The art of the chemical probe

Stephen V Frye

Chemical biologists frequently aim to create small-molecule probes that interact with a specific protein *in vitro* in order to explore the role of the protein in a broader biological context (cells or organisms), but a common understanding of what makes a high-quality probe is lacking. Here I propose a set of principles to guide probe qualification.

"We may be lost, but we're making good time." —*Yogi Berra*

Though much has been said on this topic already[1–4], the concept of a 'chemical probe' that can be used to investigate biology has become embedded in the scientific lexicon without an accepted definition. Indicative of this lack of consensus is the finding that a panel of experts that reviewed the US National Institutes of Health Molecular Libraries' initial 64 probes found little correlation between the confidence they placed in probes and objective measures such as potency, solubility or 'rule of five' compliance[1]. Having recently spent seven years leading a chemistry group responsible for all small-molecule 'lead' discovery at a major pharmaceutical company, I have also struggled with the best means to reach a judgment about the ability of a small molecule to effectively probe biology. During this period the discovery chemistry group at GlaxoSmithKline had more than 200 leads versus >75 molecular targets move forward into lead optimization. Although drug leads and probes differ in important ways, exposure to this scale of chemical biology has shaped my perspective on the academic endeavor I have now joined. I believe the scientific promise at the interface of chemistry and biology will not be realized without greater appreciation for the appropriate properties of chemical probes. For example, staurosporine is the archetype of the 'probe gone bad', with more than 8,000 publications (search PubMed, http://www.ncbi.nlm.nih.gov/pubmed/, with 'staurosporine')—many attempting to utilize this promiscuous kinase inhibitor[5] as a tool to explore specific questions in signaling science[2,6]. While attempting to "purify the dialect of the tribe" (T.S. Eliot, *Little Gidding*) is an endeavor fraught with dangers, my intent is to propose a set of principles to be fulfilled by high-quality chemical probes.

Why not just follow the rules?

Though rule-based systems for judging chemical probe quality hold the promise of objectivity, they inevitably fail when the number of parameters against which quality is judged exceeds a handful, and when the applicability of the rules to the envisioned utilization of the probe is subject to debate. Even if we focus only on the molecular target-driven paradigm for probe discovery, a short list of quality probe parameters around which rules could arguably be arrayed include: (i) *in vitro* potency versus the target (and orthologs if more than one species will be 'probed'), (ii) *in vitro* selectivity versus paralogs of the target, (iii) *in vitro* selectivity versus other molecular targets known to interfere with the pharmacodynamic readout of interest, (iv) *in vitro* evidence of the mechanism of action and reversibility versus the target, (v) cell-based activity consistent with the presumed pharmacology of the target, (vi) cell-based assessment of cytotoxicity of the probe, (vii) chemical stability and reactivity assessment (electrophiles need not apply or must be credentialed with extensive mechanistic studies to demonstrate specificity), (viii) a structure-activity relationship profile suggestive of a specific interaction with the target, (ix) synthetic route or readily available source for the probe…and so on. With such a substantial list of criteria to meet, it is no surprise that agreeing on a rule-based system is difficult and resisted as unrealistic given the context-dependent significance of individual factors. Additionally, the difficulty of establishing such rules is most challenging for new areas of biology where the relevant set of assays to evaluate a probe may be ill-defined. Ironically, we can generally only build truly valid and objective rule-based parameters for biological systems for which quality probes already exist. When working at the frontier of biological understanding, the quality probe discovers its own path to validation.

Trust me, I'm a doctor

If rule-based systems are of limited prospective utility, are we left with only expert judgment as exercised by crowdsourcing[1], with some similarity to an 'I know it when I see it' (http://en.wikipedia.org/wiki/I_know_it_when_I_see_it, accessed 21 January 2010) approach? Unfortunately community standards are not consistent or particularly instructive when formulating a research strategy in a new area. I believe that reliance on relatively simple principles that can be clearly stated would have great value. If we assume that the primary intent of a chemical probe is to establish the relationship between a molecular target and the broader biological consequences of modulating the target, the following principles apply (see also **Box 1**):

1. Molecular profiling. A quality chemical probe has sufficient *in vitro* potency and selectivity data to confidently associate its *in vitro* profile to its cellular or *in vivo* profile.

