A crowdsourcing evaluation of the NIH chemical probes

Tudor I Oprea, Cristian G Bologa, Scott Boyer, Ramona F Curpan, Robert C Glen, Andrew L Hopkins,
Christopher A Lipinski, Garland R Marshall, Yvonne C Martin, Liliana Ostopovici-Halip, Gilbert Rishton,
Oleg Ursu, Roy J Vaz, Chris Waller, Herbert Waldmann & Larry A Sklar

Between 2004 and 2008, the US National Institutes of Health Molecular Libraries and Imaging initiative pilot phase funded 10 high-throughput screening centers, resulting in the deposition of 691 assays into PubChem and the nomination of 64 chemical probes. We crowdsourced the Molecular Libraries and Imaging initiative output to 11 experts, who expressed medium or high levels of confidence in 48 of these 64 probes.

A new kind of science

The US National Institutes of Health (NIH) Molecular Libraries and Imaging (MLI) initiative1 was designed as a cross-institutional effort to stimulate the discovery of new small-molecule tools for chemical biology. Supported by all 27 NIH institutes for a limited period (up to 9 years), the MLI pilot phase (2004–2008) stimulated the development of 10 national high-throughput screening (HTS) centers known as the Molecular Libraries Screening Centers Network (MLSCN). The MLI also encouraged scientists to submit, in a peer-reviewed process via the MLSCN, (i) assays amenable to medium and high throughput, (ii) new substances to be tested, (iii) new experimental data curated for the MLSCN and (iv) new experimental results—all to become freely accessible via the PubChem system (http://pubchem.ncbi.nlm.nih.gov/). Key to the open-access, free-access system of the MLI is the intention to complement rather than duplicate drug discovery efforts currently carried out in a pharmaceutical industrial setting. At the heart of this complementary process is the creation and screening of a compound library aimed at the identification of small-molecule modulators of biological functions that are suitable for basic research yet not necessarily ready to serve as starting points (leads) for drug discovery. It was envisioned that the MLSCN should “not [be] interested primarily in drug discovery but in the elucidation of biochemical pathways”1, by discovering small-molecule ‘tools’ or ‘probes’ to enable and support research in chemical biology. Having recognized that chemical biology relies on small-molecule tools to ‘probe’ the biochemical properties of gene-encoded proteins, the majority of which the pharmaceutical industry considers “undruggable”1, the NIH aimed not only to provide a suitable framework for the development of such “chemical probes” but also to encourage a cultural shift toward open, cross-disciplinary communication within multiple scientific and technological disciplines to facilitate probe discovery. These chemical tools are aimed at filling the gap between target identification, which is often regarded as an early step in drug discovery in which a certain macromolecule or process is ‘targeted’ for therapeutic manipulation via (often) small molecules, and lead discovery, which is an ulterior step that is often preceded by HTS and thus is based on ‘hit’ identification and optimization. Hence, the MLI framework set the stage for a new phase in academic science by broadly enabling this intermediate step of chemical probe discovery1. The MLI further strengthened the “community resource” aspect of the initiative by mandating rapid dissemination of all MLI-generated data, and ensured freedom to operate by eliminating trade secrets and intellectual property claims.

Pilot tools

Essential to this pilot phase was the creation of a chemical library: the Molecular Libraries Small Molecule Repository, or MLSMR. This library, which is accessible by querying ‘MLSMR’ in PubChem Substance, was subjected to MLSCN bioactivity screening between 2004 and 2008 for a total of 691 assays that were uploaded to PubChem. These included 242 primary, 402 confirmatory and 8 summary assays, which covered 171 targets and 29 phenotypic screens. These numbers continue to grow as pilot phase projects are completed.

Within this pilot phase, these NIH chemical probes were envisioned as having “adequate potency and adequate solubility to be useful for in vitro (e.g., cell-based) experimentation” (Box 1), but requiring further “chemical modifications […] to confer the selectivity, pharmacokinetic, and metabolic properties required for in vivo use”1. The criteria for a chemical probe evolved further, and currently require a certain potency (up to 100 nM), selectivity and aqueous solubility, and an improvement over existing probes for that target or assay (Box 1).

