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

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

BMCL Digest

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

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

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

ARTICLE INFO

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

ABSTRACT

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

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

-Niels Bohr

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

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

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

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

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

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

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

0960-894X
http://dx.doi.org/10.1016/j.bmcl.2013.03.028

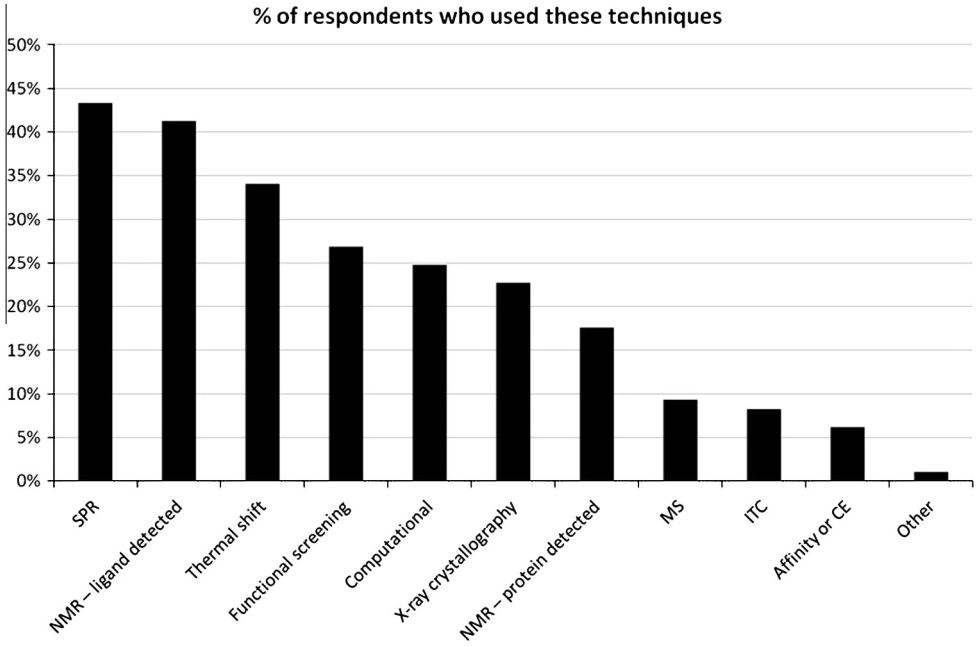


Figure 1. Methods to find fragments. These techniques were used to identify fragments, according to a poll on Practical Fragments in September 2011. There were 97 unique responses, and the average respondent used 2.4 different techniques.

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

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

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

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

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

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

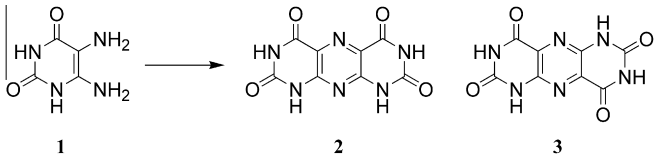
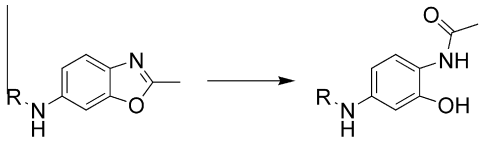


Figure 2. Unstable molecules. Compound **1** oxidizes in DMSO and dimerizes to form **2** and **3**. See text for details.

avoid this, regular analysis of the library stock solution by NMR or LCMS is good practice.[18]

In other cases, water is the problem—and because DMSO is hygroscopic, water is often (unintentionally) present in the compound stock solution. This can cause particular problems where DMSO stocks are subjected to freeze–thaw cycles.[19,20] Sandra Nelson and co-workers showed that substantially more compound degradation occurred in a test set of compounds which were subjected to freeze–thaw cycles than the same compounds stored at room temperature for the same period of time.[21] The amount of water taken up by DMSO stocks can vary dramatically depending on factors such as the position of a well in a plate, which is an additional complication in tracking compound stability. As an example, 2-methyl-benzoxazole derivatives such as compound **4** can hydrolytically open to produce compound **5** (Fig. 3). The reaction appears to be acid-catalyzed, and in fact the hydrochloride salt can decompose even when stored as a solid.[22]

Compound instability in aqueous solution can also give rise to erroneous assay data, particularly where the assay has a pre-incubation period (as is often the case). In one case, a bicyclic compound (species A) formed a tricyclic compound (species B) when stored in DMSO. Species B readily reacted with water, and formed a third compound (species C) over the course of hours in aqueous solution (Fig. 4). For the target of interest, species B was the active species, but the concentration of species B varied dramatically



4 5

Figure 3. Unstable molecules. Compound **4** is hydrolytically unstable. See text for details.

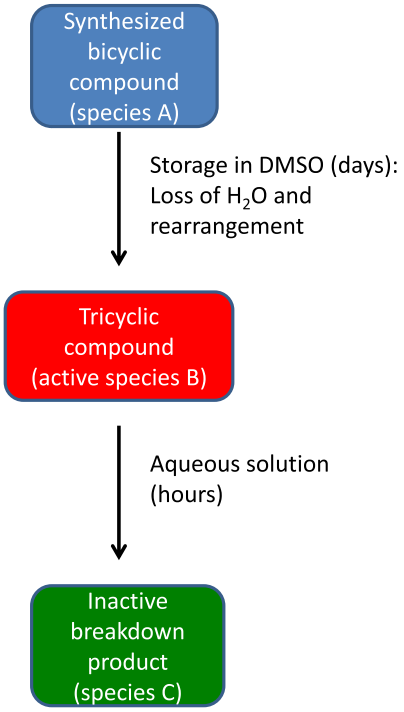


Figure 4. Schematic of example of compound stability causing problems.

depending on the details of the sample handling prior to the assay (personal communication).

Even when the compound is stable in aqueous solution, phenomena such as tautomerization should be considered. This is particularly the case for fragments, where substructures of the hits are often used as pharmacophores for subsequent evolution. An example of this is a ketone-containing compound which, in the ligand observed NMR spectra, clearly bound only in the enol form (personal communication). Moreover, compound solubility is often an issue at the high concentrations used in fragment screening. A recent analysis of fragments acquired from multiple commercial vendors found that 16% did not pass quality control standards set by Emerald Biostructures. About half of these failed due to solubility, with the remainder showing either degradation or impurities and a few with ambiguous or incorrect structures.[23] Of course, solubility is critically dependent on experimental conditions; an acidic fragment may be highly soluble at neutral pH but much less so at the low pH required for assaying, say, a lysosomal protein. This can also be an issue in crystallography, where the ionic strength or pH required for crystallization can be substantially different to that used in assays.