2. Mechanism of action. A quality chemical probe has sufficient mechanistic data versus its intended molecular target to enable interpretation of its qualitative and quantitative effect (dose dependency) on a target-dependent action in either a cell-based assay or a cell-free assay that recapitulates a physiologic function of the target.

3. Identity of the active species. A quality chemical probe has sufficient chemical and physical property data to permit utilization in *in vitro* and cell-based assays with interpretations of results attributed to its intact structure or a well-characterized derivative.

4. Proven utility as a probe. A quality chemical probe has sufficient cellular activity data to confidently address at

least one hypothesis about the role of the molecular target in a cell's response to its environment.

5. Availability. A quality chemical probe is readily available to the academic community with no restrictions on use.

Note that principles 1–4 are all evidence-based assessments of the sufficiency of data and will necessarily depend on the context of use and current state of knowledge in the particular area under investigation. Credentials for a probe are therefore founded on data—not predictions based on the structure of the probe. Though a seasoned medicinal chemist's intuition can clearly identify potential problems with small-molecule structures (for example, see the insightful comments provided by G. Rishton in the supplementary table for ref. 1), final probes must have demonstrated suitability to explore biology. Because the envisioned users of chemical probes are highly skilled, specialized biologists, these principles ensure that their expertise and experiments are focused on the best tools available; they cannot be expected to have the knowledge to effectively validate that all the principles are satisfied for less well-characterized probes.

A principled approach

Alfred North Whitehead posited that "we think in generalities, but we live in detail." Without deviating toward creation of rules, it is useful to recognize that there are some approaches to addressing these principles that will have routine utility. For example, with the power of siRNA and shRNA technologies, the assays required to satisfy principles 1, 2 and 4 can potentially be validated with these more specific, generic techniques, and excellent advice on application of these technologies has been provided[3]. Additionally, the choices of selectivity assays required for principle 1 can be informed by an appreciation for the molecular targets likely to be influenced by a modulator of the target of interest—for example, a probe versus a protein kinase would minimally be profiled versus several hundred other protein kinases to be seriously considered as a quality chemical probe today[6]. Note that principle 1 does not require single-target specificity—it only requires that the probe has been profiled for selectivity versus potentially confounding targets. This is an especially difficult boundary issue for chemical biologists, as it depends on knowing what you don't know. Perhaps it is sufficient that broad profiling is carried out where there is a knowledge base that teaches this as a necessity, as in kinase probes. In less explored areas, a greater reliance on principle 2 for characterization may be necessary. Cocrystal studies of a small molecule and its putative target that rationalize structure-activity relationships (SAR) are especially useful in supporting mechanism-of-action assertions.

The availability of adequate control compounds for principles 1–4 likely requires exploration of both SAR and alternative templates—structurally distinct probes that fulfill principles 1–4 with the same (or similar) molecular target profiles surely reinforce confidence in conclusions reached[3]. Likewise, structurally similar compounds that differ markedly in their potency or selectivity, with consequential and consistent differences in functional effects in cell-based assays, make for excellent controls. Principles 2 and 3 are essential for eliminating interference compounds[7] and activity due to reactivity, impurities or degradation products. Principle 5 is included to ensure that putative probes are freely available for road testing by the broader community.

I have intentionally neglected treating probes discovered through phenotypic, high-content or pathway-based assays as exceptions. Though these approaches are a very important complement to the traditional reductionist molecular target paradigm, I view them as having satisfied principle 4 with the necessity of working back to the molecular targets and mechanistic level (principles 1 and 2) before compounds discovered with these techniques can be considered to be high-quality probes. Chemical proteomic approaches to making these challenging connections have developed significantly and have been recently reviewed[8].

Box 1 | Five principles of a quality chemical probe

Molecular profiling. Sufficient *in vitro* potency and selectivity data to confidently associate its *in vitro* profile to its cellular or *in vivo* profile.

Mechanism of action. Activity in a cell-based or cell-free assay influences a physiologic function of the target in a dose-dependent manner.

Identity of the active species. Has sufficient chemical and physical property data to interpret results as due to its intact structure or a well-characterized derivative.

Proven utility as a probe. Cellular activity data available to confidently address at least one hypothesis about the role of the molecular target in a cell's response to its environment.

