenotypic observations in whole organisms, but this organism-based process has given way in recent decades to systematic, high-throughput assays using purified proteins, cells, or cell extracts. In vitro screens have been successful at identifying modifiers of well-understood biological processes, but they are limited in their ability to discover modifiers of processes that are poorly understood or occur only in an integrated physiological context. Small model organisms, especially the zebrafish, make it possible to combine the advantages of organism-based small molecule discovery with the technologies and throughput of modern screening. The combination of model organisms with high-throughput screening is likely to extend small molecule discovery efforts to fields of study such as developmental biology and to broaden the range of diseases for which drug screening can be performed.

Introduction
Organism-based approaches to small molecule discovery are not new. Prior to advances in in vitro screening in the 1970s, a significant proportion of biologically active small molecules were discovered based on their interactions with whole organisms. Often, serendipitous exposure of animals to naturally occurring small molecules resulted in interesting (or terrifying) phenotypes, leading ultimately to characterization of the causative molecules. For example, dicumarol was discovered in the 1930s when cattle in the northern United States began dying inexplicably of internal bleeding after having fed on rotating sweet clover [1]. A group of researchers from the University of Wisconsin identified the causative agent as dicumarol, which with its derivatives remains the most frequently prescribed anticoagulant today. More recently, observation of cyclopia and other developmental defects in sheep exposed to Veratrum californicum led to discovery of the smoothered antagonist cyclopamine [2].

More than any other animal, humans have been the subjects for such serendipitous small molecule discoveries. In 1776, William Withering observed the beneficial effects of the foxglove plant on a woman suffering from heart failure. From the foxglove he isolated digitalis, which is still an important therapy for heart failure [3]. Analgesics, antinfiammatories, antifungals, and of course many psychoactive compounds have been discovered by human exposure to small molecules.

Although the examples listed above illustrate the power of whole organism, phenotype-based compound discovery, practical and ethical factors have limited the overall utility of the approach. The examples above involved unplanned exposures of cattle, sheep, and humans to small molecules, followed by astute observation of the results. Planned, systematic exposure of mammals to uncharacterized small molecules for the purpose of compound discovery would be expensive, laborious, require large quantities of the chemical compounds, and could potentially raise ethical questions.

To accelerate small molecule discovery, some attempts to systematize whole-organism chemical screening were made. In 1907, Paul Ehrlich tested 900 arsenical compounds on mice infected with trypanosomes [4]. Although unsuccessful, these experiments laid the groundwork for the eventual discovery of the syphilis treatment Salvarsan. In the 1950s, Squibb Co. discovered the anti-tuberculosis drug isoniazid after exposing mice to more than 5000 compounds (and purportedly killing more than 50,000 mice in the process) [5]. Therefore, despite a few notable successes, the benefits of whole-organism, phenotype-based small molecule discovery have been overshadowed by the practical limitations of the approach.

During the last three decades, small molecule discovery has moved from organism-based to target-based discovery. In target-based discovery, a target protein is identified, purified, and used for in vitro screening to identify inhibitors of its biochemical activity. This target-based approach has made compound discovery more systematic, automated, and able to capitalize on the burgeoning data generated by molecular and medical research. A number of outstanding successes have resulted, including the discovery of HIV protease inhibitors and, more recently, the c-Abl kinase inhibitor Gleevec.

The Benefits of Organism-Based Small Molecule Discovery
The ascendancy of in vitro, target-based approaches has made compound discovery systematic and rational, but several important advantages of organism-based discovery have been sacrificed (Figure 1). First, while organism-based discovery can identify chemical modifiers of virtually any biological process, modern target-based approaches typically only discover modifiers of well-understood processes. Because target-based approaches are dependent upon a priori selection of a protein target, they are biased toward discovery of small molecule inhibitors of well-known proteins. As a result, this approach is less likely to reveal the functions of novel targets and provide novel fundamental insights into poorly understood phenomena. Biological understanding is a prerequisite, not a result, of target-based approaches. Despite the thousands of potential small molecule targets identified by the human genome project, it remains difficult to predict which proteins must be targeted to reverse a disease phenotype or to alter a poorly understood organismal process. Consequently,
target-based approaches continue to identify small molecule modulators of a relatively small group of targets, while novel therapeutic targets await validation by other means. Moreover, fields of study such as developmental biology receive little benefit from the discovery of new molecules. Like other chemical genetic screens [6], organismal screens permit discovery of new pathways and targets, but they also expand the range of observable phenotypes to include those affecting development, physiology, and behavior.