Reactive molecules. Even if a molecule is stable, it may react covalently with biological targets. Of course, lots of drugs work in this manner, and there is a growing trend of designing such reactive molecules, but you generally want to know this ahead of time.[24,25] In a review by Gilbert Rishton, the functionalities in Figure 5 were reported to be potentially reactive; molecules containing most of these are intentionally excluded from screening libraries for this reason.[26] However, this list is not comprehensive, and as the next sections show, molecules can pass cursory inspection but still cause problems.

PAINS. Most experienced medicinal chemists will look askance at the compounds in Figure 5, but not all problematic compounds are obvious troublemakers. After performing a number of high-throughput screens and finding that the same compounds hit many disparate targets, Jonathan Baell and Georgina Holloway christened such molecules pan-assay interference compounds, or PAINS; some of these substructures are shown in Figure 6.[14,27,28]

A common feature of many PAINS is that they contain a Michael acceptor; rhodanines (Fig. 6 upper left) are a case in point. Such molecules are soft electrophiles that can react with nucleophilic residues in proteins, either reversibly or irreversibly.[29] Although they can also bind non-covalently,[30] their propensity to form covalent bonds makes them unsuitable for drugs or even probe molecules, as selectivity will likely confound interpretation of the biology.

Some of these molecules are also photochemically reactive. For example Percy Carter and co-workers found that certain rhodanines and related molecules could form covalent adducts with a protein under ambient illumination. Although some molecules exhibited weak reversible binding in the dark, others appeared to be completely dependent on light for their activity.[31,32]

Unfortunately, molecules containing the moieties shown in Figure 6 are widely available commercially and thus found in numerous screening libraries; they often show up as hits, and too many researchers are unaware of their promiscuous nature. Indeed, Baell has aptly described such molecules as ‘polluting the literature’ with peer-reviewed false positives.

An example of how such molecules can be misleading has recently been described by R. Kiplin Guy and colleagues. SJ-172550 (compound **6** in Fig. 6) had initially been reported as an inhibitor of the p53-HDMX protein–protein interaction.[33] However, subsequent detailed mechanistic work demonstrated that not only was it a covalent inhibitor, it was also sensitive to reducing agents and possibly exhibited some of its activity by aggregation (see sec-

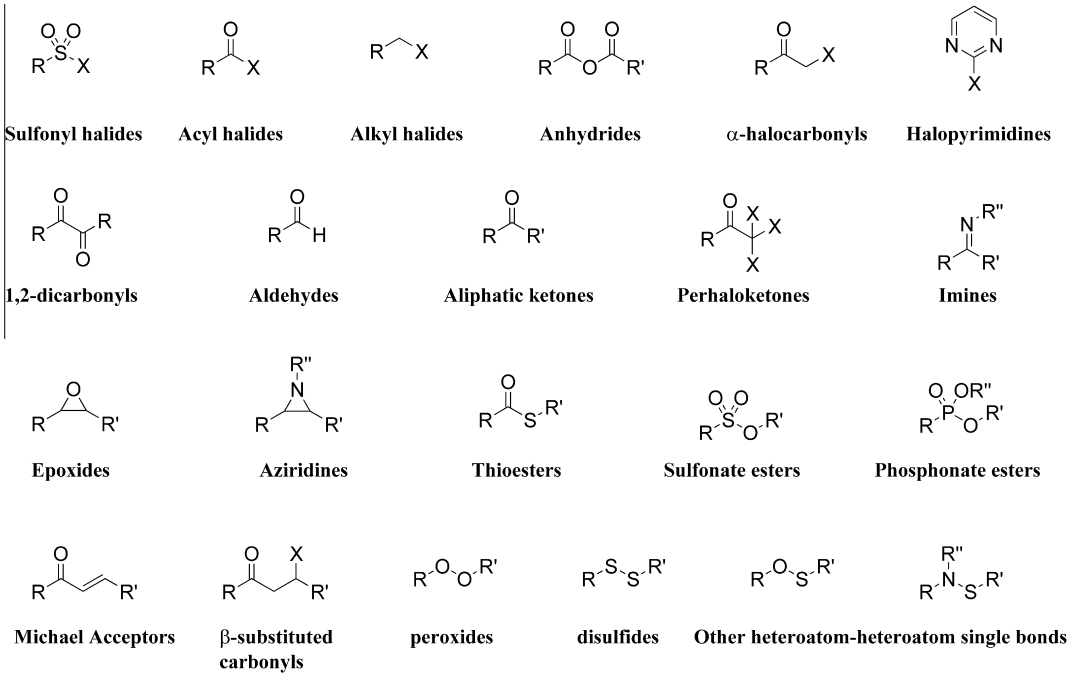


Figure 5. Reactive moieties. These functionalities are potentially reactive and should generally be avoided.

tion below). As the authors conclude, 'this complex, multimode mechanism greatly complicates the interpretation of experiments using [SJ-172550] and limits its value as a lead compound for further development as a selective MDMX inhibitor.'[34]

Redox-active molecules. A particularly nasty subset of PAINS are so-called redox cycling compounds. These molecules, many of which seem to be nitrogen-rich heterocycles, can be reduced by common buffer components such as DTT or TCEP to form reactive species that spontaneously oxidize in air, producing hydrogen peroxide in the process (see conversion of compounds **9**–**10** in Fig. 7). This in turn can readily oxidize proteins, particularly cysteine residues.[35,36] Some of these molecules are fragment-sized. They have shown up in a number of screens, and unfortunately their mechanism has not always been recognized. For example,

