Public resources for chemical probes: the journey so far and the road ahead

Albert A Antolin[1,2,3,‡] (ID), Paul Workman*[2,3,‡] (ID) & Bissan Al-Lazikani**[1,2,3,‡] (ID)

[1]The Department of Data Science, The Institute of Cancer Research, London SM2 5NG, UK
[2]CRUK Cancer Therapeutics Unit, The Institute of Cancer Research, London SM2 5NG, UK
[3]CRUK ICR/Imperial Convergence Science Centre, London SM2 5NG, UK
*Author for correspondence: Paul.Workman@icr.ac.uk
**Author for correspondence: Bissan.Al-Lazikani@icr.ac.uk
‡ Authors contributed equally

High-quality small molecule chemical probes are extremely valuable for biological research and target validation. However, frequent use of flawed small-molecule inhibitors produces misleading results and diminishes the robustness of biomedical research. Several public resources are available to facilitate assessment and selection of better chemical probes for specific protein targets. Here, we review chemical probe resources, discuss their current strengths and limitations, and make recommendations for further improvements. Expert review resources provide in-depth analysis but currently cover only a limited portion of the liganded proteome. Computational resources encompass more proteins and are regularly updated, but have limitations in data availability and curation. We show how biomedical scientists may use these resources to choose the best available chemical probes for their research.

Our understanding of biological and disease mechanisms continues to advance at a greater pace than our capacity to transform discoveries into novel medicines [1]. For example, in oncology the cancer genomes of tens of thousands of patients have now been sequenced [1] and large-scale functional genomics experiments have been carried out with hundreds of annotated cancer cell lines [2,3]. Together these approaches have resulted in the identification of several hundred potential cancer drug targets but a reasonable estimate is that less than 10% of these have been drugged or are under clinical or preclinical investigation to generate chemical modulators [2]. There is also a great need for better biological understanding and new drug targets in a whole range of other diseases. Hence the requirement has never been greater for more and better tools to investigate biological mechanisms and to further validate and mechanistically investigate novel targets.

Chemical probes are small-molecule tools that have been appropriately characterized to enable the study of specific protein targets within complex biological systems [3–6]. They are highly complementary to biological tools such as RNA interference (RNAi) or CRISPR because they provide greater control of the extent and kinetics of modulation of the target protein. They generally inhibit (or otherwise perturb) protein function rather than removing the whole protein from the biological system [4]. Accordingly, it is often advised whenever possible to use both high-quality chemical and biological tools as orthogonal and complementary reagents for mechanistic studies and target validation [3].

Interestingly, in recent years enhanced CRISPR and small-molecule targeted protein degradation approaches are bridging the gap between chemical and biological tools [3]. However, in addition, chemical probes also provide valuable evidence on whether the target can be modulated pharmacologically and thus facilitate the development of new small-molecule drugs [4].

Future Med. Chem. (Epub ahead of print)

The importance of selecting chemical probes appropriately

As an example of the value of high-quality chemical probes, bromo- and extra-terminal (BET) bromodomains were a poorly characterized protein family until JQ1 and I-BET were published and made openly accessible as first-in-class chemical probes [3,7]. This empowered unprecedented exploration of the BET family and subsequent drug discovery and development.

Conversely, the use of claimed chemical probes that are poor-quality or flawed can have devastating consequences for biomedical research [8]. A widely acknowledged, painful example is iniparib, which was developed as a first-in-class PARP1 clinical candidate until it failed in Phase III clinical trials [3,9]. This clinical trial failure cast a shadow over the development of PARP inhibitors in the clinic for some time, until it was discovered that iniparib is actually not a PARP inhibitor but rather is a nonspecific covalent cysteine modifier [3,9]. Yet despite several *bona fide* PARP inhibitors such as olaparib having now been finally approved by the FDA as cancer therapeutics, iniparib continues to be sold by many compound vendors as a PARP inhibitor. This example illustrates how online vendor catalogs should not be used as a reliable, up-to-date and scientifically accurate source of information to select the best available chemical probes. Iniparib also exemplifies how poor-quality probes lead to scientific distractions and can delay access to life-saving medicines for patients.

Neither chemical nor biological tools are exempt from limitations which must be taken into consideration before selecting and using them. For example, it is widely acknowledged that biological tools, including genetic reagents and antibodies, can exhibit off-target effects [4]. Accordingly, the use of more than one RNAi oligonucleotide and the confirmation of effects on target protein expression are widely-used to mitigate risk [3,4]. Similarly, the off-target effects of CRISPR are increasingly appreciated and refinements of this technology and use of careful controls to reduce or mitigate the off-target rate are being extensively investigated [10,11].

Chemical probes can also have off-targets [12]. These can range from hitting a small number of additional related proteins through to promiscuous activity across the proteome. In the most egregious examples, biological and biomedical experimentation has been carried out with chemical compounds, claimed to be probes, that are indiscriminately reactive chemically or that are act broadly as a result of insolubility/aggregation and assay interference. Therefore, target selectivity must be carefully defined, in addition to on-target potency. It is recommended that more than one probe from different chemical classes (chemotypes), ideally together with matched inactive controls, should be used whenever possible to reduce the likelihood of off-target effects [3]. Moreover, cell permeability is essential. Demonstrating target engagement in cells is also strongly recommended [3–6].

Of serious concern is that there is clear evidence that best practice guidelines are frequently not followed [3,8]. Although useful as an early pathfinder compound, LY294002 is a classic example of a nonselective chemical inhibitor that should no longer be used for the study of specific protein targets [3,8]. Alongside wortmannin, LY294002 was the preferred chemical inhibitor available to study PI3 kinases in the 1990s but it has been clear for two decades that this compound inhibits a plethora of kinases and other targets [3,8,13]. Unfortunately, LY294002 continues to be widely used in scientific publications (at a rate over the last five years of 500 citations per year according to PubMed), often without other chemical tools or biological controls, and the observed effects commonly being attributed exclusively and erroneously to PI3 kinases [8]. Reliance on LY294002, rather than better PI3 kinase probes such as the pan-Class I inhibitor pictilisib and more isoform-selective inhibitors, not only produces misleading results, but also further propagates the use of this poor reagent. Similarly, the promiscuous protein kinase inhibitor staurosporine has been employed widely and inappropriately [3–5].

The widespread uncritical selection and use of chemical probes based on citations in the scientific literature or standard online search engines causes major problems because these sources are usually biased towards older and commonly flawed small-molecule inhibitors. Reliance on information in vendor catalogs alone can also be dangerous. The undiscerning application of claimed chemical probes produces misleading and invalid results that can be extremely costly and contribute to the lack of robustness of biomedical research. So how can poor quality or flawed small-molecule inhibitors be spotted and avoided, and better tools identified and used with confidence?

Fitness factors to assess chemical probes

To be valuable to study specific proteins in complex biological systems, chemical probes should at least satisfy minimal fundamental criteria [3–8]. A fit-for-purpose chemical probe should bind to the intended biological target with sufficient potency to be able to modulate its activity at biologically achievable concentrations. If the target is intracellular, the probe must be able to penetrate inside cells. Also important is that the chemical probe
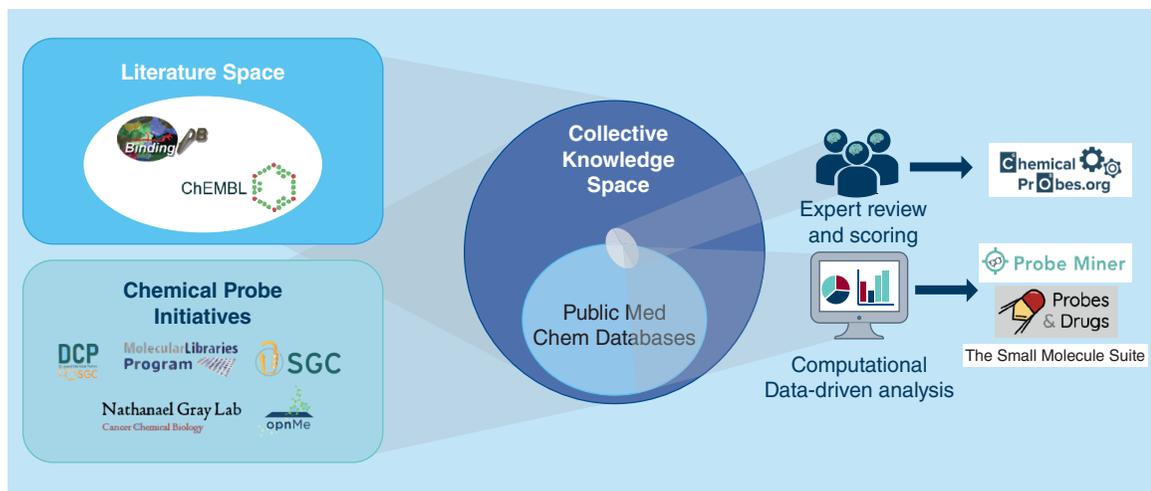
Figure 1.   Various chemical probe resources and the relationships between them. The collective knowledge on chemical probes comprises the biomedical research literature – curated in part by medicinal chemistry databases such as ChEMBL and BindingDB – and data contained in chemical probe initiatives and programmes, such as the Structural Genomics Consortium or the Donated Chemical Probes, that are not necessarily published in the biomedical literature. The shaded areas illustrate how both the Literature Space and Chemical Probe Initiatives both form the Collective Knowledge Space of chemical probes. These resources are also used by experts to peer review chemical probes and deposit this information in expert-review resources such as the Chemical Probes Portal. Other resources, such as Probe Miner and Drug & Probes, enable available data to be mined to provide objective, data-driven chemical probe assessments. For more details please refer to the main text.