The second advantage of using whole organisms for compound discovery is the physiological context they provide. Target-based discovery is usually performed using purified proteins, tissue culture cells, or cell extracts in in vitro assays. Small molecules discovered in this in vitro context may have unexpected activity in an in vivo context. Initial hits must be tested further in animals for effectiveness, side effects, toxicity, and pharmacokinetic/pharmacodynamic profile. The vast majority of small molecules discovered by in vitro target-based screening exhibit undesirable characteristics (such as lack of specificity or toxicity) when tested in the context of a whole organism. In contrast, a small molecule discovered by virtue of its ability to cause a desirable phenotype in a whole organism is more likely to be cell permeable, devoid of obvious toxicities, effective, with an acceptable pharmacokinetic/pharmacodynamic profile. Efficacy and specificity in an organismal context is the standard that small molecules ultimately seek to meet. Moving the organism to the initial stages of compound discovery combines screening and animal testing into one step.

Recently, a number of groups have attempted to combine the advantages of organism-based small molecule discovery with the technologies of modern high-throughput screening. Small model organisms including Caenorhabditis elegans, Drosophila melanogaster, Arabidopsis thaliana, and the zebrafish Danio rerio are all small enough to grow in microformat screening plates. Among these organisms, the zebrafish is most closely related to humans and is perhaps best established as a tool for small molecule discovery. This review will focus on studies involving zebrafish, with the recognition that many of the ideas discussed may be applicable to other model organisms as well.

**Figure 1. Organism- versus Target-Based Small Molecule Discovery**

Target-based discovery, such as discovery of cox-2 inhibitors (top), begins with selection of a molecular target, followed by screening for chemical modifiers of the target’s activity. The true biological effect (phenotype) is only determined at the end of the process when efficacy and safety are tested in animals and man. In contrast, organism-based discovery begins with detection of desirable phenotypes in whole organisms and may lead to discovery of novel targets. For example, the discovery of dicumarol began as an observation of hemorrhaging in cattle and proceeded to identification of vitamin K epoxide reductase as its target (bottom).

**Zebrafish as a Model Organism**

The zebrafish is rapidly gaining popularity as a model organism for developmental biology and genetics [7–10]. The attributes that have made the zebrafish such a powerful model for genetic screening also make it well suited for small molecule screening. First, unlike the fruit fly, round worm, baker’s yeast, or other popular genetic model organisms, the zebrafish is a vertebrate. It possesses discrete organs and tissues such as a brain, sensory organs, heart, liver, pancreas, kidneys, intestines, bones, muscles, etc. (Figure 2A). Zebrafish organs are remarkably similar to their human counterparts at the anatomical, physiological, and molecular levels, facilitating study of vertebrate biological processes that are inaccessible using invertebrate model organisms.

The zebrafish is also small for a vertebrate. Adults reach 3 cm in length, but during the embryonic and larval stages of life, the zebrafish is only about 1–2 mm long. During these stages, developing fish can live for days in a single well of a standard 384-well plate, surviving on nutrients stored in their yolk sacs (Figure 2B). Zebrafish are simple and inexpensive to raise, and a single pair of adults can routinely lay hundreds of fertilized eggs in a single morning. Consequently, even a small zebrafish facility can generate many thousands of embryos per day, making it possible to efficiently screen large libraries of molecules.

Another notable attribute of the zebrafish is the transparency of its embryos. Every internal organ and structure can be visualized in the intact, living organism, obviating the need to sacrifice or dissect the animal and permitting multiple observations of dynamic processes. In contrast, observing organ development in the developing mouse typically requires sacrificing the mother, followed by dissection of the embryo. The optical clarity of the zebrafish embryo becomes even more useful when combined with fluorescent markers that highlight the locations or activities of specific populations of cells (Figure 2C). For example, dozens of transgenic zebrafish lines have been created that express fluorescent proteins in locations ranging from the presomitic mesoderm [11] to the pituitary gland [12]. These lines greatly facilitate detection of anatomical changes caused by small molecules. Fluorescent assays have also been developed that report changes in zebrafish physiology. These assays include a transgenic line in which expression of
green fluorescent protein is induced upon activation of the aryl hydrocarbon receptor [13, 14]. These lines have been proposed for use as sentinels for automated detection of fresh water toxicants that activate the aryl hydrocarbon receptor. Similar lines could be engineered for high-throughput screening for chemical activators of any gene of interest. In another assay, lipid metabolism is measured in vivo based on a fluorescent signal generated by cleavage of quenched phospholipid substrates [15]. Larval zebrafish are fed fluorescently quenched phospholipids that are cleaved in vivo and transported to the gall bladder where they are readily visualized (Figures 2D and 2E). In short, the transparency of the zebrafish embryo makes it possible to rapidly assess the effects of small molecules on many aspects of anatomy or physiology.