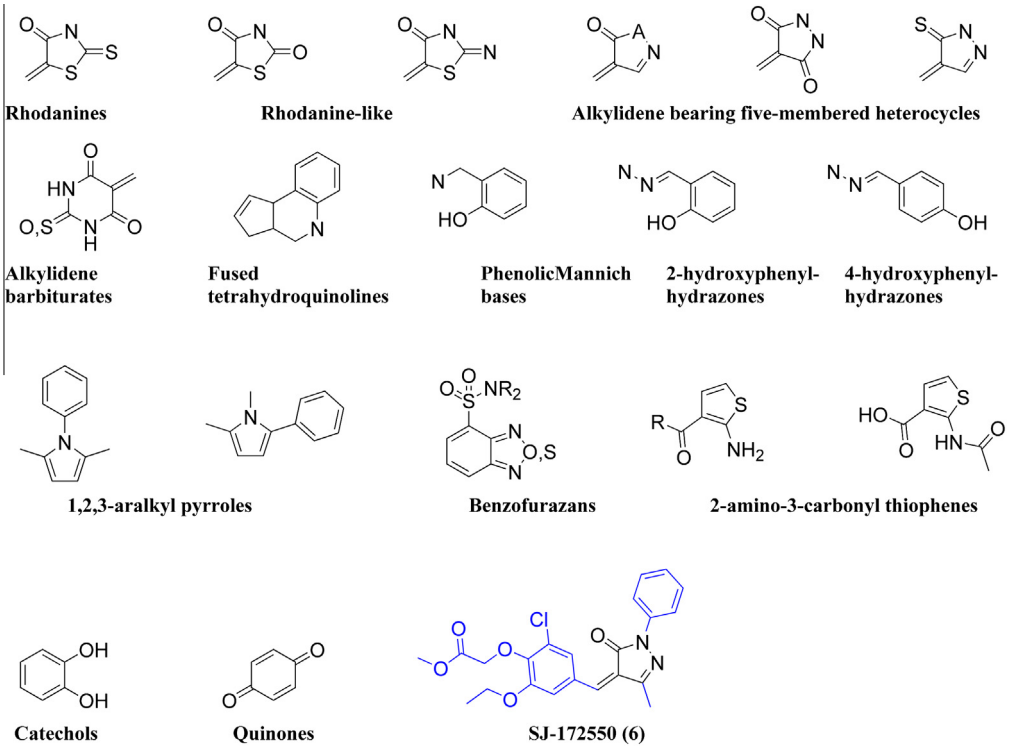


Figure 6. Pan-assay interference compounds (PAINS). These moieties have been identified as PAINS and should generally be avoided. Note that SJ-172550 (compound **6**, lower right) contains an alkylidene bearing five-membered heterocycle (top row, second from left).

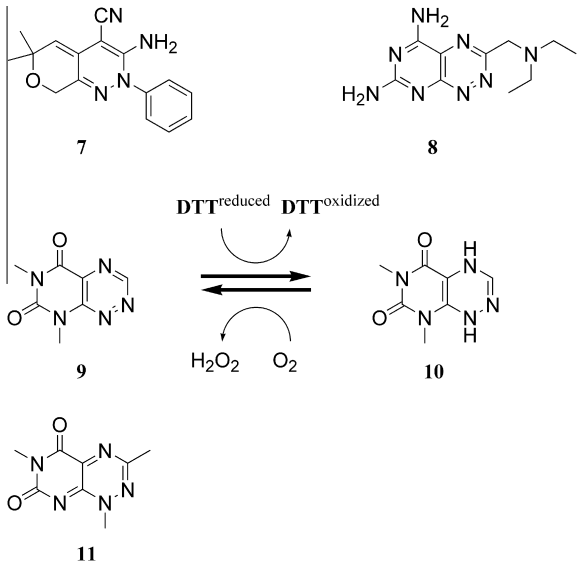


Figure 7. Redox cyclers. Compounds **7**–**9** and **11** have been shown to generate hydrogen peroxide when exposed to reducing agents such as DTT or TCEP in the presence of air (middle).

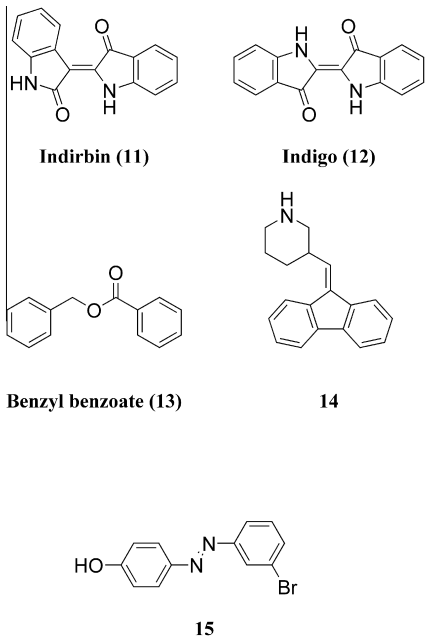


Figure 8. Aggregators. Despite their small size, these molecules can form aggregates that nonspecifically inhibit enzymes and protein–protein interactions.

compound **11** (Fig. 7) was reported to be a protein–protein interaction inhibitor, but flat SAR, discrepancies in different assays, and the presence of DTT in the assay buffer all suggest that this compound is acting by redox cycling.[37] Indeed, subsequent research identified this very molecule as a potent redox cycler.[38] Fortunately assays have been developed to weed out such compounds;[39,40] as new compounds are acquired it is essential to test them for such activity, particularly if they contain unusual heterocyclic ring systems.

Redox-active molecules and other PAINS are increasingly recognized as problematic, and Baell has recently released more than 1000 functional group and PAINS filters that can be used to cull such compounds from libraries.[14]

Aggregators. Even if molecules are stable, pure, and unreactive, they can still mislead: one of the most insidious problems is aggregation. Many small molecules can form microscopic aggregates in aqueous buffer, and these aggregates can non-specifically inhibit a variety of assays. The higher the concentration, the more likely aggregates are to form, so fragment-screening is particularly susceptible to this phenomenon. That said, aggregation can occur even at nanomolar concentrations of compounds. This phenomenon was not reported in the literature until 2002,[41] but its frequent occurrence probably dissuaded many people from pursuing low affinity fragments. Too often, unrecognized aggregators produced exciting initial results that proved recalcitrant to attempts at optimization.