should be sufficiently selective to ensure that the observed effects are due to the modulation of the intended target. Some chemical probe developers have set minimum values for these fundamental properties (e.g., *in vitro* potency <100 nM, >30-fold selectivity against other subfamilies, on-target cell activity <1 µm) [4,14]. However, there is a great shortage of probes that meet such stringent criteria. Accordingly, pragmatism dictates, and experience shows, that probes that do not quite meet these criteria may still be used in early studies as 'preprobes' and pathfinder compounds. This is provided that we know about and can mitigate their limitations, especially around off-targets [4,5]. Nevertheless, the aim should always be to seek and use very high-quality probes for the target of interest.

The important concept of 'fitness factors' was introduced to aid the evaluation of chemical probes [4]. These factors include especially potency, selectivity and cellular activity, together with physicochemical properties – such as solubility and stability – and suitability for the biological context [3,4]. Particular attention is devoted to selectivity, which should ideally have been characterized not only for close family members but also across a wider panel of diverse proteins. Of importance is the need to see evidence of concentration-dependent target engagement and modulation in cells that can be linked to biological effects in a 'pharmacological audit trail' [15,16]. Overall, it is important to evaluate carefully all fitness factors before selecting the best chemical probe(s) for a biological experiment [3,4]. Where limitations exist, experiments must be designed and controls included to address these.

Identifying and analyzing the information required to appropriately assess and select chemical probes for a given experiment can be very challenging, particularly for biologists who are not experts in chemical biology, pharmacology or drug discovery. The relevant data are commonly scattered across many publications. Fortunately, help has been forthcoming during the last five years through the emergence of a number of public chemical probe resources (Figure 1). Some of these resources provide general guidance and expert-curated advice on specific chemical probes for particular protein targets. Others facilitate a data-driven and quantitative digestion of the complex data. Here, we review current public chemical probe resources and discuss their strengths and limitations. We illustrate how, while they all have scope for improvement, the availability of these resources greatly empowers researchers to be objective and critical in their selection and application of chemical probes. Moreover, we stress the importance of getting the key messages about choosing and using chemical probes across to the general bioscience community – not just chemical biology and drug discovery specialists – and of making the chemical probe resources user-friendly to nonexperts [3].

Expert-curated chemical probe resources
The Chemical Probes Portal: expert assessment & advice

In 2015, a large group of chemical biology experts launched the Chemical Probes Portal (www.chemicalprobes.org) – a community-driven, public resource based on expert reviews, designed to facilitate the informed selection of appropriate chemical probes and reduce the use of flawed compounds [8]. The underlying principle of the Portal is that it relies on the recommendations of chemical probes for specific targets that are provided by experts. For acceptance into the review process, a proposed probe must have been published in the scholarly peer-reviewed scientific literature, deposited in PubChem (https://pubchem.ncbi.nlm.nih.gov) and have data supporting its validation that include reasonable evidence that the probe is acting selectively on the target of interest in cells. Expert reviewers then score the chemical probe for its use in cells and in model organisms using a four-star rating system [8]. The experts also provide useful comments about the strengths and limitations of each chemical probe and summaries of important properties such as selectivity. In addition, the Portal has useful links to other resources and data sources such as primary literature references, vendor websites and gene databases.

The Chemical Probes Portal is designed to facilitate user searches that can be based on either the probe or the biological target of interest. Illustrating the strength of this resource, searching for the 'PARP1' target (carried out on 8/3/2019) pointed to ten *bona fide* PARP inhibitors such as the genetically-targeted cancer drug olaparib (www.chemicalprobes.org/olaparib). Moreover, reviewer comments and target annotations clearly explain that this chemical probe is not specific for PARP1 but also inhibits closely-related family members such as PARP2 and 3. Alternative PARP inhibitors are recommended. Such information from experts facilitates a realistic understanding of the strengths and limitations of each chemical probe and suggests actions that mitigate any weaknesses.

Since its release, the Chemical Probes Portal has grown to cover more than 190 chemical probes, of which at least 125 have achieved a rating of three stars or better, acting on more than 180 biological targets. The Portal has brought together a large community of more than 140 chemical biology experts that serve as an independent Scientific Advisory Board and are the principal sources of probe proposals and reviews. In addition to reviewing current probes, the Portal flags 'historical' compounds that have been superseded by better probes and should not be employed today, as well as small-molecule inhibitors that do not qualify as chemical probes due to chemical reactivity or poor selectivity. Today, there are more than 200 historical compounds registered in the Portal, raising awareness of their drawbacks and advising against their use.

In addition, the Chemical Probes Portal has web pages with other useful advice, such as the choice of concentrations and doses of chemical probes to use in experiments and the importance of demonstrating target engagement by the probe. There is an important cautionary note stating that "*assays that demonstrate a chemical probe impacts cell proliferation and/or cell death are rarely specific enough to provide the type of validation for selectivity that we seek.*"

Over the past five years, the Chemical Probes Portal has arguably become the expert-curated resource most widely accessed and with the largest number of peer-reviewed chemical probes to help researchers select chemical probes, attracting more than 5000 users per month. However, as for any resource, it is not exempt from limitations. Currently (meaning at the time of writing throughout), the Portal covers only probes that are published in peer-reviewed scientific articles and some would argue for inclusion of those that have not necessarily been published in the scholarly literature but have been otherwise expert-assessed, such as those from the Donated Chemical Probes Initiative (see below). However, the key challenges faced by any expert-curated resource are primarily the scale of coverage and the ability to keep pace with the rapidly changing literature. At this time, the portal covers mainly kinases, GPCRs, phosphodiesterases, epigenetic proteins and BCL2 family proteins and it would be extremely valuable to the research community if the breadth of content was expanded to include a larger proportion of the liganded proteome. In addition, it is important that the probe recommendations remain timely and are updated when existing current probes are superseded by better ones. With further enhancements planned, the Chemical Probes Portal is likely to remain as an essential, trusted and user friendly 'go-to' source for guidance on chemical probes – especially since it is difficult if not impossible to replace the considered advice of experts [17].

Of note is that the American Association of Cancer Research's Chemistry in Cancer Working group has a brief but useful web page on chemical probes, pointing researchers to other resources, key review publications and also a list of vendors and services. This helps investigators determine whether a particular compound is selective against a target of interest or will exhibit target engagement in cells – noting that researchers (we would add academic biologists in particular) may find it challenging to access broad profiling technologies (www.aacr.org/Membership/PAGES/SCIENTIFIC%20WORKING%20GROUPS/CICR-GENERAL-RESOURCES.ASPX).

Resources from chemical probe developers

Several organizations have released public resources to facilitate access to key information about the chemical probes that have been developed under the auspices of a particular large-scale research programme (Figure 1). To our knowledge, the earliest example was the NIH Molecular Libraries Program (MLP) that was launched in 2004 with the aim to expand the availability, flexibility and use of small-molecule chemical probes for basic research [18]. The MLP was one of the first programmes that brought small-molecule screening platforms, which had hitherto been largely operational only in the pharmaceutical industry, to academic laboratories. This decade-long project produced 375 small-molecule chemical probes across a wide diversity of biological targets – including kinases, GPCRs, GTPases, proteases and RNA-binding proteins [19]. Data on these probes can be publicly accessed via PubChem [20] and the Probe Reports from the MLP website (www.ncbi.nlm.nih.gov/books/NBK47352/). The early phase of the programme also contributed to a fruitful discussion in the chemical biology community about how chemical probes should be evaluated, and illustrated that hits from screens were rarely chemical probes without extensive chemical optimization [4,21]. Despite some initial criticism of the quality of chemical probes being produced [21], more than a decade later it seems clear that the identification of several first-in-class chemical probes for under-studied proteins, together with their translational impact on our understanding of several pathological processes, have shown the value that the MLP provided in the evolution of the overall ecosystem [22].

