Drug Discovery Today: Technologies
Vol. 23, 2017

The evolution of library design: crafting smart compound collections for phenotypic screens☆

Kerry L. Spear[1,*], Scott P. Brown[2]

[1]Blue Oak Pharmaceuticals Inc., Concord, MA, USA
[2]Sunovion Pharmaceuticals Inc., 84 Waterford Dr., Marlborough, MA 01752, USA

The (re)emergence of phenotypic drug discovery has been marked by a growing interest in screening campaigns that utilize phenotypic assays. The key objectives of phenotypic screens are different from those of target-based screens and can require alternate library-design strategies. Designing a library that is appropriate to the selected assay increases the likelihood of identifying better quality hits, which can reduce both timelines and overall cost of the drug-discovery process. Here, we provide an overview of small-molecule library design principles as applied to phenotypic screening.

Introduction

For researchers engaged in target-based drug discovery (TDD), the single most important program decision is selection of a molecular target. If the target is not clinically relevant, there can be no drug. Protein targets are components of complex cellular systems that are both robust and adaptive. Isolation of a single target from a cellular system may facilitate assay design, or simplify data collection, but the relevance of active compounds to a systems context can be lost. This realization has led to a renewed interest in phenotypic drug discovery (PDD). The recognition that diseases can arise from defects in biological systems (rather than defects in molecular-target function) has spurred the shift toward phenotypic assays, which can be more physiologically relevant. With PDD, knowledge of the molecular target is not required. Instead, programs can be driven entirely by assessing functional endpoints and so may be more effective in identifying molecules that engage elements of signaling pathways or key regulatory nodes. This understanding has contributed to the emergence of PDD as an alternative and complimentary approach to TDD [1].

Screening can be an effective mechanism for drug discovery in either a TDD or a PDD setting. One crucial aspect of successful screening is the design of the compound library. The quality of the hits that emerge from screens will influence all subsequent project decisions. Simply put, a well-designed library should contain high-quality compounds, and any molecule with identifiable liabilities should be excluded from further consideration as it represents a resource-consuming distraction. Fewer hit liabilities translate into more efficient

We thank Thomas H. Large for his careful reviews and helpful suggestions during the preparation of the manuscript.
*Corresponding author:: K.L. Spear (kerry.spear@blueoakpharma.com)

 http://dx.doi.org/10.1016/j.ddtec.2017.05.001

project timelines and greater likelihoods of developing Advanced Leads that have better capabilities of successfully navigating hurdles in preclinical and clinical development.

The goal of this review is to present an overview of small-molecule library design principles as applied to phenotypic screening. The scope is not intended to be comprehensive but rather to provide a summary of recent developments with an emphasis on work published in the last two years.

A selective history of library design: ``What's past is prologue'' [2]

Prior to the early 1990s, screening campaigns relied on libraries that were more likely to be assembled than designed. These were typically idiosyncratic collections of corporate compounds created during the course of drug-discovery programs that could well date back decades. This began to change with the emergence of combinatorial chemistry as well as the sudden commercial availability of often structurally novel compound collections from academic groups in Eastern Europe. High-throughput screening (HTS) methodologies facilitated and accelerated this changing landscape by enabling large collections of molecules to be rapidly and systematically tested. The early exuberance associated with combinatorial chemistry and HTS led to a belief that screening millions of compounds was a desirable goal. Only in retrospect did it become apparent that the added expense did not lead to improvements in overall success rates. Consequently, awareness of the value of library-design principles began to emerge as a means of managing library size and compound quality.

Two important trends in the evolution of library design are illustrated in Fig. 1a. Over time, there has been a general reduction in the number of compounds screened [3], which has been accompanied by overall improvements in compound quality. A consequence of low hit rates from TPP screening campaigns resulted in most early libraries being designed using probabilistic strategies. This encouraged a culture of quantity, resulting in very large library sizes. Eventually, alternate approaches emerged that emphasized compound quality over library size. Efforts to devise more careful compound selection strategies had two complimentary objectives; to improve hit rates [4] or to reduce the attrition rates found during hit triage [4] and beyond [5]. One result was improved effectiveness of the screening process.

Our understanding of what it means to be a quality compound has also evolved (Fig. 1b). Early attempts to refine quality largely focused on improving the structural properties of the compounds. Initially, this meant optimizing structural diversity within a broadly defined chemical space as an alternative to assembling a random collection of compounds. This often relied on the experiences (and biases) of the medicinal chemists leading the process. Later, computational methodologies (e.g., Tanimoto-based similarity comparisons) were introduced to make the selection process more objective. One of the most significant developments in structurally focused design was the concept of the drug-like molecule [6]. This led to implementation of filtering protocols to triage compound collections based on physicochemical properties. It was also quickly extended to include concepts such as lead-like properties [6] and ''bad actor'' motifs [7] (e.g., reactive substructures and pan assay interference compounds). The goal was to increase chemical relevance by enriching libraries with compounds having desirable structural properties, effectively biasing libraries toward specific regions of chemical space.

A design strategy that began to gain traction in the late 1990s explicitly targets parts of biologically relevant chemical space (e.g., gene-family libraries [8]). Biologically relevant

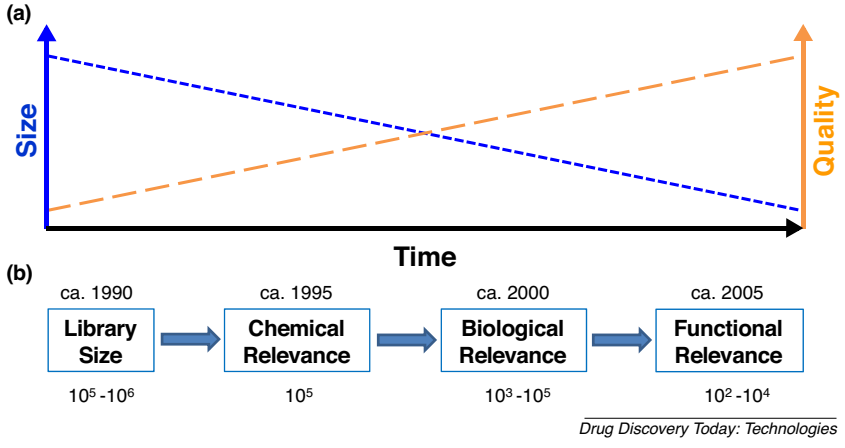

Figure 1. (a) Trends in library design illustrating the evolving role of size and quality over time. (b) Each box indicates a significant milestone in the evolution of design considerations that contributed to the trends exemplified in (a). Above each box we indicate an approximate year of adoption. Typical values for the kinds of library sizes appropriate to each design strategy are indicated below each box.