Aggregation can occur with a wide variety of molecules, including fragment-sized ones (Fig. 8). The easiest means to avoid this problem is to add a non-ionic detergent such as Triton X-100 or Tween to the assay; doing so will almost always cause a dramatic decrease in activity for aggregators, but should have little or no affect on legitimate binders.[42–44] In one example, a quantitative HTS screen of 70,563 molecules found 1274 hits, of which 1204 lost activity in the presence of detergent.[45] However, detergent is not a panacea: Wells and co-workers described a class of molecules that form fibrils, even in the presence of detergent. Procaspase molecules bind to the fibrils, bringing them close together and apparently causing them to proteolytically activate one another.[46,47]

Aggregators typically show steep Hill slopes in dose–response assays, and rarely show competitive enzyme kinetics; of course, some legitimate inhibitors also show these properties. Aggregates can also be detected using electron microscopy, FACS, or dynamic light-scattering, and can sometimes be centrifuged. Finally, compounds that aggregate sometimes show shallow SAR, meaning that relatively significant structural changes may have little effect on activity.[48] Nonetheless, it is important to always remain vigilant: in one example, a miscommunication in the amount of detergent in the assay led to a compound series being optimized for aggregation rather than binding.[49] Some of the best compounds in this series displayed IC_{50} values of 200 nM under low-detergent conditions, but no activity at higher detergent levels.

Aggregators are probably one of the worst 'pollutants' in the chemical literature. For example, previously reported inhibitors of prostaglandin E2 synthase, including one that is sold commercially as a research tool, have recently been found to be aggregators,[50,51] calling into question any research done with these molecules. Similarly, several reported inhibitors of the anticancer target lactate dehydrogenase A have also been unmasked as aggregators; a paper reporting one of them has already been cited more than 100 times.[52] Unfortunately these are not isolated examples. Both of the authors have repeatedly seen molecules reported in the literature that, upon testing, turned out to be aggregators. Researchers need to convince themselves that their molecules are not aggregators, and reviewers and journal editors need to ensure that papers are as rigorous about their assays as they are about compound quality.

Assay behavior. Besides general problems that can occur with compounds and targets, each individual fragment-finding method has its own unhappy constellation of artifacts, despite what boosters of a particular technology may say. Indeed, a key to success in fragment screening is to use multiple orthogonal methods and focus on those hits that confirm in different methods. This can be a sobering exercise: in one recent case,[53] nuclear magnetic resonance (NMR) and surface plasmon resonance (SPR) screens of essentially the same library both yielded a number of hits, but very few of these were in common, and none of the crystallographically confirmed hits were picked up in both the primary NMR and SPR screens. Technique-specific problems are discussed below.

Nuclear magnetic resonance (NMR). NMR is widely regarded as the gold standard technique for characterizing intermolecular interactions in solution, and was in fact the first technique to demonstrate that fragment-based approaches to lead discovery could be practical.[6,54] Nonetheless, there are limitations and potential artifacts associated with each of the different experiments typically used to identify ligands. To some extent these problems can be circumvented by the application of multiple complementary NMR experiments, but the limitations of each method should be considered carefully when the aggregated data are analyzed. Some of the potential artifacts and issues associated with the more common NMR binding experiments are detailed in Table 1.[55]

A common limitation with ligand-observed NMR techniques is that the experiments are typically run with a large molar excess of ligand over receptor. Thus, the experiment detects the effect on the bulk ligand population of interactions between ligand and receptor; if the ligand dissociation rate from the receptor is slow, the net effect on the bulk population will be small and no binding will be observed. In practical terms, this means that ligands which bind tighter than high nM to low μM, depending on the exchange dynamics, may be missed. While such fragments are likely to be quite rare, they are probably the most interesting!

A second limitation with many of the ligand-observed NMR experiments is that no indication of the binding site is obtained. The use of a competitor ligand can provide some information, although partial displacement is frequently observed, and it is also possible for a ligand to cause allosteric displacement of an active-site probe.

Finally, since fragment screens are often run at high concentrations, it is possible for fragments that contain acidic or basic moieties to actually change the pH of the solution, leading to chemical shift perturbations which are simply the result of altered pH rather than fragment binding.

X-ray crystallography. A picture is worth a thousand words, according to the old cliché, but in fragment optimization a picture can be far more valuable. Crystallographic data are considered the gold standard for moving compounds forward: seeing exactly how a fragment binds can provide ideas on how to improve the affinity. But by the same token, it is easy to forget that a beautiful molecular model is just that—a model—and thus susceptible to artifacts and over-interpretation. As Andrew Davis et al. noted, 'many scientists who use structural information seem to be unaware of the fact that an X-ray crystal structure is one crystallographer's subjective interpretation of an experimental electron density map expressed in terms of an atomic model.'[57] For example, ambiguous or incomplete electron density can result in fitting small molecules or moieties in the wrong orientation.[58] In some cases simple inspection combined with chemical intuition could reveal the problem (for example, if a primary amide is reversed, contacts with surrounding residues may not make sense), but in other cases, particularly if the crystal is soaked with a racemic mixture, structures could be much harder to interpret. Sometimes solvent or buffer components are misinterpreted as the desired small molecule ligand. One group came close to publishing a structure of a fragment only to notice that the partial density was actually part of a long PEG chain running through the binding site (personal communication). Unfortunately, such cases are not uncommon, and are not always caught before being published.[59]

Ironically, these problems seem to be exacerbated for the more interesting structures. After analyzing all the structures deposited in the Protein Data Bank (PDB), Eric Brown and S. Ramaswamy reported that 'the most striking result is the association between structure quality and the journal in which the structure was first published. The worst offenders are the apparently high-impact general science journals. The rush to publish high-impact work in the competitive atmosphere may have led to the proliferation of poor-quality structures.'[60] Even journals such as *Nature* are not immune, as a recent example in which a structure was retracted from the PDB illustrates.[61] In another recent study, only 17% of 728 crystal structures previously used to validate computational docking methods passed a rigorous set of quality criteria.[59]

Another issue is that crystallography usually produces a single, static structure, and this could be significantly different than the structure adopted by the protein–ligand complex in solution. Moreover, individual molecules in a crystal may adopt different conformations, but these may be averaged out during crystallographic refinement. As Mark DePristo et al. noted, 'in many cases the problem is not selecting the best single conformation, but that several conformations are equally plausible interpretations of the electron-density map, especially at lower resolution.'[62]

Crystal packing contacts between individual protein molecules can also affect the binding sites of ligands. In some cases this may not be obvious when looking at the structure of just a single protein–ligand interaction, as the ligand could make direct or water-mediated interactions with a different (symmetry-related) protein in the crystal lattice.[63]