Given the interdependence between the library being screened and a particular assay, not all assays can yield new chemical probes. Yet MLSCN scientists were tasked with nominating suitable chemical probes, based on the emergent results. Beyond these general guidelines for chemical probes, the centers were given considerable autonomy in determining the quality and druggability

A list of affiliations appears at the end of the paper.
e-mail: toprea@salud.unm.edu or lskar@salud.unm.edu
Box 1 Chemical probe evolution

Chemical Probe 1.0 (2005). A probe is a chemical compound with activity in the primary and any secondary assays with adequate potency and aqueous solubility to be useful for in vitro (that is, cell-based) experimentation. The specifications for a probe are likely to vary depending on the target and may need to be set jointly by the assay provider and the MLSCN Steering Committee. An example of the specifications for a chemical probe that would define an endpoint for MLPCN activities would be: expected potency of 1 µM and solubility in 1% DMSO. Source: http://mli.nih.gov/mli/wp-content/uploads/data-sharing-and-ip.pdf.

Chemical Probe 2.0 (MLSCN 2007). An MLSCN probe is a compound which represents an improvement over the existing art. Supporting information is required showing currently available probes, their properties and how the new probe is clearly an improvement (first or best in class).

- **Potency:** Criteria for potency is context dependent, varying with the assay and target type, and current state of the art. In general, probes identified via biochemical assays are expected to exhibit potencies of < 500 nM and, more ideally in the 100 nM range. Those identified in cell based assays are expected to exhibit potencies at a level of < 1 µM, however in certain instances potencies in the 10 µM may be acceptable. Criteria for whole organism screening is less stringent.
- **Solubility:** Sufficient solubility in relevant solvents.
- **Availability:** The probe molecule should be accessible in amounts to allow advanced studies (15–20 mg), and protocols for its preparation or isolation should be made available.
- **SAR, mode of action:** (for example, evidence of binding to target, characterization of mechanism of action) and awareness of selectivity against relevant and/or related targets is expected.
- **Appropriate data on toxicity, permeability, etc. of probes are strongly encouraged.**


Chemical Probe 3.0 (2008). One of the primary goals of the MLPCN is the production of new and useful chemical probes for biomedical research. The minimum characteristics that a probe compound will need to have to be a useful research tool have been determined by the MLPCN Steering Committee and the MLP Project Team. According to the definition currently used by the MLPCN, these characteristics include <100 nM affinity, >10-fold selectivity against related targets and solubility in aqueous solutions (possibly including a low concentration of DMSO). Most importantly, a chemical probe must represent an improvement over existing probes for the designated target (https://mli.nih.gov/mli/wp-content/uploads/mlpcn_x02_faq.pdf). The NIH recognizes that whatever the characteristics of the probes, further modification may be necessary to produce compounds that are useful for in vivo studies. However, such additional efforts are outside of the activities called for in this funding opportunity announcement. It is the desire of the NIH that during the production phase, all probes and all related biological and chemical data are made available to all researchers. Source: http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-08-005.html.


necessary for a compound to be nominated as a probe.

As output, this $385 million initiative has produced “a compound repository, a database and technology” as well as 64 chemical probes so far, with additional probes pending approval, which subjected it to critical analysis. The initiative is also intended to nucleate discovery science throughout the research community by providing for individual investigators as well as centers to produce targets, probes, chemistry, new hardware and new software tools. In this Commentary, we have two aims. First, we will evaluate the quality of the NIH chemical probes, as perceived by a network of experts (“wisdom of the crowd” or “crowdsourcing”). Second, as the MLI enters a second phase (the Molecular Libraries Probe Production Centers Network, or MLPCN), we will discuss strategies for increasing the quality and usefulness of nominated chemical probes, including some of those implemented already by the MLPCN.