As zebrafish have become more widely used, additional technologies have been developed that have increased the utility of the system even further. The zebrafish genome project is now nearly complete, and DNA microarrays have been generated for expression profiling studies [16, 17]. Cloning of zebrafish by nuclear transfer has been accomplished [18], and antisense morpholino oligonucleotides have proven to be an effective means of “knocking down” gene function [19]. More recently, reverse genetic approaches have been developed for the zebrafish, enabling researchers to generate mutations in virtually any gene of interest [20]. Thus, the zebrafish is rapidly becoming a mature model organism armed with an impressive collection of genomic and experimental tools. These tools are also broadening the scope of whole-organism chemical screens that can be imagined.

**Discovery of Chemical Probes for Developmental Biology**

Perhaps the most obvious application for zebrafish chemical screening is the discovery of small molecule probes for developmental biology. The zebrafish is well established as a model organism for developmental biology. Forward genetic screens have been successful at identifying thousands of mutations that affect development of nearly every organ system [21, 22]. The screening methodologies used for identifying zebrafish genetic mutants are readily adapted for chemical screening, and the collection of mutants provides us with some sense of the kinds of phenotypes that may be expected.

The simplest screens for chemical modifiers of development involve arraying water and wild-type embryos into the wells of 96- or 384-well plates, adding small molecules from a chemical library to the water in the wells, and allowing development to proceed. At predetermined stages of development, the embryos are screened visually for developmental perturbations in the system(s) of interest (Figure 3). In one such screen, four organ systems—the central nervous system, the cardiovascular system, the ear, and the skin—were examined using a dissecting microscope [23]. After screening 1100 small molecules, modifiers of all four systems were identified. One compound called 32N5 causes a malformation of the hindbrain. Another, 32P6, causes a heart patterning alteration in which the ventricle forms within, rather than adjacent to, the atrium. Two compounds affect the development of melanocytes in the skin. One of these blocks pigment production in all cells by inhibiting the enzyme tyrosinase. The other prevents devel-
Figure 3. Screening for Chemical Modifiers of Vertebrate Development

Adult zebrafish lay hundreds of fertilized eggs each morning. Embryos are arrayed in assay plates, and compounds from small molecule libraries are added to the water in each well. Embryos are allowed to develop and are screened visually for developmental defects. Examples of specific developmental phenotypes include elongation of the notochord (A), absence of blood (B, untreated, upper; treated, lower), and loss of a single otolith in the ear (C).

Development of only a subset of pigmented cells that are derived from the neural crest. Apparently, this compound blocks the specification of neural crest cells to become melanocytes or their subsequent proliferation. In other screens, small molecules have been identified that block differentiation of blood cells, prevent formation of the eyes, alter the length of the notochord, or affect fin length (unpublished results). Thus, the range of developmental phenotypes that can be identified by small molecule screens seems almost limitless.

Several of the compounds discovered by small molecule developmental screens appear to be potent and specific [23]. The molecule 32P6 affects heart chamber patterning with an EC50 of 2 nM. Although less potent, another molecule (31N3) causes a remarkably specific developmental phenotype. Embryos treated with 31N3 fail to form the two tiny otoliths of the inner ear, while the rest of embryo appears to develop normally. The specificity of the phenotypes identified by chemical screening seems to approach those identified by genetic screens. In fact, several small molecules phenocopy specific genetic mutations.

Given that thousands of developmental mutants have already been identified, what additional value will be added by identifying chemical modifiers of development? Having temporal control over disruption of developmental pathways is one significant benefit, especially since development is inherently such a time-dependent process. Signaling pathways involved in early development often perform additional roles at later stages of development. The sonic hedgehog pathway, for example, is believed to be involved in development of dozens of structures at various developmental time points [24]. Because genetic mutation is generally permanent, it is often difficult to distinguish the role of a gene at a later stage of development from secondary effects of disrupting the gene at earlier stages. In contrast, small molecules can be added and removed at any stage of development.

The sonic hedgehog (Shh) pathway is one of the few developmental pathways for which a specific small molecule modifier exists. Cyclopamine, which antagonizes the sonic hedgehog effector smoothened [25], has proven invaluable for determining the role of the hedgehog pathway in later developmental events. One recent illustration of this was the use of cyclopamine to determine how different muscle cell types are specified in the developing myotome [26]. During zebrafish myotome development, the specification of three muscle cell types—the muscle pioneers, the superficial slow fibers, and medial fast fibers—are all dependent upon Shh. Wolff et al. elegantly showed that the correct specification of the proper cell type is a function of both the level of hedgehog to which they are exposed and the timing of the exposure [26]. To do this, they treated embryos with cyclopamine at various times and concentrations and demonstrated that the cell identities of the myotome are altered by the resulting changes in the timing and strength of hedgehog signaling. Such an analysis would not have been possible without the temporal and quantitative control offered by the small molecule cyclopamine. It is probable that small molecules discovered by whole-organism screening in zebrafish will, like cyclopamine, be useful tools for dissecting other aspects of vertebrate development. One such molecule has already
been useful for studying the mechanism of determining heart chamber orientation during development [27].