Perhaps the most fundamental but easily overlooked problem with crystallographic data is the fact that it provides no affinity information. Given the high concentrations of fragments some-

Table 1

Experiment	Species observed	Limitations or potential artifacts
Protein observed NMR		
^{15}N-^{1}H HSQC	Protein	Shifts resulting from solvent binding or pH changes. Spectral overlap
^{13}C-^{1}H HSQC	Protein	Shifts resulting from solvent binding or pH changes. Spectral overlap
Ligand observed NMR		
STD	Bound ligand (saturation occurs only in bound state)	False positives resulting from direct irradiation of ligand. False negatives resulting from incomplete saturation of protein. No binding site information
Water-LOGSY	Bound and free ligand (magnetization transfer occurs in both free and bound states)	False positives from self-association of ligands. Contribution from both free and bound populations. Chemical exchange. No binding site information[55]
Relaxation filtered 1D	Free ligand (bound ligand is attenuated)	False positives or negatives resulting from unusually short or long relaxation rates. No binding site information
ILOE	Ligands binding in close proximity to one-another	Aggregation of ligands can appear to be positive signal.[56] No binding site information
TINS	Free ligand (bound ligand is attenuated)	Severe line broadening from matrix. Heterogeneous phase Requirement for reference protein. No binding site information
^{19}F direct observe	Free ligand (bound ligand is attenuated)	Requirement for ^{19}F in fragment. No binding site information
^{19}F probe displacement	Indirect effect of bound ligand (probe is displaced from bound state)	Requirement for labeled probe molecule

times used for soaking experiments, it is possible to obtain co-crystal structures of fragments that show no detectable binding or inhibition in other assays. While these can be valuable, they can also identify fragments that bind so weakly that even heroic efforts would be wasted on advancing them.

On the other end of the spectrum, even well-behaved fragments may fail to yield crystallographic structures for a variety of reasons. The affinity and/or solubility of fragments in crystallization or soaking buffers may be considerably less than in other buffers. Often multiple crystal forms or soaking conditions may need to be investigated, and even then some binding sites may be occluded by crystal contacts and thus be inaccessible. Although such false negatives will not lead to the same wasted effort as do false positives, an over-reliance on crystallography could cause promising chemical series to be prematurely abandoned.

Isothermal titration calorimetry (ITC). ITC can provide thermodynamic binding parameters (ΔG, ΔH, and ΔS) for protein–ligand interactions. When performed carefully, this is an extremely powerful and accurate technique. However, systematic errors, such as incorrectly measured concentrations or failure to properly account for heats of dilution, can significantly affect the results.[64] In one notable example, 14 separate labs were each provided with a model system, bovine carbonic anhydrase II (CA II) and 4-carboxybenzenesulfonamide (CBS), and asked to use ITC to determine the enthalpy of binding and binding constant.[65] Although most of the values clustered relatively close together, there were significant outliers: the ΔH ranged from −4.3 to −15.4 kcal/mol (with the average being −10.4 kcal/mol) and even the stoichiometry ranged from 0.55 to 1.25 (with the average being 0.94). The fact that such variation is observed for a robust, well-studied model system in which participants were all using the very same reagents is a useful reminder that all data needs to be taken with a grain of salt, particularly in less well-defined systems.

Surface-plasmon resonance (SPR). As seen in Figure 1, SPR has become a dominant method for finding fragments. Typically, a target protein is immobilized on a chip, and varying concentrations of a small molecule are allowed to flow over it. When a ligand binds to the protein it causes a change in the reflective properties that depends on the ratio of the ligand mass to the protein mass. Though small fragments require very sensitive detection, many commercially available instruments are currently suitable for fragment screening.

Because SPR experiments are straightforward to set up and run, it is easy for novices to make mistakes or misinterpret their data. As Rebecca Rich and David Myszka noted after surveying 1413 papers that used SPR in 2008, 'less than 30% would pass the requirements for high-school chemistry'.[66] An analysis of the 1514 publications in 2009 that used SPR was equally discouraging: only 20% were given a passing grade by a panel of 22 experts.[67] Fortunately there are now good resources for how to establish an SPR screen; one of the best is a nearly 50-page book chapter by Tony Giannetti.[68]

Some of the same aggregation phenomena that disrupt biochemical screens can also muddle SPR screens. The problem is that aggregates can cause an SPR signal, and this can be misinterpreted as fragment binding by inexperienced users.[69] That said, legitimate fragments too can sometimes cause problems. For example, a fragment shown to bind to the enzyme Pin1 by both NMR and crystallography showed super-stoichiometric binding by SPR.[70] In this case the fragment was a relatively 'flat' aromatic heterocycle, and analogs that incorporated a stereocenter showed better behavior. It is possible that such 'three-dimensional' fragments may generally be less prone to aggregation, though the authors are unaware of any systematic study. Indeed, some fragments may behave poorly under some conditions with some targets but be well behaved with different targets. For example, a negatively charged protein may cause cationic fragments to bind nonspecifically around it, a phenomenon that Markku Hämäläinen calls 'selective promiscuous binders.'[71]

In theory, SPR can provide on-rates and off-rates, but in practice binding kinetics are rarely observed for fragments due to their low affinity. In fact, unusually slow dissociation can be a sign of aggregation.[69] That said, small fragment-sized molecules can show high affinity binding and slow kinetics; one 14-atom ligand of the protein D-amino acid oxidase (DAAO) binds with low nanomolar affinity with an off-rate on the order of 0.07 h^{-1}.[72] In another example, a fragment identified as binding to a kinase by ligand-observed NMR was found to have a dissociation constant of 90 nM by SPR.[73]

Another question, not limited to SPR, is how specific a fragment should be. Within a given target class, for example kinases, it is possible to optimize a promiscuous fragment to a highly selective molecule (though it is also possible to start with a selective fragment and lose selectivity).[74] However, sometimes an unrelated protein is used as a counter-screen to weed out promiscuous binders; this could unnecessarily exclude useful fragments that just happen to bind to both targets. For example, in a recent study, several fragments were shown by SPR to bind to two completely unrelated proteins, and these were initially classified as promiscuous. However, subsequent crystallography demonstrated that they did in fact bind to the protein of interest, HIV-1 integrase.[53] Indeed, due to the fact that fragments generally lack 'molecular complexity' it makes sense that certain privileged pharmacophores may bind to multiple proteins.[75]

Biochemical and functional assays. Enzymatic assays and binding assays such as fluorescence polarization (FP) or fluorescence resonance energy transfer (FRET) assays are simple and fast, and often useful as primary screening methods, but one needs to be willing to wade through a lot of false positives. For example, in an inhibition assay of the enzyme Pin1, 40 hits were identified, but only five could be confirmed by NMR.[76] Biochemical assays can fall victim to the full panoply of pathological impurities, reactive molecules, and aggregators discussed above. Finally, in the case of cascade assays involving multiple proteins one needs to verify that hits actually inhibit the target of interest rather than, say, an enzyme involved in producing a signal.