Molecular confidence or dubiosity

Molecular confidence or dubiosity (Box 2) is a subjective or objective property that can be attributed to small molecules in certain cycles of preclinical drug discovery. Its subjective aspect relates to what skilled medicinal chemists would describe as ‘gut reaction’—an intuitive or emotional response they experience when deciding which HTS hits, or confirmed actives, should be progressed further from a particular HTS assay. In its more subjective form, the scientific opinion of a medicinal chemist regarding the potential of a compound to be optimized into a clinical candidate drug is based on what Polanyi described as “tacit” knowledge, based on learning from experience. In its more objective aspect, molecular confidence or dubiosity is rooted in empirical observations, and it is usually derived by filtering HTS hits using a variety of parameters: (i) a predetermined list of substrates to remove reactive species; frequent hitters; aggregators or fluorescent compounds (as observed, for example, via flow cytometry); (ii) a predetermined list of toxicity filters; Lipinski’s rule-of-five (Ro5); Pfizer-like rules, central nervous system penetration; drug-like properties or lead-like properties; (iii) criteria for enrichment or removal of chemotypes based on similarity or dis-similarity to desired or unwanted chemicals; (iv) additional model-based filters for downstream drug safety issues such as human ERG binding (to rule out cardiac toxicity), cytochrome P450 inhibition (to reduce the risk of drug-drug interactions) and interactions with other potential antitargets; and (v) all of the above. Thus, confidence or doubt in a molecule stems from any number of undesired properties, and the potential to circumvent them within a reasonably small period (typically, less than six months), as judged by drug discovery scientists.

Crowdsourcing

A key challenge in evaluating the quality of the chemical probes resulting from the MLI initiative is a lack of skilled experts who can individually determine, with a reasonable expectation of success, the quality of these probes. In the absence of a completely objective way to evaluate these chemical probes, we used the ‘wisdom of the crowd’ to crowdsource (Box 2)
this evaluation to a team of 11 scientists with diverse backgrounds in small-molecule discovery, most of them experienced in industrial drug discovery and currently engaged in both academia and industry. Their task: Express their level of confidence in the 64 NIH chemical probes.

The University of New Mexico team, which did not vote, conducted exhaustive (as of October 10, 2008) queries for each chemical probe; the results were synthesized in tabular format and provided to the voting group. The ‘crowdsourced group’ (CSG) was then asked to evaluate these probes in the absence of any information related to the principal investigator or the institution that performed the assay and proposed the chemical probe. As a final step to ensure a blind evaluation, voters were asked not to use PubChem or any other database during the evaluation.

We anticipate that this crowdsourcing approach will accomplish two goals. It will (i) address the question “what is the quality of these chemical probes for the targets for which they were proposed.”

Confidence scores
Because molecular confidence is neither exact nor accurate, the NIH chemical probes evaluation was performed on a qualitative ranking of 0 (low confidence, high dubiosity) to 10 (high confidence, low dubiosity) (Box 3). Of the 64 probes, 16 (25%) were evaluated as having low confidence or high dubiosity (that is, their empirical rank scores are between 0 and 2); 16 probes (25%) were assigned medium confidence or dubiosity (that is, their empirical rank scores are between 3 and 4); and half the probes (32) were considered high in confidence or low in dubiosity (that is, their scores are between 5 and 10) (Fig. 1). Exact scores, together with comments inserted by the 11 members of the voting group, are provided in Supplementary Table 1 online.