One major obstacle that may limit the utility of small molecule developmental screens is the current difficulty of identifying molecular targets for novel compounds. Once a small molecule is identified that disrupts a developmental process of interest, no systematic process exists for identifying the responsible target. Several approaches to small molecule target identification have proven successful in the past, including candidate gene approaches, expression cloning, and affinity chroma-
tography using immobilized small molecules [28]. However, not every approach works in every situation, and the selection of compounds with activity against the microbe of developmental biology to benefit from the discovery of chemical probes will be greatly enhanced. Several methods for systematizing target identification have been proposed, including synthesis of chemical libraries with preattached linkers. One such library was used to identify a small molecule that affects zebrafish brain develop-
ment [29]. The preattached linker allowed facile generation of an affinity matrix and biochemical purification of a protein binding partner.

**Zebrafish Screens for Drug Discovery**

Beyond discovery of small molecule probes for fundamental biological research, it may be possible to use zebrafish screens to identify lead compounds with therapeutic potential. If human diseases can be accurately modeled in zebrafish, chemical screens using zebrafish disease models as substrates could be used to identify compounds that ameliorate the disease phenotypes. Importantly, even diseases without a known, druggable target may be amenable to this approach, because no prior assumptions about the mechanism of disease amelioration are necessary.

Many diverse zebrafish disease models have already been developed and are reviewed elsewhere [30–32]. Most of these models fit within one of three categories. In the first, the human disease and the zebrafish model share the same phenotypic manifestation and are known to share the same underlying cause. For example, humans with hereditary hemorrhagic telangiectasia and zebrafish *violet beauregarde* mutants both have arteriovenous malformations and share mutations in the activin receptor-like kinase 1 gene [33]. Disease models in this category are obvious candidates for zebrafish-based drug lead discovery. In the second category, human and zebrafish disease phenotypes are similar, but are not known to share the same underlying cause. For example, causes a malformation of the aorta that is similar in many ways to coarctation of the aorta in humans [34].

Success has also been achieved generating zebrafish models of infection, and there may be advantages to screening for antimicrobial compounds in the context of a whole organism. Some infectious pathogens cannot be cultured outside of a host, and a zebrafish model may allow these pathogens to be subjected to high-throughput chemical screening. Furthermore, screening for antimicrobials in an organismal context would allow selection of compounds with activity against the microbe but without undue toxicity to the host.

Two recent papers highlight the potential of zebrafish for modeling infectious diseases. Davis et al. demonstrated that zebrafish embryos can be readily infected with *Mycobacterium marinum*, a close relative of the agent that causes tuberculosis in humans, *Mycobacterium tuberculosis* [35]. In zebrafish, the mycobacteria cause chronic infection of macrophages and result in formation of tuberculous granulomas exhibiting many of the hallmarks of tuberculosis. Granuloma-specific mycobacterium genes are also activated, and the infection reliably results in death of the zebrafish by 9 days post infection. Mycobacteria engineered to express green fluorescent protein were used, making it easy to observe progression of the infection. Van Der Sar et al. performed similar experiments using DsRed-labeled *Salmonella typhimurium* and demonstrated that infection progresses much as in established murine models [36]. The bacterium multiplies in macrophages and at blood vessel epithelial cells and is ultimately lethal. In both examples, it is easy to envision high-throughput screening for compounds that prevent progression of the infection, either by observing the fluorescently la-
beled pathogen directly or by selecting compounds that permit survival of the zebrafish host beyond the time of expected lethality.

One question that remains to be answered is whether small molecules that modify a disease phenotype in zebrafish will have similar activity in humans. That question has not been answered, but it has been shown that several drugs with known effects in humans cause analogous effects in zebrafish. In one study, Milan et al. exposed zebrafish to 23 drugs known to lengthen a portion of the cardiac cycle in humans known as the QT interval [37]. Of the 23 drugs, 22 also caused an analogous prolongation of the cardiac cycle in zebrafish. Other drugs with similar activities in fish and humans include cholesterol synthesis blockers, vasodilators, angiogenesis inhibitors, and anticoagulants, as reviewed elsewhere [38]. Therefore, it appears that drug binding sites are generally well conserved between humans and zebrafish, and many compounds that are active in zebrafish may have similar activities in humans. This fact combined with the existence of good zebrafish disease models suggests that zebrafish-based screening for new drug leads may be possible. Several such screens are currently underway.