The Structural Genomics Consortium (SGC) is a well-known precompetitive, open-access, public–private partnership that develops chemical probes. Established in 2003, the SGC's original and achieved aim was to obtain large numbers of crystal structures of human proteins across a broad representation of families and place them in the public domain [23]. In 2010, a collaboration between chemical biologists at the SGC and the Dana-Farber Cancer Institute published the discovery of the first bromodomain chemical probe, JQ1 [24]. JQ1 was made openly accessible to the community and has had a profound impact both on understanding bromodomain biology and on the number and speed of new bromodomain inhibitors entering clinical trials [25]. Since then, the SGC has participated in the discovery and public release of over 65 chemical probes against more than 65 biological targets, often representing the first available chemical probe for a protein target and with a special focus on epigenetics [7]. The SGC policy of releasing the probes to the scientific community to facilitate broader understanding of their biological effects has proved highly effective. The data generated on chemical probes by SGC and its partners can be accessed openly through the SGC website (www.thesgc.org/chemical-probes). The SGC also generates Target Enabling Packages (www.thesgc.org/tep).

The Gray Laboratory Probes resource currently lists 38 small-molecule chemical probes and targeted protein degraders against 38 targets (https://graylab.dana-farber.org/probes.html). Useful information for each chemical probe can be downloaded in a PDF file as well as a list of 96 current best-in-class tool inhibitors for 90 kinases based on review of publications and also patents, which are another source of information.

Since the late 2000s, in an effort to counteract the increased competition/duplication and decreased R&D productivity, pharmaceutical companies have started to develop open innovation strategies [26]. Many such open innovation platforms have emerged, some focused on incorporating external compounds [26], others on solving specific challenges [27], and yet others releasing internal compounds screened against, for example, neglected tropical diseases [28]. More recently, some pharmaceutical companies have added chemical probes to their open innovation strategies. Boehringer Ingelheim has launched the opnMe Open Innovation Portal to freely share selected molecules with the scientific community [29]. Since 2017, opnMe has released over 30 chemical tools against more than 30 targets, and comprehensive information about their selectivity and other properties can be accessed online at the opnMe website (https://opnme.com).

In 2018, a group of seven pharmaceutical companies associated with the SGC undertook each to donate ten high-quality chemical probes to the scientific community under the auspices of the Donated Chemical Probes initiative [30]. This was a result of the increasing appreciation that the public release of previously inaccessible industry compounds and data can provide value both to the wider scientific community in general and also to the companies involved, particularly regarding target validation in disease models [31]. The probes are required to meet typically stringent criteria for modern high-quality chemical probes. A major benefit of the Donated Chemical Probes initiative is the release of the detailed data associated with each chemical probe, of the type that are often unavailable or located in patents or supplementary files. Also very important are the data provided on matched, 'inactive' (defined as at least 100-times less active) control compounds [30]. Use of active/inactive compound pairs

is considered part of the best practice approach alongside application of two structurally distinct active chemotype compounds [3,4]; however, matched inactive compounds are only rarely available and applied (e.g., see [32] and [33]).

Clear instructions are provided on the use of the Donated Chemical Probes. They are required to meet specific high-quality criteria including potency, selectivity and cellular activity and are profiled extensively – against >500 assays, including broad panels of pharmacologically-relevant targets, such as GPCRs, kinases, ion channels and proteases, to identify off-target activities. Several probes have sufficient druglike character to make them suitable for *in vivo* use in animal models. Probes donated by a company are scrutinized by other partners and also by an external expert committee to ensure provision of high-quality data that are made openly accessible on the public website (https://opensciencorobes.sgc-frankfurt.de). Supplies of the Donated Chemical Probes and, importantly, the corresponding inactive controls, can be requested on the website through an Open Science Trust Agreement. Of note is that the Donated Chemical Probes may not at the time of release be published in the peer reviewed literature – which some would see as an advantage in terms of early access, whereas others might argue that formal peer review at a scholarly journal remains important. The information package described above is not currently available for every Donated Chemical Probe, but this will build with time.

Although commercial compound vendors are often at fault for providing uncritical information in their catalogs and selling poor-quality and flawed compounds as chemical probes, it should be noted that some suppliers do have useful advice on their websites (e.g., www.tocris.com/resources/tocris-blog/using-chemical-probes-brief-guide; www.promega.co.uk/resources/pubhub/features/advancing-biomedical-research-with-quality-chemical-probes).

Overall, more than five different websites from chemical probe developers (as distinct from vendors) can be accessed online, illustrating the increase in programmatic chemical probe initiatives that are strongly enabling for the scientific community. Despite their value as repositories of information, no single chemical probe developer can discover a chemical probe for every target and each website alone covers only a limited number of chemical probes against a limited number of targets. Therefore, for the average scientist trying to select chemical probes for their target of interest, more centralized resources would be very helpful.

Another very important limitation of these expert-curated resources is the form of the valuable data that they contain. Currently, the important information hosted in these resources is reported differently between the various resources, in different files and with different formats – which makes it difficult for researchers to reach a general overview and prioritization for a given target. No programmatic access facility or bulk download of data is currently available to enable users to integrate and compile this valuable information. Such aggregated and integrated data would be highly useful for multiple applications both in chemical biology and beyond, especially in view of the still low abundance of highly characterized compounds and in particular to enable exploitation of the large datasets using Big Data approaches, such as machine learning and AI [34].

Quantitative, data-driven, computational chemical probe resources
Data, datasets & early algorithms
The development of free-to-use computational resources relies on access to large-scale bioactivity data in the public domain. Since the 2000s, large amounts of these data have become publicly available in resources such as PubChem, BindingDb (www.bindingdb.org), IUPHAR/BPS Guide to Pharmacology (www.guidetopharmacology.org) and ChEMBL (www.ebi.ac.uk/chembl) [20,35–38]. Together, these resources host data for millions of compounds against thousands of protein targets, and they underpin many modern chemoinformatics tools, resources and applications (Figure 1). Note that SciFinder (www.cas.org/products/scifinder) is a commercial, curated chemical/pharmacological resource that is used extensively by medicinal chemists [39], but it is not freely available and cannot be integrated with public chemical probe resources in an open innovation setting.

Before 2018, only very few small and focused computational efforts had attempted to use public databases to score or prioritise chemical probes [40,41]. As an example, compounds available in ChEMBL – itself an invaluable curated public database of chemical and bioactivity bioactive data that are manually abstracted from the primary published medicinal chemistry literature – were filtered by adapting the potency and selectivity criteria defined by the SGC to identify 407 best-in-class compounds for 278 protein targets [41]. This initial use of computational methods for chemical tool prioritisation paved the way for subsequent, larger scale studies.

The first comprehensive study was led by Novartis researchers, who performed a meta-analysis of bioactivity data and supervised machine learning to derive a 'tool score' that could be used to prioritize compounds with higher selectivity for use in future phenotypic screening [42]. The team integrated large-scale, heterogeneous data from both private and public sources, incorporating expert weighting to compute the tool score as a product of potency and

selectivity. The tool score was used to assess 384 compounds that were further experimentally screened, and a list of 68 nonselective compounds that were frequently used in the literature was disclosed [42]. Despite the importance of this initial comprehensive analysis, the use of proprietary databases and expert weighting meant that its usability by the wider community, lacking access to these resources, was very limited. Moreover, the analysis was performed at a specific point in time, whereas data are constantly accumulating. Hence there is clearly an important need for public, quantitative, data-driven, computational chemical probe resources that could maximally benefit the whole scientific community (Figure 1).

Probe Miner

Probe Miner is a public resource for the objective, quantitative and data-driven assessment of chemical probes [43]. As a data source, Probe Miner uses information from several public chemical pharmacological resources, such as ChEMBL and BindingDB, that are in turn integrated within the canSAR knowledgebase (https://cansarblack.icr.ac.uk) [44]. Probe Miner implements six different scores inspired by the fitness factors [4] – namely, target potency, target selectivity, cell potency, structure–activity relationships (SAR), inactive analogs and pan-assay interference (PAINS) – using a statistical methodology that allows probe comparison and ranking. At the time of the launch and publication, a total of 1.8 million compounds were analysed and triaged. Of these, >300,000 active compounds were assessed against >2200 human targets using the six chemical probe scores [43].