Thermal shift. As shown in Figure 1, thermal shift assays are widely used for fragment screening. A key advantage is that they are fast and inexpensive to perform: a protein is mixed with a putative ligand, heated, and the 'melting temperature' of the protein is measured; ligands typically stabilize their hosts against thermal denaturation, so molecules that raise the melting temperature are considered hits.[77] Operationally, the assay is typically conducted by adding a dye that changes its fluorescence properties when it binds to the denatured form of a protein. The screen can be conducted in microtiter plates using widely available instruments designed for RT-PCR.

Of all the methods for finding fragments, thermal shift assays seem to be the most controversial. At the FBLD 2012 meeting in San Francisco, some speakers described them as extremely unreliable, while others found them to be quite useful. Certainly some proteins are more suited to the technique than others. However, one needs to be cautious about using the technique as a primary assay: in a screen for stabilizers of mutant p53, Fersht and coworkers found a much lower hit rate than for an NMR method, and suggested that fluorescence quenching by fragments could lead to a high false-negative rate.[78] On a similar note, 14 of 15 fragments that bound to PARP15 as assessed by SPR were confirmed by NMR, while only one was confirmed by differential scanning fluorimetry.[79] False positives are an issue too: only 26 of 56 fragments that stabilized CYP121 towards thermal denaturation were confirmed by NMR to bind at the active site, though others may have bound outside the active site.[80]

No matter what your primary assay, it is essential to be aware of the potential pitfalls, particularly if you are new to the technology or to fragment-based screening. Ideally, fragment hits should be confirmed in at least two completely different assays before embarking on any sort of chemistry optimization. And of course, even after you have found validated fragments, there are many challenges to advancing these from low affinity ligands to leads and, eventually, to a clinical candidate.

To conclude, every step in any fragment screen can yield misleading information. Knowing about possible problems can help you recognize them before investing additional resources or embarrassing yourself publicly. Although it may seem paranoid, it is probably safest to assume any hit is guilty of being an artifact until proven innocent. As Richard Feynman noted, 'The first principle is that you must not fool yourself–and you are the easiest person to fool. So you have to be very careful about that.' Hopefully this Digest will help to illuminate some of the darker aspects of fragment screening. Only then can the real fun can begin.

Acknowledgements

The authors would like to thank the numerous colleagues and members of the community who provided anecdotes and data for this review, one of the reviewers of this manuscript for reminding us of the light-dependent inhibitors, Monya Baker for a careful reading of the manuscript, and Martin Davis for the cartoon graphical abstract.