Table 1 Distribution of scores

<table>
<thead>
<tr>
<th>Distribution type</th>
<th>Score average</th>
<th>Score median</th>
<th>Score above average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>2.50</td>
<td>2.97</td>
<td>1</td>
</tr>
<tr>
<td>25% quartile</td>
<td>3.78</td>
<td>3.49</td>
<td>3</td>
</tr>
<tr>
<td>Median</td>
<td>4.19</td>
<td>4.18</td>
<td>5</td>
</tr>
<tr>
<td>75% quartile</td>
<td>5.13</td>
<td>5.25</td>
<td>6</td>
</tr>
<tr>
<td>Maximum</td>
<td>7.84</td>
<td>8.00</td>
<td>11</td>
</tr>
</tbody>
</table>

Box 2 Definitions
Dubiosity (plural dubiosities). The state or characteristic of being dubious. Synonyms for dubiosity include doubtfulness, incertitude, skepticism, suspicion and uncertainty. In Ulysses, episode 16, James Joyce (1922) writes: “Possibly perceiving an expression of dubiosity on their faces, the globetrotter went on adhering to his adventures” (James Joyce: http://en.wikipedia.org/wiki/James_Joyce).

Crowdsourcing. First suggested by J. Howe26, crowdsourcing refers to an outsourcing ‘open innovation’27 model that uses distributed problem solving (http://en.wikipedia.org/wiki/Crowdsourcing). Its research and development applications are used by InnoCentive (http://www.innocentive.com/), YourEncore (http://www.yourencore.com/jsp/index.html) and NineSigma (http://www.ninesigma.net/), among others. In its simplest form, problems are crowdsourced when a large (often unknown) group of solvers (the crowd) is tasked with solving a problem or a particular set of problems. Often, the crowd ranks the possible solutions and selects the optimal ones. If crowdsourcing occurs within a business parameter, the crowdsourcer owns the solution, while the best-ranked solutions are compensated financially or otherwise28.

Wisdom of crowds. This concept, popularized by Surowiecki29, describes group decision-making based on the aggregation of independent, individual decisions, where the average decision is more accurate than any individual decision. The four elements of a wise crowd are independence, diversity of opinion, decentralization and aggregation. The survey on the NIH chemical probes fulfils the requirements of a wise crowd.

Box 3 The evaluation of confidence
By definition, low confidence would make us doubt the overall potential of a chemical probe related to (for example) selectivity, its probability to act as promiscuous binder and its future use as a chemical probe and starting structure for lead discovery. Each voting member was asked to assign a score, from 0 to 10, to each of the chemical probes according to their own judgment. However, they were under no obligation to score each probe; indeed, two voters did not score all chemicals (7.8% missing values). Higher scores imply higher dubiosity for the molecule, whereas low values imply higher confidence. To derive a qualitative evaluation of confidence, we applied the following steps and criteria:

1. Compute the average score across all voters (see Table 1); note that the average across all averaged scores was 4.54335.
2. Compute the median score across all voters (see Table 1).
3. Count the number of times a probe was scored above the average of all scores (4.54335)—that is, the score above average (see Table 1).
4. Empirically rank confidence by adding 3 if the score above average or the score median is above the 75% quartile, by adding 2 if the score above average or the score median is between the median and 75% quartile, by adding 1 if the score above average or the score median is between the 25% quartile and the median, and by adding 0 if the score above average or the score median is below the 25% quartile.
5. Qualitatively rank probes as follows: if the empirical rank is 5 or 6, high dubiosity (16 probes); if it is 3 or 4, medium (16 probes); if it is 0, 1 or 2, high confidence (32 probes). Further discussions are based on this qualitative rank.
For some probes, most of the CSG members were in agreement (horizontal lines that are predominantly one color in Fig. 1). Such consensus implies that some of these molecules should perhaps not have been nominated as probes (for example, where predominantly red and orange lines are found). Eleven probes ranked lowest in confidence (score of 6) and had a median score of 6 or higher. The CSG experts are in agreement that these structures should not have been nominated as chemical probes given that their bioactivity spectrum is quite nonspecific, thus making them useful only under carefully controlled biochemical conditions. Some of them share a polyphenolic dye character: alizarin yellow (PubChem CID 5281855), myricetin (CID 5281672) and the diketone CID 665013. A known fungicide (dazomet; CID 10788) contains both an amimal and a thiourea group, and an immunosuppressant drug (azathioprine; CID 2265) is known to hydrolyze in vivo to 6-mercaptopurine, a thiol-reacting group. Idarubicin (CID 151582) is a well-characterized cytotoxic CSG members expressed concern that ebselen (CID 3194) may be a false positive owing to its aliphatic ester and pyrimidiny1 sulfone combination, and CID 16725315 is an acylhydrazine. Finally, two probes contain what toxicologists agree are troublesome chemical moieties (an ammonium group, CID 5716367; and an N-oxide nitrile, CID 1756).