To empower the community to use this data-driven chemical probe assessment and facilitate access to non-chemical biology experts, Probe Miner was published alongside the launch of a public, user-friendly, interactive website (https://probeminer.icr.ac.uk). The resource is designed to be searched by biological target as the aim is to provide the researcher with a quantitative assessment of all the compounds for the target of interest that are available in public databases, together with useful data summaries [43]. The compounds are prioritized using a 'global score' that integrates the six different scores with predefined weightings. However, the weightings can be customized according to the individual researchers' preferences and needs. Minimum-quality thresholds of potency (<100 nM), selectivity (tenfold against any other protein) and cell potency (<10 µM) are used to highlight potential probe limitations. Clearly visible icons are displayed to draw attention to the strengths and limitations of each chemical probe. Chemical probes that have also been reviewed by experts in the Chemical Probes Portal are highlighted and web links are provided to encourage the use of both resources in conjunction. If desired, users can dive deeper into the data and navigate to chemical probe websites where all the data used to calculate the probe scores can be found. In addition, expert users can download all the data in Probe Miner for their own analysis [43].

When comparing the computational Probe Miner scores with expert reviewed ratings in the Chemical Probes Portal, it was found that in 67% of cases Probe Miner ranked the probes recommended in the Portal in the top 20 for the target of interest [43]. Disagreement is mainly caused by public medicinal chemistry databases not containing key data. This is often because the chemical probe in the Portal was originally published in a journal that is outside the medicinal chemistry literature abstracted by ChEMBL, whereas it is known to the expert Portal reviewers. Another source of disagreement comes from when more comprehensive or up-to-date data are available in the databases which have not yet been identified or curated by the Portal reviewers. A definite advantage of the Probe Miner is that when a user carries out a search using the Probe Miner web resource the results and rankings returned will always be based on the latest update of the underlying data. Overall, the comparisons between Probe Miner and the Chemical Probes Portal demonstrate the complementarity between the two resources and value of using the two alongside each other for chemical probe assessment and selection.

Probes & Drugs

Probes & Drugs (https://www.probes-drugs.org) was developed as a tool for the exploration of bioactive compound space by bringing together data on focused libraries of bioactive compounds with data on commercially available screening libraries, while at the same time offering an intuitive and powerful chemical filtering and visualization system [45]. Probes & Drugs is assembled from 29 established public and commercial chemical libraries, with special attention being given to chemical probes and drugs. The compound libraries are also enriched with additional data, such as the target annotation provided by compound vendors, bioactivities from public resources such as ChEMBL, and pathways in which these targets take part [45]. The resource supports many applications, such as the identification of compounds that are common between libraries or the identification of commercial compounds with desired properties. It can be searched both by target and by compound.

Probes & Drugs has assembled several valuable datasets of chemical probes, including one originally developed by Nature Chemical Biology from articles describing a new chemical probe and published in that journal and that is no longer available at the journal website. There are also two datasets manually curated from literature articles [46,47]. Moreover, in addition the resource integrates probes from chemical probe developers and includes the >2500 compounds assessed by Probe Miner that pass minimum quality criteria. In total, Probes & Drugs currently lists over 4000 actual or potential chemical probes. It provides a very useful integration of scattered chemical biology resources and facilitates the identification commercially available chemical probes among these resources. However, Probes & Drugs initially relied on the annotation of chemical probes provided by the individual resources that were integrated, and it was originally not straightforward to compare different probes for the same target between resources according to their fitness factors. Moreover, Probes & Drugs lists as chemical probes >2,500 compounds that, while they pass minimum quality criteria using Probe Miner's assessment, are not scrutinized as to whether they are among the best compounds available for a given target. Very recently, Probes & Drugs has incorporated a potency-selectivity score, inspired by the Probe Miner scoring methodology, which facilitates the comparison between chemical probes. The potency-selectivity scoring awaits comprehensive validation or benchmarking. Probes & Drugs is a very useful resource in chemical biology. Note, however, that older probe datasets and targets curated from vendor catalogues should be used with caution.

The Small Molecule Suite

Very recently, the Small Molecule Suite (https://labsyspharm.shinyapps.io/smallmoleculesuite) has been developed to compare, score and create chemical libraries based on computational analysis of selectivity, target coverage and induced cellular phenotypes, among other parameters, from public resources such as ChEMBL [48]. The applications of this resource go beyond the selection of individual chemical probes for specific targets, with the set of algorithms and software tools allowing, for example, the assembly of compound sets for screening with the lowest possible off-target overlap. The researchers evolved the tool score developed by Novartis scientists [42] and established other scoring metrics to compile a chemical library of 2026 compounds that covers 1852 targets [48]. Moreover, they have developed the SelectivitySelectR web-based application to display the affinity and selectivity of compounds in the HMS-LINCS collection (https://lincs.hms.harvard.edu) for a gene of interest, which can be used to identify and explore potentially useful chemical probes.

Taken together, chemical biologists, pharmacologists and drug discovery scientists have never before had access to such a diversity of empowering chemical biology resources. Integrative resources such as canSAR [44] or Open Targets (www.opentargets.org) [49] already provide links to several of the above-mentioned chemical probe resources. However, how best to use these resources depends on the experience and needs of the particular researcher – and in particular whether the user is an expert in chemical biology and the selection and application of chemical probes, or alternatively if they are a biologist or biomedical researcher with no or limited probe experience and have a one-off or occasional need to find a chemical probe for a particular line of experimentation in their model system. Experts can use the various resources to explore existing public data and keep track of advances in the field, but they ultimately rely on their accumulated expertise for chemical probe assessment and selection, gained with experience. Researchers with less experience can use the resources as a guide to select chemical probes but may need advice on navigating the various resources. In the next section we exemplify how nonexperts could use the available resources for chemical probe assessment and selection.

Complementary use of expert-curated & computational resources for optimal chemical probe assessment

