Bioorganic & Medicinal Chemistry Letters 23 (2013) 2844–2852

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

BMCL Digest

Learning from our mistakes: The 'unknown knowns' in fragment screening

Ben J. Davis [a,*], Daniel A. Erlanson [b,*]

[a] Vernalis Ltd., Granta Park, Great Abington, Cambridge CB21 6GB, UK
[b] Carmot Therapeutics Inc., 409 Illinois Street, San Francisco, CA 94122, United States

ARTICLE INFO

Article history:
Received 12 February 2013
Revised 1 March 2013
Accepted 8 March 2013
Available online 18 March 2013

ABSTRACT

In the past 15 years, fragment-based lead discovery (FBLD) has been adopted widely throughout academia and industry. The approach entails discovering very small molecular fragments and growing, merging, or linking them to produce drug leads. Because the affinities of the initial fragments are often low, detection methods are pushed to their limits, leading to a variety of artifacts, false positives, and false negatives that too often go unrecognized. This Digest discusses some of these problems and offers suggestions to avoid them. Although the primary focus is on FBLD, many of the lessons also apply to more established approaches such as high-throughput screening.

An expert is a person who has made all the mistakes that can be made in a very narrow field.

-Niels Bohr

Fragment-based lead discovery (FBLD) is now widespread throughout academia and industry and has delivered more than two dozen drugs into clinical trials. The approach entails screening small libraries of very small molecules, typically less than 300 Da. Because there are fewer possible fragment-sized molecules than lead-sized or drug-sized molecules, chemical space can be explored much more efficiently than by traditional high-throughput screening (HTS), even with a library of just a few thousand fragments. Fragments also make potentially better starting points for lead discovery because they contain fewer interfering moieties than HTS hits. The theory and practice of fragment-based lead discovery have been extensively reviewed in the literature as well as in five books.[1–5]

Clearly the approach works, but that is not to say it is easy. This Digest focuses on an area we believe is still insufficiently appreciated: the myriad pitfalls and artifacts that can befall a fragment-screening program. For the sake of brevity, we have chosen to focus on the problems that can hinder or derail an experimental fragment screening campaign; a full discussion of issues around fragment library design, virtual fragment screening, and fragment evolution is best dealt with elsewhere.

The first challenge facing FBLD is simply finding fragments which can be confidently identified as binding to the target. Having this confidence in the validity of a fragment hit is key, particularly since the risks of being misled by experimental artifacts are so much greater for fragments than when identifying tightly binding specific ligands.

Since fragments generally have low affinities for their targets—sometimes weaker than 1 mM—it is essential to have sensitive and robust methods for detecting weak interactions. In 1996 researchers at Abbott demonstrated that protein-detected NMR could be used both to discover low affinity fragments and inform how to link them; this paper is widely credited with popularizing the field.[6]

Today many techniques are used to identify fragments (Fig. 1),[7] each with its own strengths. Importantly, however, each of these techniques also has unique limitations. While expert users are generally aware of these and readily pick out the signal from the noise, newcomers are often deceived by spurious signals. This can lead to resources wasted following up on artifacts. In the worst cases—unfortunately all too common—researchers may never realize that they have been chasing false positives, and publish their results. At best, this is an embarrassment, with the researchers sometimes none the wiser. At worst it can cause other research groups to waste their own resources. Two recent reports have demonstrated that literature results are not nearly as robust as one would hope.[8,9] Although these were not focused on fragments, FBLD may be particularly prone to artifacts given its multidisciplinary nature and the number of neophytes in the field.

All the pitfalls described below are known, yet they continue to show up on a regular basis in internal programs and, unfortunately, in the literature. Thus, they can be categorized as what Mike Hann memorably christened *unknown knowns*: 'Those things that are

* Corresponding authors. Tel.: +44 1223 895555 (B.J.D.), +1 415 978 2159 (D.A.E.).
E-mail addresses: b.davis@vernalis.com (B.J. Davis), derlanson@carmot.us (D.A. Erlanson).