**High-Content Characterization of Chemical Libraries**

In addition to their use for small molecule screening itself, zebrafish assays can play supporting roles in the process of small molecule discovery. Because zebrafish assays are “high-content” assays, a single embryo can provide a great deal of information about the molecule
Figure 4. High-Throughput Assays of Cardiovascular Physiology

(A) Automated video microscopy captures 15 s videos of zebrafish hearts in 96- or 384-well plates. Data from one 384-well plate can be acquired in less than 2 hr. Light intensity in the region of the heart is plotted as a function of time.

(B) Computer algorithms automatically extract physiological data from the videos, including heart rate.

(C) High-throughput screens can be used to detect individual compounds or combinations that affect cardiovascular physiology [37]. In this example, a QT prolonging interaction between erythromycin and cisapride was detected. Atrial rate is plotted against increasing doses of cisapride. Erythromycin concentrations are indicated in the inset.

Erythromycin and cisapride alone have little effect on heart rate, but together they cause severe bradycardia. (Images courtesy of David Milan.)

to which it is exposed. As such, zebrafish assays are excellent for characterizing collections of compounds at various stages in the compound discovery process. This characterization can take the form of testing libraries for functional diversity and biological activity. Particularly after synthesis of novel chemical libraries via combinatorial chemistry, it is important to gauge the likelihood that the libraries contain compounds with biological activity. Because disrupting any one of thousands of gene products in the zebrafish causes an observable phenotype, testing even a relatively small number of library compounds is often sufficient to assess a library’s potential for biological activity. Zebrafish developmental assays have been used in this way for preliminary characterization of several novel combinatorial libraries [39–41].

During drug discovery and development, early detection of compounds that cause unacceptable toxicity can save much wasted time and expense. Inexpensive, high-throughput zebrafish toxicity assays can be used to eliminate toxic compounds before time and money are invested in their further development. This can occur prior to screening by testing whole libraries or after screening as a means of prioritizing hits for further development. The zebrafish is becoming a well-accepted model for toxicologic pathology [42], and at least one company has been founded with the intent to provide zebrafish screening services for drug toxicity profiling. High-throughput assays for specific types of toxicity have also been developed. For example, prolongation of the QT interval is a common culprit behind abandonment of previously promising drug leads. Milan et al. developed an automated system for measuring zebrafish heart rate data in 96- and 384-well plates and showed that the system can be used to predict QT prolongation [37]. With few exceptions, individual drugs and drug combinations that cause QT prolongation in humans also cause bradycardia in zebrafish and can be detected by the automated zebrafish assay (Figure 4). As with most emerging technologies, more experimentation and practical application will have to be performed before the real utility of zebrafish toxicity screening can be determined. However, the ability to perform whole-animal studies rapidly, inexpensively, and in large numbers is an appealing possibility.

Conclusions

In vitro, nonorganismal screens are likely to remain a mainstay of small molecule discovery efforts, particularly those directed at well-understood biological processes. However, organism-based screens address some of the shortcomings of modern small molecule discovery, notably the lack of appropriate physiological context provided by in vitro assays and the difficulty of approaching biological problems for which previously validated targets do not exist. Drug discovery efforts may benefit from the increased efficiency of combining screening and animal tests into one step. Zebrafish screens may also impact drug discovery by allowing diseases to be tackled that were previously intractable by target-based methods.

At present, small molecule discovery is largely the purview of the pharmaceutical industry. Perhaps the most significant contribution of organism-based screening will be in helping to increase the feasibility and appeal of performing small molecule discovery in academic settings. Far from requiring the equipment and automation of a high-throughput screening facility, any lab with a few fish tanks, a chemical library, and a microscope can perform a zebrafish chemical screen. And
because understanding a biological process is an aim, not a prerequisite, of organism-based screens, the kinds of biological questions that can be approached are limited only by an investigator’s imagination. Therefore, like traditional genetic screens in model organisms, zebrafish small molecule screens hold great potential as a mechanism for discovering novel biological pathways. With luck, such efforts will lead both to increased biological insight and to the discovery of small molecules for the treatment of human disease.

Acknowledgments

The authors thank Shuo Lin, Ning-Al Liu, and David Milan for providing images and Stanley Shaw, Jordan Shin, and Ashok Srinivasan for helpful comments on the manuscript.

References