How should a researcher use public chemical probe resources to make better decisions when selecting chemical probes? Figure 2 provides a flow diagram which suggests how this might be approached. It is worth stressing again that the ideal scenario would always be to have at least two chemical probes, from different chemotypes, that are both highly selective for the target of interest (preferably with distinct off-target profiles) together with an inactive analog and a target engagement biomarker. However, when this ideal scenario does not exist, probes may still be useful, but experiments should be considered to mitigate the risk. For example, if even the best available probe is not as selective as one would like, it would be essential to use a secondary chemotype with different off-targets and to demonstrate biomarker modulation to reduce the risk of off-targets effects confounding the results in the system under study. Similarly, when no secondary chemical probe with a different chemotype is available, the researchers should consider the use of biological tools to functionally inactivate the target (Figure 2).
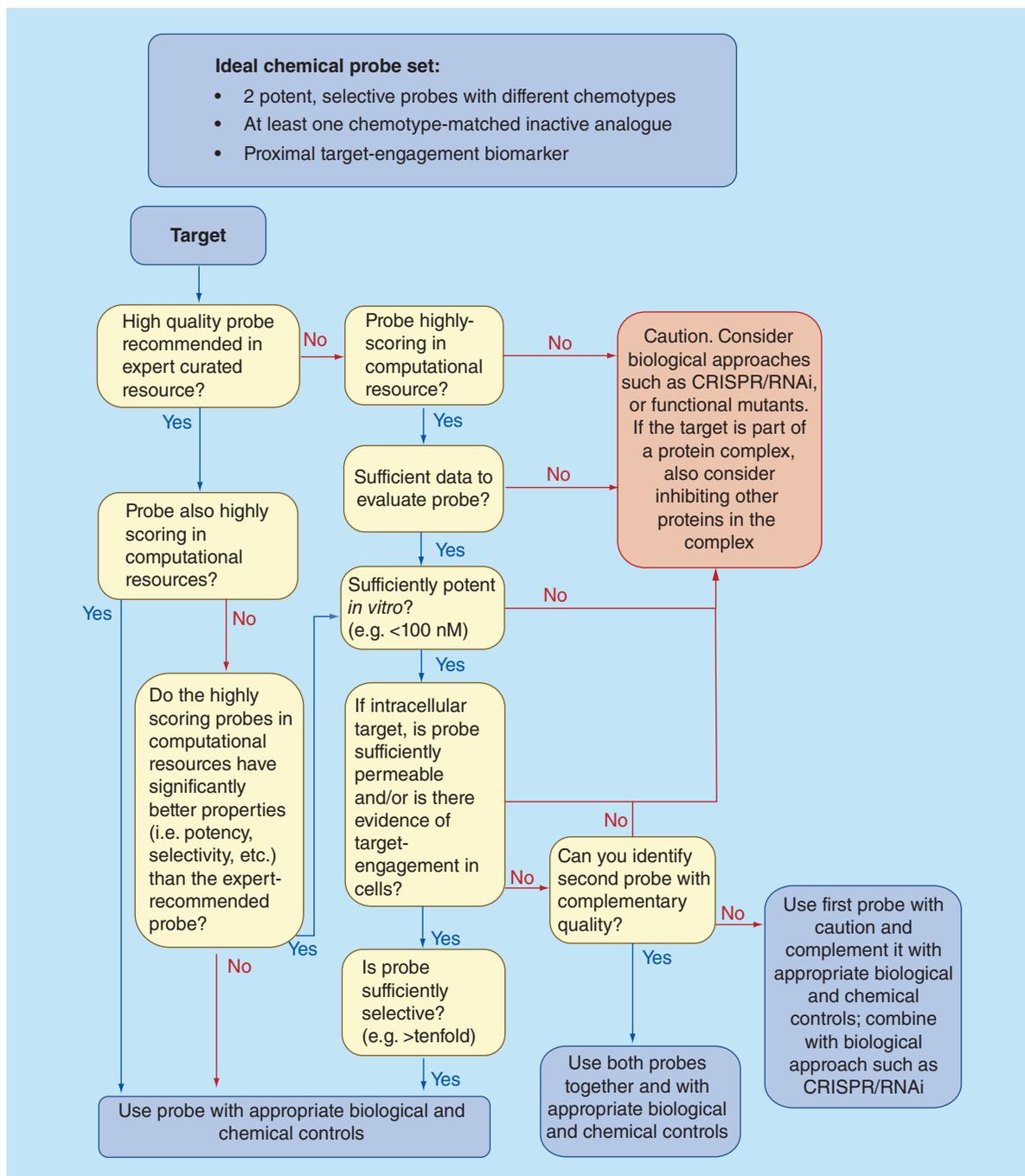
Figure 2. Schematic illustrating how researchers can use public chemical probe resources in a coordinated manner to assess and select chemical probes for their research. In the top box we summarize the ideal chemical probe set. The flowchart below illustrates how for a selected target of interest, a researcher can proceed to consult public chemical probe resources to identify the best chemical probes for that target. We recommend beginning with expert-reviewed resources, particularly the Chemical Probes Portal, which contains assessments of chemical probes carried out by chemical biology experts, together with useful information about their strengths, limitations and usage, plus a list of 'historical compounds' that should no longer be used. In addition, or whenever a target is not yet covered by an expert-review resource, we recommend using a computational resource, such as Probe Miner. Bearing in mind that computational resources are not expert-curated, we suggest extra steps should be taken to ensure that the chemical probes selected through Probe Miner are of sufficient quality. In the flow diagram we suggest several questions that the researcher can ask to make sure that there is enough information available to make an assessment that the chemical probe is sufficiently potent, selective and permeable, and that there is evidence of target engagement and/or modulation. Whenever minimal criteria are not met, we propose caution and/or additional controls that can be implemented to reduce the risk. Overall, the coordinated use of expert-curated and computational chemical probe resources can be strongly enabling to facilitate chemical probe assessment and selection. In addition to the above, we always recommend checking the recent literature for modulators of the target, and consulting with a chemical biology colleague if possible. For more details and examples please see main text.

In the flow diagram in Figure 2, we describe key questions that could serve as a guide for nonexperts on how to use available public chemical probe resources. We recommend initially to search expert-curated resources such as the Chemical Probes Portal for chemical probes suggested for the target of interest. If chemical probes are recommended in the Portal, we would advise users also to compare them with respect to the information available in computational resources like Probe Miner or Probes & Drugs. Despite their limitations, computational resources are regularly and more easily updated than expert-reviewed resources and often incorporate more recent and better-characterized probes that have not yet been reviewed publicly by experts. Accordingly, it would be important to check if the probes recommended in the Portal are highly ranked in computational resources – and if they are not, to see if other compounds that are highly ranked in Probe Miner or Probes & Drugs have improved properties (e.g., selectivity) that would justify their prioritization over the probes recommended by experts.

Unless there is a clear improvement in fitness factors, we advise following the recommendation of expert-curated resources. If there is no recommended probe in expert-curated resources, computational resources should be used instead, but we advise additional checks to ensure the quality of the potential chemical probes as the output of computational resources will not have been reviewed publicly by experts. For example, it is important to verify that a potential probe identified in public computational resources such as Probe Miner or Probes & Drugs has been tested sufficiently for selectivity against other targets – and to be sure that there are sufficient data available to evaluate this fitness factor. Moreover, we also recommend at least minimum thresholds of biochemical potency (<100 nM) and selectivity (>tenfold) and if these are not met, we suggest caution and/or the use of additional controls to mitigate risk – including avoidance of excessively high concentrations that would increase the likelihood of nonspecific effects. It is important to stress that the tenfold selectivity threshold is a minimum requirement and might not be sufficient in some cases, as was recently illustrated for small-molecules having off-target affinity for BET bromodomains – where even 30-fold selectivity was not sufficient to prevent the strong BET-dependent response [50]. It is always advisable to validate rigorously the connection between the target of interest and the phenotype using additional chemical or biological controls whenever possible. Below, we discuss three examples that represent real-world scenarios in which the flow diagram (Figure 2) is used as a guide to the selection of chemical probes using public resources.
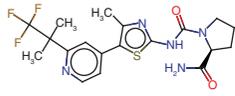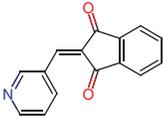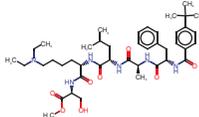
PIK3CA

At the time of submission of this article, a search for the PI3 kinase 'PIK3CA' in the Chemical Probes Portal highlighted two chemical probes: the recently approved PIK3CA-selective drug alpelisib (BYL719) and the pan-Class I inhibitory drug pictilisib (GDC-0941; Figure 3A). Both probes are potent and active in cells but alpelisib is more selective for PIK3CA whilst pictilisib also inhibits other Class I PI3K isoforms, which may be desirable, undesirable or neutral in the particular biological context of interest to the probe user. Probe Miner and Probes & Drugs also give high scores for these two chemical probes, alongside other pan-Class I PI3K and dual PI3K/mTOR inhibitors, illustrating that they are likely the best available probes as there do not seem to be newer and more selective chemical probes available. Accordingly, we recommend the use of both alpelisib and pictilisib as chemical probes, bearing in mind that pictilisib will inhibit other PI3K isoforms.

It is worth repeating, as referred to in the introductory section, that the use of LY294002 is not recommended. This is in spite of it continuing to maintain a high citation rate and being sold by vendors as a PI3K inhibitor. It is a flawed, unselective, historical compound. It is in fact flagged as an historical probe and not recommended in the Chemical Probes Portal and has a low ranking in Probe Miner and Probes & Drugs due to its weak potency and its off-target effects.

PRC1

If a researcher is interested in studying PRC1 and searches the existing literature, online search engines or vendor catalogs, they could come across PRT4165 that is sold as a Bmi1/Ring1A inhibitor. However, a search for 'BMI1' or 'RING1A' in the Chemical Probes Portal will not identify any compound recommended as a chemical probe for these targets. Moreover, Probe Miner returns no published compounds inhibiting them. This should raise an alarm and careful consideration is required (Figure 2). PRT4165 is listed as a BMI1/RING1A inhibitor in Probes & Drugs because it is displayed as such by a compound vendor. As discussed earlier, vendor catalogues should not be used to select chemical probes because they do not undergo any curation or scrutiny of the original information source. In the case of PRT4165, this compound was published as an unoptimized hit from a high-throughput screen measuring a Bmi1/Ring1A ubiquitination. The reported potency is poor (IC$_{50}$ = 3.9 μm) and no information is

Figure 3. Three examples illustrating the strengths and limitations of two actual or potential chemical probes for each of three specific targets. (A) PIK3CA. **(B)** PRC1. **(C)** PARP15. At the top of each panel, the intended target is displayed. Below it, the chemical probe name and chemical structures are represented followed by the icons of the resources that recommend this chemical probe for the target of interest. Under the heading Fitness Factors, we summarize and color code the most relevant fitness factors for each of the compounds shown (blue: high-quality; grey: medium-quality; red: low quality). Finally, we provide some recommendations for the use (or not) of each compound as a chemical probe for the target of interest.

publicly available concerning its selectivity [51]. Thus, given the lack of useful assessment or evaluation metrics from Probe Miner and the Chemical Probes Portal and its currently known properties in the public domain, we strongly caution against the use of this compound as a chemical probe for PRC1.