In some cases (for example, CID 665013), CSG questioned whether compounds with such chemical liabilities (polyphenol and diketone) are indeed an improvement over the state of the art, given that the intended target, HIV-1 reverse transcriptase, has already been targeted successfully by known drugs. Of the 16 low-confidence probes, 8 (12.5% of the total) are flagged for potential toxicity alerts and one is a pesticide (dazomet). Three (4.7%) of these probes are or have been marketed drugs (azathioprine, idarubicin and ebselen), and another four (6.25%) are perceived as drug-like. The low confidence score awarded to these 16 chemical probes is primarily rooted in their apparent lack of immediate chemical liabilities. Although 12 (18.75%) of them are aromatic amines, this potential mutagenicity risk was not deemed relevant at the chemical probe stage by the CSG.

Can we bring objectivity to this process?
One of the frustrating aspects of chemical probe evaluation is the lack of an objective function, or metric, by which their quality can be judged. How did the few empirical criteria for chemical probes correlate with the results of the voting group? Potency for the designated target and the estimated aqueous solubility (computed via ALOGPS20) were central to the NIH requirements (Probe 2.0, Box 1). Current NIH guidelines mandate that probes should have affinity below 100 nM (Probe 3.0, Box 1). Yet 45 out of 64 probes (70.3%) have activity above 100 nM, with 3 probes above 10 µM. Among high-confidence, low-dubiosity probes, 11 (17.2% of the total) have potency below 100 nM, and another 13 (20.3%) are between 100 nM and 1 µM. As 7 of the 17 low-confidence, high-dubiosity probes were reported as having bioactivity of 100 nM or lower, while 8 high-confidence and 11 medium-confidence probes have activity worse than 1 µM, it appears that target potency is regarded as a tunable parameter—that is, one that can be further optimized. Although 70% of the probes would not meet the current (Probe 3.0) activity cut-off, this aspect should be regarded in the context of the pilot phase (MLI), where fewer resources were available to meet this desirable goal. As the production
phase proceeds, no doubt more medicinal chemistry resources will be allocated for improving affinity and selectivity. However, potency was deemed as less important in both the MLSCN nomination process and in our own CSG evaluation. In this context, perhaps the NIH should consider reverting to Probe 2.0 criteria with respect to affinity.

We did not observe an apparent relationship between activity and computed solubility (Fig. 2). All probes except one low-confidence probe have estimated (ALOGPS) solubility above 1 µM, so this criterion was largely met, but it does not distinguish high-ranked probes from low-ranked probes. Even the Ro5 criteria (Fig. 3) fail to discriminate between low-confidence and high-confidence probes. 60 chemical probes violate none of the Ro5 criteria, and 2 violate 1 rule. Only 2 probes are outside the Ro5 parameters, which perhaps indicates the profound effect that the Ro5 criteria have had on the lead and drug discovery community. These two exceptions are peptidomimetic cathepsin L inhibitors that owe their low confidence score in part to an acylhydrazine moiety. By setting the polar surface area (PSA) value above 97 Å², 8 low-confidence probes are filtered, in contrast to 4 high-confidence and 4 medium-confidence probes. Twenty-eight high-confidence probes, 11 medium-confidence probes and 9 low-confidence probes (75% of the total) satisfy the condition PSA ≤ 97 Å². Based on a simple set of two-dimensional descriptors, and in the absence of supervised learning tools, we could find no criterion that adequately separates probes based on assessed quality.