References and notes

1. Jahnke, W., Erlanson, D. A., Eds.; Fragment-Based Approaches in Drug Discovery. In *Methods and Principles in Medicinal Chemistry*, series editors, Mannhold, R., Kubinyi, H., Folkers, G.; Wiley-VCH: Weinheim, Germany, **2006**. Vol. 34.
2. *Fragment-Based Drug Discovery: a Practical Approach*; Zartler, E., Shapiro, M., Eds.; John Wiley & Sons: Hoboken, NJ, **2008**.
3. Kuo, L. C., Ed.; Fragment-Based Drug Design: Tools, Practical Approaches, and Examples. In *Methods in Enzymology*, Elsevier: San Diego, **2011**; Vol. 317.
4. Davies, T., Hyvonen, M., Eds.; Fragment-Based Drug Discovery and X-ray Crystallography. In *Topics in Current Chemistry*, Springer: New York, **2012**; Vol. 317.
5. *Library Design, Search Methods, and Applications of Fragment-Based Drug Design*; Bienstock, R., Ed.; American Chemical Society: Washington, DC, **2012**.
6. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531.
7. Erlanson, D. A. Poll Results: Fragment Screening Methods. *Practical Fragments*. Published Online: **2011**. http://practicalfragments.blogspot.com/2011/10/poll-results-fragment-screening-methods.html.
8. Prinz, F.; Schlange, T.; Asadullah, K. *Nat. Rev. Drug Disc.* **2011**, *10*, 712.
9. Begley, C. G.; Ellis, L. M. *Nature* **2012**, *483*, 531.
10. Hann, M. M. *Med. Chem. Commun.* **2011**, *2*, 349.
11. Erlanson, D. A. Looking for Trouble. *Practical Fragments*. Published Online: **2013**. http://practicalfragments.blogspot.com/2013/01/looking-for-trouble.html.
12. Halford, B. *Chem. Eng. News* **2012**, *90*, 34.
13. Ward, R. A.; Brassington, C.; Breeze, A. L.; Caputo, A.; Critchlow, S.; Davies, G.; Goodwin, L.; Hassall, G.; Greenwood, R.; Holdgate, G. A.; Mrosek, M.; Norman, R. A.; Pearson, S.; Tart, J.; Tucker, J. A.; Vogtherr, M.; Whittaker, D.; Wingfield, J.; Winter, J.; Hudson, K. *J. Med. Chem.* **2012**, *55*, 3285.
14. Baell, J. B. *J. Chem. Inf. Model.* **2013**, *53*, 39.
15. Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N.-E.; Scott, N.; Petersen, A.; Gabriel, S.; Vu, B. T.; George, K. M.; Narayanan, A.; Li, S. H.; Qian, H.; Beatini, N.; Niu, L.; Gan, Q.-F. *ACS Med. Chem. Lett.* **2013**, 4.
16. Kozikowski, B. A.; Burt, T. M.; Tirey, D. A.; Williams, L. E.; Kuzmak, B. R.; Stanton, D. T.; Morand, K. L.; Nelson, S. L. *J. Biomol. Screen.* **2003**, *8*, 205.
17. Prochazkova, E.; Jansa, P.; Brezinova, A.; Cechova, L.; Mertlikova-Kaiserova, H.; Holy, A.; Dracinsky, M. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6405.
18. Baurin, N.; Aboul-Ela, F.; Barril, X.; Davis, B.; Drysdale, M.; Dymock, B.; Finch, H.; Fromont, C.; Richardson, C.; Simmonite, H.; Hubbard, R. E. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157.
19. Oldenburg, K.; Pooler, D.; Scudder, K.; Lipinski, C.; Kelly, M. *Comb. Chem. High Throughput Screening* **2005**, *8*, 499.
20. Erlanson, D. A. Fragment Problems, In *Fragment-based Drug Discovery group of LinkedIn*. Published Online: January **2013**.
21. Kozikowski, B. A.; Burt, T. M.; Tirey, D. A.; Williams, L. E.; Kuzmak, B. R.; Stanton, D. T.; Morand, K. L.; Nelson, S. L. *J. Biomol. Screen.* **2003**, *8*, 210.
22. McElhinny, C. J., Jr.; Lewin, A. H.; Mascarella, S. W.; Runyon, S.; Brieaddy, L.; Carroll, F. I. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6661.
23. Begley, D. Grooming Your Fragment Library: Checking the Quality of Commercially-Sourced Fragments. *Emerald Insights Blog*. Published Online: **2013**. http://www.emeraldbiostructures.com/blog/entry/grooming_your_fragment_ library_checking_the_quality_of_commercially_sourced.
24. Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Disc.* **2011**, *10*, 307.
25. Noe, M. C.; Gilbert, A. M. *Ann. Rep. Med. Chem.* **2012**, *47*, 413.
26. Rishton, G. M. *Drug Discovery Today* **2003**, *8*, 86.
27. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53*, 2719.
28. Baell, J. B. *Future Med. Chem.* **2010**, *2*, 1529.
29. Powers, J. P.; Piper, D. E.; Li, Y.; Mayorga, V.; Anzola, J.; Chen, J. M.; Jaen, J. C.; Lee, G.; Liu, J.; Peterson, M. G.; Tonn, G. R.; Ye, Q.; Walker, N. P.; Wang, Z. *J. Med. Chem.* **2006**, *49*, 1034.
30. Hughes, S. J.; Millan, D. S.; Kilty, I. C.; Lewthwaite, R. A.; Mathias, J. P.; O'Reilly, M. A.; Pannifer, A.; Phelan, A.; Stuhmeier, F.; Baldock, D. A.; Brown, D. G. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 6586.
31. Carter, P. H.; Scherle, P. A.; Muckelbauer, J. K.; Voss, M. E.; Liu, R. Q.; Thompson, L. A.; Tebben, A. J.; Solomon, K. A.; Lo, Y. C.; Li, Z.; Strzemienski, P.; Yang, G.; Falahatpisheh, N.; Xu, M.; Wu, Z.; Farrow, N. A.; Ramnarayan, K.; Wang, J.; Rideout, D.; Yalamoori, V.; Domaille, P.; Underwood, D. J.; Trzaskos, J. M.; Friedman, S. M.; Newton, R. C.; Decicco, C. P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11879.
32. Voss, M. E.; Carter, P. H.; Tebben, A. J.; Scherle, P. A.; Brown, G. D.; Thompson, L. A.; Xu, M.; Lo, Y. C.; Yang, G.; Liu, R. Q.; Strzemienski, P.; Everlof, J. G.; Trzaskos, J. M.; Decicco, C. P. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 533.
33. Reed, D.; Shen, Y.; Shelat, A. A.; Arnold, L. A.; Ferreira, A. M.; Zhu, F.; Mills, N.; Smithson, D. C.; Regni, C. A.; Bashford, D.; Cicero, S. A.; Schulman, B. A.; Jochemsen, A. G.; Guy, R. K.; Dyer, M. A. *J. Biol. Chem.* **2010**, *285*, 10786.
34. Bista, M.; Smithson, D.; Pecak, A.; Salinas, G.; Pustelny, K.; Min, J.; Pirog, A.; Finch, K.; Zdzalik, M.; Waddell, B.; Wladyka, B.; Kedracka-Krok, S.; Dyer, M. A.; Dubin, G.; Guy, R. K. *PLoS One* **2012**, *7*, e37518.
35. Guertin, K. R.; Setti, L.; Qi, L.; Dunsdon, R. M.; Dymock, B. W.; Jones, P. S.; Overton, H.; Taylor, M.; Williams, G.; Sergi, J. A.; Wang, K.; Peng, Y.; Renzetti, M.; Boyce, R.; Falcioni, F.; Garippa, R.; Olivier, A. R. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2895.
36. Tjernberg, A.; Hallen, D.; Schultz, J.; James, S.; Benkestock, K.; Bystrom, S.; Weigelt, J. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 891.
37. Yi, F.; Regan, L. *ACS Chem. Biol.* **2008**, *3*, 645.
38. Soares, K. M.; Blackmon, N.; Shun, T. Y.; Shinde, S. N.; Takyi, H. K.; Wipf, P.; Lazo, J. S.; Johnston, P. A. *Assay Drug Dev. Technol.* **2010**, *8*, 152.
39. Johnston, P. A. *Curr. Opin. Chem. Biol.* **2011**, *15*, 174.
40. Baell, J. B. *Drug Discovery Today* **2011**, *16*, 840.
41. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712.
42. McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 4265.
43. Ryan, A. J.; Gray, N. M.; Lowe, P. N.; Chung, C. W. *J. Med. Chem.* **2003**, *46*, 3448.
44. Feng, B. Y.; Shoichet, B. K. *Nat. Protocols* **2006**, *1*, 550.
45. Babaoglu, K.; Simeonov, A.; Irwin, J. J.; Nelson, M. E.; Feng, B.; Thomas, C. J.; Cancian, L.; Costi, M. P.; Maltby, D. A.; Jadhav, A.; Inglese, J.; Austin, C. P.; Shoichet, B. K. *J. Med. Chem.* **2008**, *51*, 2502.
46. Zorn, J. A.; Wille, H.; Wolan, D. W.; Wells, J. A. *J. Am. Chem. Soc.* **2011**, *133*, 19630.
47. Zorn, J. A.; Wolan, D. W.; Agard, N. J.; Wells, J. A. *J. Biol. Chem.* **2012**, *287*, 33781.
48. Shoichet, B. K. *Drug Discovery Today* **2006**, *11*, 607.
49. Ferreira, R. S.; Bryant, C.; Ang, K. K.; McKerrow, J. H.; Shoichet, B. K.; Renslo, A. R. *J. Med. Chem.* **2009**, *52*, 5005.
50. Koeberle, A.; Zettl, H.; Greiner, C.; Wurglics, M.; Schubert-Zsilavecz, M.; Werz, O. *J. Med. Chem.* **2008**, *51*, 8068.
51. Wiegard, A.; Hanekamp, W.; Griessbach, K.; Fabian, J.; Lehr, M. *Eur. J. Med. Chem.* **2012**, *48*, 153.
52. Kohlmann, A.; Zech, S. G.; Li, F.; Zhou, T.; Squillace, R. M.; Commodore, L.; Greenfield, M. T.; Lu, X.; Miller, D. P.; Huang, W. S.; Qi, J.; Thomas, R. M.; Wang, Y.; Zhang, S.; Dodd, R.; Liu, S.; Xu, R.; Xu, Y.; Miret, J. J.; Rivera, V.; Clackson, T.; Shakespeare, W. C.; Zhu, X.; Dalgarno, D. C. *J. Med. Chem.* **2013**.
53. Wielens, J.; Headey, S. J.; Rhodes, D. I.; Mulder, R. J.; Dolezal, O.; Deadman, J. J.; Newman, J.; Chalmers, D. K.; Parker, M. W.; Peat, T. S.; Scanlon, M. J. *J. Biomol. Screen.* **2013**, *18*, 147.
54. Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 5818.
55. Hubbard, R. E.; Davis, B.; Chen, I.; Drysdale, M. J. *Curr. Top. Med. Chem.* **2007**, *7*, 1568.
56. Sledz, P.; Silvestre, H. L.; Hung, A. W.; Ciulli, A.; Blundell, T. L.; Abell, C. *J. Am. Chem. Soc.* **2010**, *132*, 4544.
57. Davis, A. M.; St-Gallay, S. A.; Kleywegt, G. J. *Drug Discovery Today* **2008**, *13*, 831.
58. Malde, A. K.; Mark, A. E. *J. Comput. Aided Mol. Des.* **2011**, *25*, 1.
59. Warren, G. L.; Do, T. D.; Kelley, B. P.; Nicholls, A.; Warren, S. D. *Drug Discovery Today* **2012**, *17*, 1270.
60. Brown, E. N.; Ramaswamy, S. *Acta Crystallogr., Sect. D Biol. Crystallogr.* **2007**, *63*, 941.
61. Chen, L.; Yang, S.; Jakoncic, J.; Zhang, J. J.; Huang, X.-Y. *Nature* **2011**, *476*, 240. http://dx.doi.org/10.1038/nature10281.
62. DePristo, M. A.; de Bakker, P. I.; Blundell, T. L. *Structure* **2004**, *12*, 831.