To study the function of the PRC1 complex one could instead consider chemical probes for other protein components such as CBX7. The CBX7 protein has the chemical tool UNC3866 recommended by experts in the

Table 1. A selection of available public resources for chemical probes.		
Resource Type	URL	Coverage
Expert-curated		
Chemical Probes Portal	www.chemicalprobes.org/	>190 chemical probes against >180 biological targets and >200 'historical' compounds
NIH Molecular Libraries Program	www.ncbi.nlm.nih.gov/books/NBK47352	375 chemical probes
Structural Genomics Consortium	www.thesgc.org/chemical-probes	65 probes against 65 biological targets
Nathanael Gray laboratory	https: //graylab.dana-farber.org/probes.html	38 chemical probes against 38 targets; 96 best-in-class chemical probes for 90 kinases
opnMe	https://opnme.com	>30 chemical probes against >30 targets
Donated Chemical Probes	https://openscienceprobes.sgc-frankfurt.de	70 chemical probes provided by industry
Quantitative, objective, data-driven, computational		
Probe Miner	https://probeminer.icr.ac.uk	Scores >300,000 compounds as potential chemical probes against >2600 human targets
Probes & Drugs Portal	www.probes-drugs.org	> 4000 potential/actual chemical probes
Small Molecule Suite	https://labsyspharm.shinyapps.io/smallmoleculesuite/	2026 compounds against 1852 targets

Chemical Probes Portal (Figure 3B). Moreover, this chemical probe also ranks top against CBX7 in Probe Miner. Its selectivity has been characterized, it has demonstrated cellular activity, and its *in vitro* affinity is in the low nanomolar range (Kd = 97 nM). The Chemical Probes Portal and Probe Miner clearly highlight the strengths and limitation of UNC3866, such as its limited selectivity over CBX4. Accordingly, our recommendation would be to use UNC3866, checking for target modulation and carefully controlling for whether CBX4 could be important in the system under study, including also using complementary biological tools [52].

PARP15

There will be instances where a chemical probe approach is not a good option for a particular target. This applies in the case of PARP15. The Chemical Probes Portal does not recommend any chemical probe for this target (Figure 2). Probes & Drugs displays olaparib while Probe Miner ranks canSAR2285520 as the top small-molecule inhibitor for PARP15 (Figure 3C). However, data in both resources clearly indicate that these chemical inhibitors are not good candidates. Olaparib is far more potent against other PARP family members compared to PARP15 while canSAR2285520 is only weakly potent and poorly characterized. Our recommendation would be not to use chemical inhibitors to study this target and to rely on biological tools until better chemical probes are developed.

These three examples above illustrate how the use of expert-curated and computational chemical probe resources is strongly enabling and facilitates the assessment and selection of chemical probes, particularly for researchers who are not experienced in chemical biology or drug discovery. We counsel that researchers should be especially cautious about relying on data obtained exclusively from vendors and recommend the combined use of computational and expert-curated resources like the Chemical Probes Portal to help choose the best available probes. These resources can also be valuable to make researchers aware of the limitations that should be considered for a given chemical probe, for example the degree and extent of selectivity, and inform on the appropriate controls to mitigate risks such as off-target effects. Of course, consulting the primary literature to further evaluate the fitness properties of a given compound is always important before selecting a chemical probe for a given experimental use. However, care should be taken to avoid historical compounds that have impressive citation rates but which have been superseded by more suitable modern tools. Obtaining advice from expert chemical biologists can be very helpful, particularly if the non-expert has doubts about which chemical probe to use after consulting available resources (Table 1).

Future perspective

The last 5-year period has seen the development of several public chemical probe resources. While each resource provides unique functionalities, they all empower researchers in the assessment and selection of chemical probes. Despite their respective limitations, they facilitate the assessment of fitness factors for potential chemical probes, point to the best available chemical tools, enable the identification of drawbacks on selected probes that should be carefully controlled for, and discourage the choice of low-quality or flawed compounds that continue to be sold

by chemical vendors without alerting on their problems. Accordingly, these resources contribute to addressing the frequent and ongoing mistakes in chemical probe selection and use and to improving best practice in the wider biomedical research sphere.

We envisage that a major challenge will be the long-term maintenance and sustainability of these resources. Appropriate sources of funding must be identified to ensure that public access is secured. Moreover, it will be essential that experienced chemical biology experts continue to support expert-curated resources by providing key information on the suitability of probes, contributing to reviews and sharing their experience to advise the broader research communities on the importance of using the right chemical probes.

Though a laudable goal, and very good progress has been made, it will require major efforts from the chemical biology and drug discovery community to ensure substantially increased proteome coverage so that every liganded target is supported by expert advice on the best available chemical probes in expert-curated resources such as the Chemical Probes Portal. It will be challenging to keep peer-reviewed resources updated as new and better chemical probes become available for a given target. By contrast, coverage and timeliness are the strengths of computational resources. Accordingly, experts could use computational resources to help them increase the target coverage and track the release of better new chemical probes for subsequent review in the Chemical Probes Portal.

Since all resources have their limitations, combined use is recommended. Moreover, we believe that an enhanced level of integration between the chemical probe resources would be hugely beneficial to the bioscience community. The Chemical Probes Portal would be the ideal venue at which to provide centralized integration of expert-curated input and information from chemical probe developers – linking to the quantitative objective assessment resources such as Probe Miner which can help both probe reviewers and users – and hence continue to evolve as a user-friendly one-stop-shop for scientists wishing to identify the best chemical probes for their target and research question of interest. Activity is underway to enhance the overall capability and the data and resource integration, as well as target/probe coverage of the Chemical Probes Portal [53]. Meanwhile, we recommend the complementary use of the different resources for chemical probe assessment and selection, as summarized in Figure 2. In addition to the initial use of recommended public web resources, we advise that researchers should always check the latest scientific literature and if in doubt consult an expert chemical biology colleague.

The biggest current limitation of computational approaches lies in data quality and availability. To make the most out of computational resources, it will be important to better capture chemical biology data that are currently missing from resources that are built on medicinal chemistry databases, and that are scattered across many locations, including probe developer resources and journals outside the medicinal chemistry literature. Improving data quality, curation and standardization will also be important for the benefit of current and future computational analysis methods and resources. We foresee continued development of computational approaches and scoring algorithms for chemical probes. In particular, the use of machine learning and AI approaches will likely be particularly enabling to enhance current computational assessments [54].

Initiatives such as Donated Chemical Probes that aim to release chemical probes that are otherwise hidden in big pharma vaults and to characterize them using state-of-the-art standards will be continue to be important, despite the current modest scale. Debate is likely to continue regarding the need for publication of recommended probes in the scientific literature, but meanwhile the expert review, broad off-target profiling and release of inactive controls are highly valuable features. The emergence of new chemical probe types such as activity-based probes [55] or targeted protein degradation probes [56] should be also monitored and included where appropriate. Moreover, the recent publication of the Target 2035 proposals illustrates the increase in chemical probe development initiatives and raises the hope of achieving the ambitious goal of developing a chemical probe for every human protein [57]. However, it will also be important to continue to develop better probes for targets that currently have only limited- or low-quality chemical reagents available. In addition, it is important to generate the key data demonstrating that some claimed chemical probes are not valid, such as was exemplified very well recently for inhibitors of human NMT2 and Keap1−Nrf2 protein–protein interaction inhibitors [58,59]. An important recent publication used CRISPR technology to show that for several drugs undergoing clinical development the claimed principal target is not in fact required by cancer cells and that the drugs in question do not work by acting on the ostensible target, but rather through other, off-target effects [11]. Public chemical probe resources will be key to prevent the future use of low-quality or invalidated chemical reagents but how to facilitate capturing this information in a timely manner and to discourage uncritical reading of the old literature that was generated using these flawed small-molecule inhibitors will remain challenging.

Overall, chemical probe resources are on the rise, but perhaps the biggest challenge is how to achieve greater uptake of these by the wider scientific community. Ensuring that the resources are user-friendly to non-specialists while still valuable to experts is essential. It is a pressing need that bioscience researchers avoid the use of standard search engines and vendor catalogs and instead refer to specialized public resources described here for chemical probe selection in order to avoid the pitfalls of using – and perpetuating the use of – flawed probe reagents. Various ways of promulgating the message include articles in the more mainstream biological and general science journals [3,60], sessions at biological and biomedical conferences, blogs and social media [61–63].

