

New screening tools for lead compound identification

Tomi K Sawyer

The identification of promising lead compounds from high-throughput screens is still a very complex problem. A new high-throughput assay for identifying aggregation-based false positives could help.

The 'thrill of victory and agony of defeat' that many Olympians have experienced in their quest for the gold finds analogy in the challenges of lead compound generation and optimization to advance drug discovery. Despite the mind-boggling tools and 'core' technologies that exploit various aspects of chemistry, biology and *in silico* drug design to experimentally analyze thousands to millions of small molecules from chemical libraries or collections, there are hurdles to validating true lead compounds and eliminating false positives. In a paper published in this issue of *Nature Chemical Biology*¹, Feng, Guy, Shoichet and colleagues have introduced new high-throughput assays focused on pinpointing aggregation-based, false-positive lead compounds that effect nonspecific inhibition of a key enzyme target.

The 'Lipinski rule of five' characterizes properties, such as the number of hydrogen-bond donors and acceptors, molecular weight, and lipophilicity, that make a molecule more likely to be a successful drug². Feng *et al.* demonstrate that a rather high percentage of small molecules showed aggregation and promiscuous inhibition of enzymes tested¹, despite passing scrutiny in terms of the Lipinski rule of five for drug-like properties. This work is particularly important for both industrial and academic groups that use high-throughput screening of large chemical libraries or collections to identify initial lead compounds (frequently dubbed 'hits'). The authors show how a rather simple, robust enzyme assay was useful for exploiting detergent sensitivity to pinpoint aggregation-dependent inhibitors of lactamase. They also

present a more direct analysis of promiscuous aggregation (more exactly, particle formation) using dynamic light scattering. In addition to biochemical studies conducted on more than 1,000 small molecules to sort out such false-positive lead compounds, the authors describe a computational strategy to explore the behavior of such compounds so that predictive methods might be developed to further validate the process for lead compound identification (Fig. 1).

Lead compound identification is a complex process, with many decision-making points converging upon the selection of promising compounds to compete in the race for lead compound optimization and, ultimately, clinical candidate development. Initially, the focus was on having a large, chemically diverse library or collection of small molecules for testing to maximize the chances of finding a lead. Issues such as purity, chemically reactive functional groups and drug-like properties have been addressed over recent years to enhance the chances of finding a successful lead. Other issues, including physical properties such as solubility and aggregation, are also critical to the filtering of hits versus non-hits. Depending on the decision-making strategy, lowering the concentration of lead compounds used during screening may reduce the impact of physical properties such as poor solubility or high propensity for aggregation. However, lowering the concentration also increases the risk of missing new small molecules that have the potential to be optimized by iterative drug design and biological screening. Thus, these new methods¹ provide some experimental and computational strategies for limiting the interference from solubility and aggregation, while still maintaining a higher screening concentration.

Expanding the Lipinski rule of five to *in silico* and experimental methods may

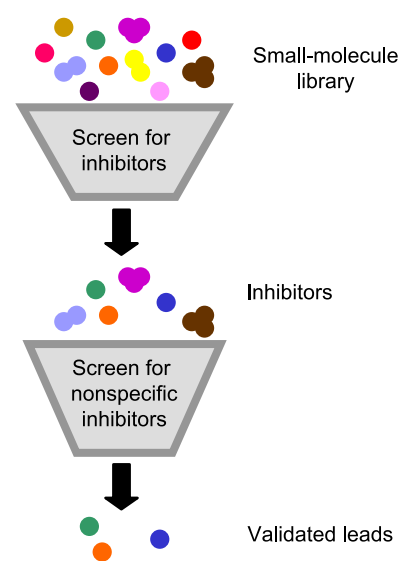


Figure 1 Illustration of high-throughput biochemical and computational screens to support lead compound identification with respect to validation.

enhance the predictability of lead compound generation. Sophisticated cheminformatics and related experimental molecular analyses to address relevant chemical and biological properties, including (among others) solubility, aggregation, bioavailability and metabolic stability, are critical components of the drug-discovery toolbox. Training scientists, like hopeful athletes, to use and understand the strengths and weaknesses of such technologies effectively will be an important aspect of the drug-discovery process within the scope of the multidisciplinary field of chemical biology.

1. Feng, B.Y. *et al.* *Nat. Chem. Biol.* **1**, 146–148 (2005).
2. Lipinski, C.A., Dominy, B.W. & Feeney, P.J. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).

Tomi K. Sawyer is at Ariad Pharmaceuticals, Inc., 26 Landsdowne Street, Cambridge, Massachusetts 01772, USA.
e-mail: tomi.sawyer@ariad.com