63. Sondergaard, C. R.; Garrett, A. E.; Carstensen, T.; Pollastri, G.; Nielsen, J. E. *J. Med. Chem.* **2009**, *52*, 5673.
64. Tellinghuisen, J.; Chodera, J. D. *Anal. Biochem.* **2011**, *414*, 297.
65. Myszka, D. G.; Abdiche, Y. N.; Arisaka, F.; Byron, O.; Eisenstein, E.; Hensley, P.; Thomson, J. A.; Lombardo, C. R.; Schwarz, F.; Stafford, W.; Doyle, M. L. *J. Biomol. Technol.* **2003**, *14*, 247.
66. Rich, R. L.; Myszka, D. G. *J. Mol. Recognit.* **2010**, *23*, 1.
67. Rich, R. L.; Myszka, D. G. *J. Mol. Recognit.* **2011**, *24*, 892.
68. Giannetti, A. M. *Methods Enzymol.* **2011**, *493*, 169.
69. Giannetti, A. M.; Koch, B. D.; Browner, M. F. *J. Med. Chem.* **2008**, *51*, 574.
70. Potter, A. J.; Ray, S.; Gueritz, L.; Nunns, C. L.; Bryant, C. J.; Scrace, S. F.; Matassova, N.; Baker, L.; Dokurno, P.; Robinson, D. A.; Surgenor, A. E.; Davis, B.; Murray, J. B.; Richardson, C. M.; Moore, J. D. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 586.
71. Erlanson, D. A. FBDD Down Under. *Practical Fragments*. Published Online: **2012**. http://practicalfragments.blogspot.com/2012/11/fbdd-down-under.html.
72. Lange, J. H.; Venhorst, J.; van Dongen, M. J.; Frankena, J.; Bassissi, F.; de Bruin, N. M.; den Besten, C.; de Beer, S. B.; Oostenbrink, C.; Markova, N.; Kruse, C. G. *Eur. J. Med. Chem.* **2011**, *46*, 4808.
73. Hubbard, R. E.; Murray, J. B. *Methods Enzymol.* **2011**, *493*, 509.
74. Bamborough, P.; Brown, M. J.; Christopher, J. A.; Chung, C. W.; Mellor, G. W. *J. Med. Chem.* **2011**, *54*, 5131.
75. Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856.
76. Potter, A.; Oldfield, V.; Nunns, C.; Fromont, C.; Ray, S.; Northfield, C. J.; Bryant, C. J.; Scrace, S. F.; Robinson, D.; Matossova, N.; Baker, L.; Dokurno, P.; Surgenor, A. E.; Davis, B.; Richardson, C. M.; Murray, J. B.; Moore, J. D. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6483.
77. Kranz, J. K.; Schalk-Hihi, C. *Methods Enzymol.* **2011**, *493*, 277.
78. Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R. *Chem. Biol.* **2010**, *17*, 46.
79. Erlanson, D. A. Seventh Annual Fragment-Based Drug Discovery Meeting. *Practical Fragments*. Published Online: **2012**. http://practicalfragments.blogspot.com/2012/04/seventh-annual-fragment-based-drug.html.
80. Hudson, S. A.; McLean, K. J.; Surade, S.; Yang, Y. Q.; Leys, D.; Ciulli, A.; Munro, A. W.; Abell, C. *Angew. Chem., Int. Ed.* **2012**, *51*, 9311.