It is important to emphasize that public chemical probe resources are a community effort provided for the benefit of the community. Accordingly, the community should not shy away from providing critical input, information and experience. Alongside expert researchers, stakeholders such as funding bodies and journals also have a big responsibility. Many reputable journals now demand minimum-quality checks and short tandem repeat DNA profiling to avoid past pitfalls with misidentification, cross-contamination and misuse of mammalian cell lines (www.ncbi.nlm.nih.gov/books/NBK144066/). Similarly, journals should more consistently require that any research publication using chemical probes includes justification for how the selected probes were chosen in the context of other available probes for the same target, together with references or any new key data – at least on potency, selectivity and permeability – to support their choice. The exact chemical structure should also be included or referenced. Although there is still more to do, biological researchers have become more aware, and journal reviewers and editors more demanding, with respect to the use of biological/genetic reagents and the need for essential controls to mitigate off-target activity [3,64], including rescue experiments [65] – yet this does not generally apply to the use of chemical probes, especially in many biology journals.

In conclusion, the chemical probe resources discussed here are highly valuable – yet they can be further improved and their use needs to be promoted much more widely and effectively. Good practice is facilitated by cross-disciplinary collaborations and conversations between biologists and chemists but we need to reach out more effectively to biomedical researchers who do not currently work with medicinal chemists and chemical biologists. We need to find ways to encourage biologists to be at least equally (if not more) circumspect with chemical probes as they are with biological reagents, including active outreach beyond the specialist journals and conferences where chemical biologists and probe experts converse with each other. The coordinated effort of all the chemical probe stakeholders – from probe developers and users to journal editors, reviewers, chemical vendors and funding agencies – will be necessary to truly eliminate the misuse of flawed chemical probes that currently contaminate the scientific literature and reduce the robustness and reproducibility of biological and biomedical research.

Executive summary

- Chemical probes are commonly not being selected using objective, data-driven assessment of fitness factors, which results in the use of low-quality or flawed small-molecule inhibitors that generate misleading and invalid results in the scientific literature.
- Several public chemical probe resources have been developed recently that greatly empower researchers in chemical probe assessment and selection and that support best practice guidelines.
- The Chemical Probes Portal is the most comprehensive expert-curated resource but would benefit from increased target and probe coverage and the facility for bulk data downloads and programmatic access to the Portal's data.
- Computational resources such as Probe Miner or Probes & Drugs offer objective, computational analysis, larger coverage and timely updates but suffer from limitations in publicly available data; improved capture and curation of chemical biology data are needed to improve computational resources.
- We propose the complementary use of both expert-curated and objective, computational resources and provide advice and a useful flowchart to facilitate the assessment and selection of chemical probes, particularly for researchers who are not chemical biology experts.
- The chemical biology community should promulgate the use of these resources if we are to address current challenges in chemical probe selection that affect the robustness and reproducibility of bioscience research.
- The coordinated effort of all the chemical probe stakeholders – from developers and users to journal editors, reviewers, chemical vendors and funding agencies – is required to improve the selection and use of chemical probes, and greater outreach to the biomedical research community will be extremely important.

Open access

References

Papers of special note have been highlighted as: ● of interest

1. Workman P, Al-Lazikani B. Drugging cancer genomes. *Nat. Rev. Drug Discov.* 12(12), 889–890 (2013).

2. Behan FM, Iorio F, Picco G *et al.* Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature* 568(7753), 511–516 (2019).

3. Blagg J, Workman P. Choose and use your chemical probe wisely to explore cancer biology. *Cancer Cell* 32(1), 9–25 (2017).
● **Comprehensive review and perspective article on choosing and using chemical probes – notably aimed at biological scientists and containing background, examples of good and bad usage in oncology, advice on best practice, and useful 'dos and don'ts'. Focuses on the cancer field, but generally applicable.**

4. Workman P, Collins I. Probing the probes: fitness factors for small molecule tools. *Chem. Biol.* 17(6), 561–577 (2010).
● **Original detailed review and perspective article, setting out and discussing the 'fitness factors' that represent essential criteria for chemical probes. Provides important background and examples for developers and users of chemical probes.**

5. Frye SV. The art of the chemical probe. *Nat. Chem. Biol.* 6(3), 159–161 (2010).
● **Excellent short commentary where a set of five principles to guide chemical probe qualification are proposed and discussed.**

6. Bunnage ME, Chekler ELP, Jones LH. Target validation using chemical probes. *Nat. Chem. Biol.* 9(4), 195–199 (2013).
● **This valuable concise commentary from pharma industry authors discusses the use of 'fully profiled' chemical probes to help identify or validate new molecular targets that drive phenotypic responses, emphasizing the helpful 'four pillars' concept.**

7. Brown PJ, Müller S. Open access chemical probes for epigenetic targets. *Future Med. Chem.* 7(14), 1901–1917 (2015).

8. Arrowsmith CH, Audia JE, Austin C *et al.* The promise and peril of chemical probes. *Nat. Chem. Biol.* 11(8), 536–541 (2015).
● **Influential community commentary by academic and industry experts, describing the opportunities and dangers of chemical probes and outlining the launch of the Chemical Probes Portal.**

9. Mateo J, Ong M, Tan DSP, Gonzalez MA, De Bono JS. Appraising iniparib, the PARP inhibitor that never was – what must we learn? *Nat. Rev. Clin. Oncol.* 10(12), 688–696 (2013).

10. Kocak DD, Josephs EA, Bhandarkar V, Adkar SS, Kwon JB, Gersbach CA. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37(6), 657–666 (2019).

11. Lin A, Giuliano CJ, Palladino A *et al.* Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* 11(509), eaaw8412 (2019).
● **An important paper using CRISPR technology to show that for several drugs undergoing clinical development the claimed principal target is not in fact required by cancer cells and that the drugs in question do not work by acting on that target, but rather through other off-target effects.**

12. Antolin AA, Mestres J. Distant polypharmacology among MLP chemical probes. *ACS Chem. Biol.* 10(2), 395–400 (2015).

13. Davies SP, Reddy H, Caivano M, Cohen P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* 351(1), 95–105 (2000).

14.   Structural Genomics Consortium. Chemical Probe Criteria (2019).www.thesgc.org/chemical-probes.

15.   Rossanese O, Eccles S, Springer C *et al.* The pharmacological audit trail (PhAT): use of tumor models to address critical issues in the preclinical development of targeted anticancer drugs. *Drug Discov. Today Dis. Model.* 21, 23–32 (2016).

16.   Simon GM, Niphakis MJ, Cravatt BF. Determining target engagement in living systems. *Nat. Chem. Biol.* 9(4), 200–205 (2013).

17.   Editorial. Probing questions. *Nat. Chem. Biol.* 11, 533 (2015).

18.   Austin CP, Brady LS, Insel TR, Collins FS. NIH molecular libraries initiative. *Science (80-.).* 306, 1138–9 (2004).

19.   Wang Y, Bryant SH, Cheng T *et al.* PubChem BioAssay: 2017 update. *Nucleic Acids Res.* 45(D1), D955–D963 (2017).

20.   Sayers EW, Barrett T, Benson DA *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 40(D1), 5–16 (2012).

21.   Oprea TI, Bologa CG, Boyer S *et al.* A crowdsourcing evaluation of the NIH chemical probes. *Nat. Chem. Biol.* 5(7), 441–447 (2009).

22.   Schreiber SL, Kotz JD, Li M *et al.* Advancing biological understanding and therapeutics discovery with small-molecule probes. *Cell.* 161(6), 1252–1265 (2015).

●   **Detailed account of the progress made in the NIH Molecular Libraries Program, exemplifying the novel small-molecule probes identified and their use in enabling the exploration of biological pathways and therapeutic hypotheses.**

23.   Williamson AR. Creating a structural genomics consortium. *Nat Struct Biol.* 7(11s), 953 (2000).

24.   Filippakopoulos P, Qi J, Picaud S *et al.* Selective inhibition of BET bromodomains. *Nature.* 468(7327), 1067–1073 (2010).

25.   Scott AR. Chemical probes: a shared toolbox. *Nature.* 533(7602), S60–S61 (2016).

26.   Nilsson N, Felding J. Open innovation platforms to boost pharmaceutical collaborations: evaluating external compounds for desired biological activity. *Future Med. Chem.* 7(14), 1853–1859 (2015).

27.   Bender E. Crowdsourced solutions. *Nature.* 533(3), S62–S64 (2016).

28.   Gamo FJ, Sanz LM, Vidal J *et al.* Thousands of chemical starting points for antimalarial lead identification. *Nature* 465(7296), 305–310 (2010).