**Lessons learned?**

Having completed the pilot phase (MLSCN), the NIH Molecular Libraries Program (MLP) moved into the production phase by launching the MLP in September 2008. The MLP aims to explore the interaction between small molecules and biological function at the molecular, cellular and in vivo levels, by encouraging biomedical researchers to combine large-scale screening with medicinal chemistry and informatics, in order to identify chemical probes to study the functions of genes, cells and biochemical pathways. The MLP has four Comprehensive Centers (the Broad Institute, the Burnham Institute for Medical Research, the NIH Chemical Genomics Center and the Scripps Research Institute) that will provide all three types of services (assay, cheminformatics and medicinal chemistry); three Specialized Screening Centers (Johns Hopkins University, Southern Research Institute and University of New Mexico) focused on specific assay technologies and informatics; and two Specialized Chemistry Centers (University of Kansas and Vanderbilt University), which will provide medicinal chemistry and cheminformatics support for chemical probe optimization.

Because low confidence was expressed in only 25% of the MLSCN chemical probes, we anticipate that through collective experience, chemical probe optimization will become an effective effort both in the Chemistry Centers and in the Comprehensive Centers as well. We also envision a dynamic MLSMR library, where protein-reactive electrophiles, warhead-containing agents, known frequent hitters and aggregators will gradually disappear as our collective knowledge of artifact-generating chemicals and their negative impact on chemical biology interactions increases. Indeed, the fact that despite more than a decade of cumulative experience from the pharmaceutical industry, such chemical structures are even found among the NIH chemical probes suggests that a rigorous evaluation of this problem, and how it influences the MLSMR library, is appropriate.

With few exceptions, each MLSCN center had a share of high-confidence and low-confidence chemical probes. Low-confidence probe nominations could perhaps be explained by a variety of factors: confidence in the chemists’ ability to overcome potential liabilities; no known promiscuity at the time of filing the probe; or perhaps knowledge that the perturbation of the intended biological target or process is not interfered with by these liabilities. MLP has already instigated a more rigorous review process for vetting both probe development plans and probes themselves before they are accepted.
From our experience, rigorous assessment of probe quality requires multiple communication rounds performed in an integrative manner involving: (i) chemical and drug informatics, including data mining, liabilities and virtual screening; (ii) biology, from cellular processes to *in vivo* observations wherever possible; and (iii) chemistry, with respect to structure-activity analyses, synthesis planning, solubility and purity information. First and foremost, both the biology underlying the assay and the chemistry on which the probes are derived need to be ascertained and confirmed in independent experiments, according to strict guidelines. Furthermore, any information regarding frequent hitter behavior, scaffold-specific chemical liabilities and potential/known off-target activities for the compounds of interest needs to be discussed in the context of structure-activity relationships. Finally, the biological activity of the putative probe should be confirmed by a third party, using fresh samples of the chemical in question. These types of cycles of probe testing at individual centers, and at the site originating the bioassay (the “target provider”), are now commonplace.

This type of further integration between biological experiments and chemical synthesis, together with intensive data mining and *in silico* tools, should lead to substantial improvements in chemical probe discovery. This approach is likely to be well complemented by some of the Challenge grants (RC1; [http://grants.nih.gov/grants/working/working_challenges/](http://grants.nih.gov/grants/working/working_challenges/)) and Grand Challenge grants (RC2; [http://grants1.nih.gov/grants/working/working_publications.html](http://grants1.nih.gov/grants/working/working_publications.html)) that have been launched by the NIH under the American Recovery and Reinvestment Act. The US National Institute of General Medical Sciences has also launched a Grand Opportunity (RC2) challenge related to the “development and application of statistical and computational data analysis methods for DNA sequence, variation, GWAS, genomic function, chemical biology and related genomic data sets” specifically for the development of “chemist- and biologist-friendly” tools, to support medicinal chemistry optimization by further improving the data analysis and integration components of screening ([http://www.genome.gov/27530674#2](http://www.genome.gov/27530674#2)).