0960-894X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.03.028

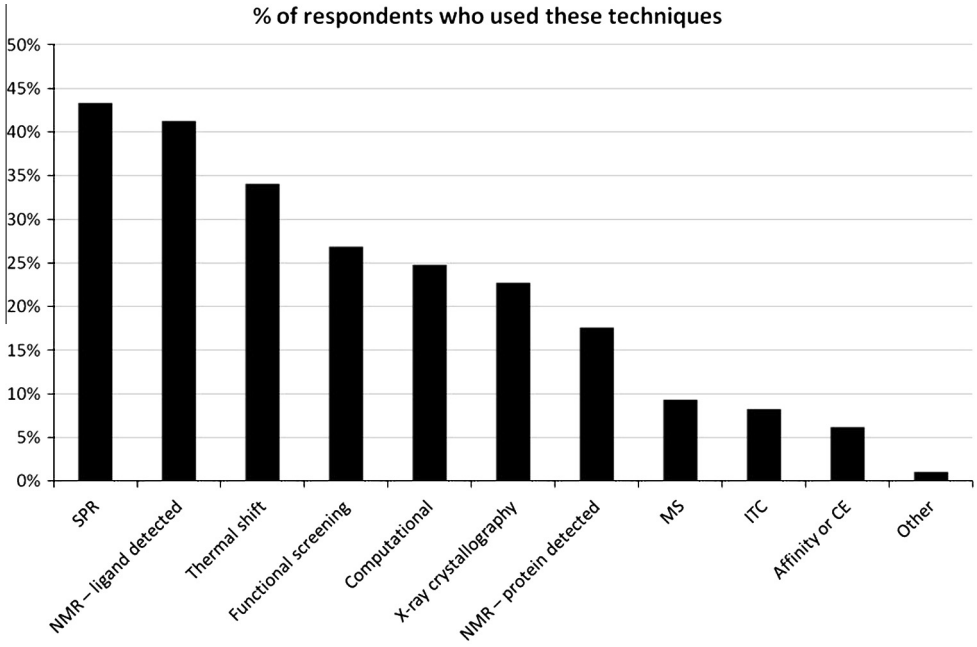

Figure 1. Methods to find fragments. These techniques were used to identify fragments, according to a poll on Practical Fragments in September 2011. There were 97 unique responses, and the average respondent used 2.4 different techniques.

known but have become unknown, either because we have never learnt them, or forgotten about them, or more dangerously chosen to ignore.'[10] It is our hope that this Digest can go some way towards transforming these pitfalls into known knowns. While most of the examples are taken from the literature, some have been reported in meetings, and others come from discussions with practitioners, who in some cases wish to remain anonymous; these are referenced as personal communications.[11]

Compound behavior. In order to be confident in the results of a fragment screen, you need to be confident in the quality of your hits. Before committing expensive chemistry resources, how do you guarantee that your fragment is what you think it is, that it remains what you think it is, and that it is actually doing what you think it is doing—i.e., making favorable interactions with a target?

Compound identity. Although it may seem trivial, it is always worth checking to make sure that the compound you think you have is really what you have. A fragment may simply be incorrectly registered in a database. More seriously, a purchased compound may not be what it says it is; both the authors have experienced this. If you are lucky, any follow-up chemistry will fail. If not, it might work, but not give you what you think you have. Depending on what your QC processes are, the error can propagate quite some way. In one example, a compound purchased for inclusion in a fragment library was found to be an isomer of the structure claimed by the vendor; worryingly, despite unambiguous data proving the catalog structure was incorrect, the vendor refused to remove the compound from sale 'because no-one else had complained' (personal communication). In another particularly notorious example, more than a dozen vendors were discovered to be selling the wrong isomer of the clinical stage kinase inhibitor bosutinib.[12]

Low-level impurities. Because fragment screening is typically performed at high concentrations, small amounts of reactive intermediates can wreak havoc: a 1% impurity will be present at 10 μM if a screen is run at 1 mM. Characterizing fragments by NMR and HPLC-MS is useful, but silent impurities can still sneak past. Metals are often used in organic synthesis, and can sometimes co-purify with compounds. For example, residual silver was found to cause a number of false positives in one assay,[13] as has gadolinium.[14] Similarly, several assays at Roche were found to be sensitive to low micromolar levels of zinc, a contaminant in a number of compounds.[15] In fact, zinc binding was even detectable by surface plasmon resonance. One of the projects was a fragment screen run at 250 μm, and the researchers note that fragment screens, 'which are typically run at much higher compound concentrations, should be more prone for false-positive signals from zinc and metal-contaminated compounds.'

It is possible for small amounts of potent impurities to contaminate a chemical sample during synthesis, purification, or compound management and plating. In one case, a fragment was contaminated with a trace of a potent generic kinase inhibitor, causing severely misleading results when that fragment was later screened against a kinase. Fortunately, in that instance, the use of orthogonal techniques identified the issue before significant resources were engaged (personal communication).

Compound stability. Compounds can degrade over time, sometimes quite unexpectedly: medicinal chemists generally strive to make molecules that will be stable in vivo, so it can be disconcerting to find that they fall apart during storage. One culprit is the commonly used solvent DMSO, which is a mild oxidant.[16] For example, pyrimidine derivatives such as compound **1** are colorless, but when dissolved in DMSO change color and oxidatively dimerize to form **2** and **3** within a matter of hours (Fig. 2).[17] Since compounds are often stored for months or more as stock solutions in DMSO, this degradation can become a serious issue. In order to

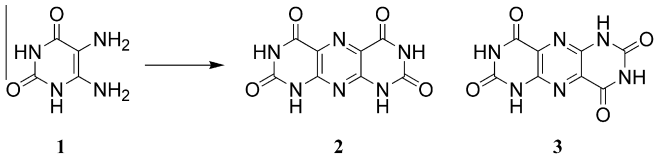

Figure 2. Unstable molecules. Compound **1** oxidizes in DMSO and dimerizes to form **2** and **3**. See text for details.