29.   Mullard A. Boehringer Ingelheim experiments with open-access chemical probes. *Nat. Rev. Drug Discov.* 17(1), 7 (2018).

30.   Muller S, Ackloo S, Arrowsmith CH *et al.* Donated chemical probes for open science. *Elife.* 7, e34311 (2018).

31.   Lee WH. Open access target validation is a more efficient way to accelerate drug discovery. *PLoS Biol.* 13(6), e1002164 (2015).

32.   Clarke PA, Roe T, Swabey K *et al.* Dissecting mechanisms of resistance to targeted drug combination therapy in human colorectal cancer. *Oncogene.* 38, 5076–5090 (2019).

33.   Whittaker SR, Barlow C, Martin MP *et al.* Molecular profiling and combinatorial activity of CCT068127: a potent CDK2 and CDK9 inhibitor. *Mol. Oncol.* 12(3), 287–304 (2018).

34.   Vogt M, Jasial S, Bajorath J. Extracting compound profiling matrices from screening data. *ACS Omega.* 3(4), 4706–4712 (2018).

35.   Chen WL. Chemoinformatics: past, present, and future. *J. Chem. Inf. Model.* 46(6), 2230–2255 (2006).

36.   Chen X, Liu M, Gilson M. BindingDB: a web-accessible molecular recognition database. *Comb. Chem. High Throughput Screen.* 4(8), 719–725 (2001).

37.   Harding SD, Sharman JL, Faccenda E *et al.* The IUPHAR/BPS guide to pharmacology in 2018: updates and expansion to encompass the new guide to immunopharmacology. *Nucleic Acids Res.* 46(D1), D1091–D1106 (2018).

38.   Gaulton A, Hersey A, Nowotka M *et al.* The ChEMBL database in 2017. *Nucleic Acids Res.* 45(D1), D945–D954 (2016).

●   **Introduces the highly influential ChEMBL database containing binding, functional and ADMET information, all manually abstracted from the medicinal chemistry literature. Used as a source of chemical and bioactivity data for other resources such as canSAR and Probe Miner.**

39.   Krallinger M, Rabal O, Lourenço A, Oyarzabal J, Valencia A. Information retrieval and text mining technologies for chemistry. *Chem. Rev.* 117(12), 7673–7761 (2017).

40.   Dimova D, Bajorath J. Computational chemical biology: identification of small molecular probes that discriminate between members of target protein families. *Chem. Biol. Drug Des.* 79(4), 369–375 (2012).

41.   Butler KV, MacDonald IA, Hathaway NA, Jin J. Report and application of a tool compound data set. *J. Chem. Inf. Model.* 57(11), 2699–26706 (2017).

42.   Wang Y, Cornett A, King FJ *et al.* Evidence-based and quantitative prioritization of tool compounds in phenotypic drug discovery. *Cell Chem. Biol.* 23(7), 862–874 (2016).

43.   Antolin AA, Tym JE, Komianou A, Collins I, Workman P, Al-lazikani B. Objective, quantitative, data-driven assessment of chemical probes. *Cell Chem. Biol.* 25(2), 194–205 (2018).

●   **Describes the public Probe Miner resource, which enables objective, quantitative, data-driven assessment of chemical probes based on regularly updated large-scale data – recommended for use in conjunction with the expert review-based Chemical Probes Portal.**

44. Coker EA, Mitsopoulos C, Tym JE *et al.* canSAR: update to the cancer translational research and drug discovery knowledgebase. *Nucleic Acids Res.* 47(D1), D917–D922 (2018).

45. Skuta C, Popr M, Muller T *et al.* Probes & Drugs portal: an interactive, open data resource for chemical biology. *Nat. Methods.* 14(8), 759–760 (2017).

•   **Introduces the Probes & Drugs resource which utilizes proprietary commercial as well as public data.**

46. Scheer S, Ackloo S, Medina TS *et al.* A chemical biology toolbox to study protein methyltransferases and epigenetic signaling. *Nat. Commun.* 10(1), 1–14 (2019).

47. Wu Q, Heidenreich D, Zhou S *et al.* A chemical toolbox for the study of bromodomains and epigenetic signaling. *Nat. Commun.* 10(1), 1915 (2019).

48. Moret N, Clark NA, Hafner M *et al.* Cheminformatics tools for analyzing and designing optimized small molecule collections and libraries. *Cell Chem. Biol.* 26(5), 765–777 (2019).

•   **Describes a data-driven approach to the design of compound libraries, linked to the public Small Molecule Suite online tool.**

49. Koscielny G, An P, Carvalho-Silva D *et al.* Open targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* 45(D1), D985–D994 (2017).

•   **Presents the public Open Targets resource as a platform to aid identification of therapeutic targets.**

50. Galdeano C, Ciulli A. Selectivity on-target of bromodomain chemical probes by structure-guided medicinal chemistry and chemical biology Future. *Future Med. Chem.* 8(13), 1655–1680 (2016).

51. Alchanati I, Teicher C, Cohen G *et al.* The E3 ubiquitin-ligase Bmi1/Ring1A controls the proteasomal degradation of Top2α cleavage complex – a potentially new drug target. *PLoS ONE.* 4(12), e8104 (2009).

52. Stuckey JI, Dickson BM, Cheng N *et al.* A cellular chemical probe targeting the chromodomains of polycomb repressive complex 1. *Nat. Chem. Biol.* 12(3), 180–187 (2016).

53. Mullard A. A probe for every protein. *Nat. Rev. Drug Discov.* 18(10), 733–736 (2019).

54. Workman P, Antolin AA, Al-Lazikani B. Transforming cancer drug discovery with Big Data and AI. *Expert Opin. Drug Discov.* 14(11), 1089–1095 (2019).

55. Jones LH. Cell permeable affinity- and activity-based probes. *Future Med. Chem.* 7(16), 2131–2141 (2015).

56. Paiva SL, Crews CM. Targeted protein degradation: elements of PROTAC design. *Curr. Opin. Chem. Biol.* 50, 111–119 (2019).

57. Carter AJ, Kraemer O, Zwick M, Mueller-Fahrnow A, Arrowsmith CH, Edwards AM. Target 2035: probing the human proteome. *Drug Discov. Today* (2019). https://doi.org/10.1016/j.drudis.2019.06.020. (Epub ahead of print).

•   **An aspirational commentary calling for the chemical biology community to develop a chemical probe or antibody for every protein in the human proteome by 2035.**

58. Kallemeijn WW, Lueg GA, Faronato M *et al.* Validation and invalidation of chemical probes for the human N-myristoyltransferases. *Cell Chem. Biol.* 26(6), 892–900.e4 (2019).

•   **An excellent recent example of identifying compounds that should or should not be used as chemical probes, in this case for human N-myristoyltransferases.**

59. Tran KT, Pallesen JS, Solbak SMØ *et al.* A comparative assessment study of known small-molecule Keap1–Nrf2 protein–protein interaction inhibitors: chemical synthesis, binding properties, and cellular activity. *J. Med. Chem.* 62(17), 8028–8052 (2019).

60. Baker M. Reproducibility: check your chemistry. *Nature.* 548(7668), 485–488 (2017).

61. Lowe D. Chemical probe compounds: time to get real (2015). https://blogs.sciencemag.org/pipeline/archives/2015/07/22/chemical_probe_compounds_time_to_get_real.

62. Kostic M. Chemical probes: magic bullets, or smoke and mirrors? (2017). http://crosstalk.cell.com/blog/chemical-probes-magic-bullets-or-smoke-and-mirrors.

63. Workman P. More on chemical probes – from Indiana Jones to open science (2018). www.icr.ac.uk/blogs/the-drug-discoverer/page-details/more-on-chemical-probes-from-indiana-jones-to-open-science.

64. Kaelin WG. Common pitfalls in preclinical cancer target validation. *Nat. Rev. Cancer* 17(7), 441–450 (2017).

•   **A thoughtful and important perspective, outlining the most common pitfalls in preclinical cancer target identification with both genetic and pharmacological reagents and some potential approaches to mitigate them. Covers discussion of reproducibility versus robustness; causation versus correlation; necessity versus sufficiency; 'up' versus 'down' assay readouts; on-target versus off-target effects; use of positive and negative controls; the importance of 'rescue' experiments; and the value of orthogonal corroboration.**

65. Zhang C, Stockwell SR, Elbanna M *et al.* Signalling involving MET and FAK supports cell division independent of the activity of the cell cycle-regulating CDK4/6 kinases. *Oncogene.* 38(30), 5905–5920 (2019).