**What are the measures of MLI success?**

Despite significant increases in research and development funds in the top 50 major pharmaceutical companies over the last decade, the number of new chemical entities is somewhere between 15 and 30 per year. A rigorous analysis indicates that the total cost of developing new drugs increased from $350 million per drug in 1991 to over $800 million per drug in 2003 (normalized for US$ in 2000). Though MLI and its pilot phase budget and superficially modest productivity were subject to industrial criticism, it should be noted that small-molecule discovery and innovation in both academia and the pharmaceutical sector can substantially benefit from a focus on innovative discovery science. Through this public initiative, all the results can be subjected to peer review through traditional academic mechanisms including publications and meeting presentations.

Although objective functions to assess confidence for both chemical probes and drug leads are not available, we offer crowdsourcing as a cross-disciplinary alternative that pools multiple levels of expertise from translational disciplines to provide a more rigorous chemical probe evaluation process within the MLPCN. We anticipate that the progressive removal of artifactual chemicals from the MLSMR, coupled with substantial improvements in integrative communication tools and data-to-knowledge transformative technologies, will result in significant improvements in chemical probe output. The three-year MLI pilot phase MLSCN was extended by one year, running in parallel with the MLPCN, giving the original ten centers a chance to complete projects that were in process. These additional probe projects are now being vetted according to the more rigorous standards instituted for the MLPCN.

The MLI has provided a blueprint for many (planned) activities in different countries. Several national chemical library initiatives are in progress, and even a European library is currently under consideration via a consortium. As a community experiment, the MLI has delivered some interesting chemical probes and has also altered the chemical biology community by providing an open-access data repository system (PubChem), a wide array of associated chemical and biological data for a large chemical library (MLSMR), and an increasingly larger number of assays. Not all chemical probes will stand the test of time, but as the NIH program continues, we expect that academic probe
discovery efforts will have an increasingly high impact in the public health sector. In the end, the MLI will be judged by the ultimate crowd-sourcing experiment when scientists around the world choose whether or not to use these chemical tools in their own research.

ACKNOWLEDGMENTS

We thank E. Bolton (NIH/PubChem) for promptly answering questions related to the MLI. B.K. Shoichet (University of California, San Francisco) voted on some of these chemical probes, as reflected in Figure 1 and Supplementary Table 1. C.A. Taylor, J. Feng, R. Yang, Y. Tang, D. Kuster and C.M. W. Ho (all from Washington University School of Medicine) contributed to the probe evaluation vote of G. Marshall, J. Benard (University of Dundee) contributed to the vote of A. Hopkins. This work was supported by NIH grant 1U54MH084690 (to the University of New Mexico).

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturechemicalbiology/.

Erratum: A crowdsourcing evaluation of the NIH chemical probes.

Tudor I Oprea, Cristian G Bologa, Scott Boyer, Ramona F Curpan, Robert C Glen, Andrew L Hopkins, Christopher A Lipinski, Garland R Marshall, Yvonne C Martin, Liliana Ostopovici-Halip, Gilbert Rishton, Oleg Ursu, Roy J Vaz, Chris Waller, Herbert Waldmann & Larry A Sklar


In the version of this article initially published, the phrase “Though MLI and its pilot phase budget and superficially modest productivity were subject to industrial criticism” originally cited reference 3, but should cite reference 2. The error has been corrected in the HTML and PDF versions of the article.
Erratum: A crowdsourcing evaluation of the NIH chemical probes

Tudor I Oprea, Cristian G Bologa, Scott Boyer, Ramona F Curpan, Robert C Glen, Andrew L Hopkins, Christopher A Lipinski, Garland R Marshall, Yvonne C Martin, Liliana Ostopovici-Halip, Gilbert Rishton, Oleg Ursu, Roy J Vaz, Chris Waller, Herbert Waldmann & Larry A Sklar


In the version of this article initially published, the phrase “Though MLI and its pilot phase budget and superficially modest productivity were subject to industrial criticism” originally cited reference 3, but should cite reference 2. The error has been corrected in the HTML and PDF versions of the article.