Availability. Is readily available to the academic community with no restrictions on use.

Is the juice worth the squeeze?

Although these principles set a very high bar, they are achieved in contemporary science. For example, Nolen *et al.* recently described the discovery of two classes of small-molecule inhibitors of the actin-related proteins Arp2 and Apr3 where satisfaction of the five principles can readily be delineated, despite the rather modest potency of the compounds (half-maximal inhibitory concentration values of 10–30 μM), which would likely have violated a rule-based assessment[9]. Use of control analogs, cell-free functional assays, cell-based assays, and cocrystal studies to explain SAR and mechanism of action all support designation of the compounds described as high-quality probes. A second example that also illustrates the use of quantitative chemical proteomics has recently appeared: Huang *et al.* identified a probe of the Wnt signaling pathway and determined that the molecular target of the probe is tankyrase[10]. Through the use of inactive close analogs of the probe, siRNA knockdowns of putative targets and overexpression rescue experiments with wild-type and catalytically inactive tankyrase, the authors carefully cross-correlated the probe and its molecular target's function. As the specificity of the question being asked by a probe increases, the requirements for characterization also increase. Recent work in the kinase area illustrates this in two noteworthy papers on members of the AGC kinase family where carefully engineered mutant proteins are used to 'probe the probes'[11–13].

As others have noted, quantitative bibliometric measures of the impact of readily available high-quality chemical probes are similar to those of top academic researchers[2]. Therefore, with some trepidation I might suggest that the investment to produce a high-quality probe be compared to that required to create a high-quality academic researcher. Realization of these principles will certainly require medicinal chemistry resources and expertise that are often lacking from

academic screening centers, as well as more extensive profiling and cell-based evaluation than typically envisioned. Fortunately, the US National Institutes of Health Molecular Libraries Probe Centers Network effort has evolved toward more chemistry-driven follow-up of hits, and this will hopefully address the primary shortcoming of current probes[1]. Pragmatically, 'pre-probes' may be as far as an initial investigator's lab can carry a project, and subsequent work by others will be required for graduation to a high-quality probe. With these principles in mind, a chemical probe begins to attain a luster worthy of the aspirations of even the most adept chemical biologist. Not satisfied to leave a 'staurosporine' as their legacy[5], they will embrace the creation of high-quality chemical probes. ■

Stephen V. Frye is in the Division of Medicinal Chemistry and Natural Products and the Center for Integrative Chemical Biology and Drug Discovery, Eshelman School of Pharmacy, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
e-mail: svfrye@email.unc.edu

References

1. Oprea, T.I. *et al. Nat. Chem. Biol.* **5**, 441–447 (2009).
2. Edwards, A.M., Bountra, C., Kerr, D.J. & Willson, T.M. *Nat. Chem. Biol.* **5**, 436–440 (2009).
3. Weiss, W.A., Taylor, S.S. & Shokat, K.M. *Nat. Chem. Biol.* **3**, 739–744 (2007).
4. Kaiser, J. *Science* **321**, 764–766 (2008).
5. Tanramluk, D., Schreyer, A., Pitt, W.R. & Blundell, T.L. *Chem. Biol. Drug Des.* **74**, 16–24 (2009).
6. Karaman, M.W. *et al. Nat. Biotechnol.* **26**, 127–132 (2008).
7. Seidler, J., McGovern, S.L., Doman, T.N. & Shoichet, B.K. *J. Med. Chem.* **46**, 4477–4486 (2003).
8. Rix, U. & Superti-Furga, G. *Nat. Chem. Biol.* **5**, 616–624 (2009).
9. Nolen, B.J. *et al. Nature* **460**, 1031–1034 (2009).
10. Huang, S.-M.A. *et al. Nature* **461**, 614–620 (2009).
11. Frye, S.V. & Johnson, G.L. *Nat. Chem. Biol.* **5**, 448–449 (2009).
12. Okuzumi, T. *et al. Nat. Chem. Biol.* **5**, 484–493 (2009).
13. Cameron, A.J., Escribano, C., Saurin, A.T., Kostelecky, B. & Parker, P.J. *Nat. Struct. Mol. Biol.* **16**, 624–630 (2009).

Competing interests statement

The author declares no competing financial interests.