Design principles for fragment libraries – Maximizing the value of learnings from Pharma fragment based drug discovery (FBDD) programs for use in academia

György M. Keserű[1]*, Daniel A. Erlanson[2], György G. Ferenczy[1], Michael M. Hann[3], Christopher W. Murray[4], Stephen D. Pickett[3]

[1]Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok körútja 2., H-1117, Budapest, Hungary

[2]Carmot Therapeutics, Inc. 409 Illinois Street, San Francisco, CA 94158, USA

[3]GlaxoSmithKline Medicines Research Centre, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, U.K.

[4]Astex Pharmaceuticals, 436 Cambridge Science Park, Milton Road, Cambridge, CB4 0QA, U.K.

*email: gy.keseru@ttk.mta.hu

Abstract

Fragment-based drug discovery (FBDD) is well suited for discovering both drug leads and chemical probes of protein function: it can cover broad swaths of chemical space and allows the use of creative chemistry. FBDD is widely implemented for lead discovery in industry, but is sometimes used less systematically in academia. Design principles and implementation approaches for fragment libraries are continually evolving, and the lack of up-to-date guidance may prevent more effective application of FBDD in academia. This Perspective explores many of the theoretical, practical, and strategic considerations that occur within FBDD programs, including the optimal size, complexity, physicochemical profile, and shape

profile of fragments in FBDD libraries, as well as compound storage, evaluation, and screening technologies. This compilation of industry experience in FBDD will hopefully be useful for those pursuing FBDD in academia.

Rules are for the obedience of fools and the guidance of wise men.

Harry Day, Royal Air Force (1898-1977)

Introduction

As academia expects to play an increasing role in target validation and drug discovery, lessons that have been hard won in industry must become available so that they can be adopted and built on, and not poorly implemented or wastefully reinvented. Fragment-based drug discovery (FBDD) has emerged over the last 20 years as a highly successful way to find quality leads for subsequent optimization into drug candidates and approved new medicines. FBDD involves screening compounds that generally have fewer heavy atoms than is typical in a high throughput screening collection. Hit-identification methods must be adapted for the smaller size of fragments, and so require sensitive biophysical methods or higher concentration biochemical assays.

In this Perspective we have tried to collate the collective experiences of authors from several companies intimately involved in the development of FBDD so as to help the wider adoption and effective implementation of FBDD in academia and smaller start-ups.

In theory, a well-chosen fragment library samples an astronomically greater proportion of chemical space than a well-chosen high throughput screening (HTS) library can ever do.[1,2]

This gives researchers the confidence that there will be fragments in the library that sample chemical space as thoroughly as possible. As a result, fragment hits form high-quality interactions with the target, usually a protein, despite being weak in potency.[3] In fact, hit rates from fragment screens are sometimes used to assess the chemical tractability of a target. Fragment hits often show unique, often enthalpy-driven thermodynamic binding profiles.[4] Even small fragments offer enormous scope for growing into larger molecules.

Structure-based drug design (SBDD) is nearly always used to help focus synthetic exploration of the fragment for ways to boost binding affinity and guide other properties needed for a successful drug.[5] The careful use of SBDD allows medicinal chemists to control the physical properties of the growing fragment, ensuring that any additional molecular weight and lipophilicity also produces an acceptable increase in affinity.[6] Ligand efficiency metrics are often used to judge whether increases in affinity are acceptable.[7] As mentioned above, an alternative use of fragments is to assess the chemical tractability of a target,[8] but the design and screening principles discussed here are broadly the same for such usage.

Design principles for fragment libraries

In 2003, Congreve *et al.* analyzed the results from a relatively small number of fragment screens and proposed that "hits seem to obey, on average, a "Rule of Three (RO3)" in which molecular weight <= 300 Daltons, the number of hydrogen bond donors <=3, the number of hydrogen bond acceptors is <=3 and cLogP is <=3".[9] The molecular weight and cLogP criteria were enthusiastically adopted by the fragment community, whereas the hydrogen bond criteria have been less frequently used, in part because of ambiguities in the way

acceptors and donors are defined.[10] Ten years later, the RO3 concept is still widely employed, but the successful design of fragment libraries incorporates many other factors, informed by years of fragment screening.

The first consideration is the number of compounds to be included in the library; this is in part driven by the detection technology to be used in fragment screening. Higher throughput techniques, such as high concentration screening (HCS) using for example a biochemical assay, are not usually as sensitive as lower throughput biophysical techniques. In general, less sensitive techniques require more potent fragments which are likely to be more complex (that is, larger) compounds. (See fragment complexity section below.)[11] This in turn requires larger libraries, as the probability of any one fragment being a hit exponentially declines with the increased complexity of the ligand. More typically, fragment libraries are screened using sensitive biophysical techniques and the fragment library need only consist of a few thousand compounds with molecular weights between about 140 and 230 Daltons.

Other factors that will be discussed in detail in this article include ensuring that libraries:

(i) Sample relevant chemical space by including key pharmacophores that can drive fragment binding.

(ii) Contain an appropriate size distribution and a balance of differently shaped fragments (see section on 3D metrics) of appropriate complexity. Overly complex fragments have decreased hit rates due to functionalities that interfere with binding; overly flexible ones could have lower affinities due to the entropic costs of binding.

(iii) Contain a diversity of synthetically accessible growth vectors so that fragment hits can be effectively optimized into lead compounds.[12]

(iv) Avoid groups known to be associated with high reactivity, aggregation in solution, or persistent false positive data.[13]

Additionally, fragment libraries will typically be subjected to extensive initial and ongoing quality control (QC) analyses under conditions relevant to both the storage and screening of the samples. These should include quantitative assessments of the purity, identity, stability and solubility of the fragments, together with more specific analyses such as the aggregation properties of the fragments.

For difficult targets, such as protein-protein interactions, a good library offers the highest chance of identifying fragment hits that might be optimizable into lead compounds. For more tractable targets, even relatively poorly constructed libraries will yield some fragment hits, but the optimization of those hits into high quality lead compounds will still be heavily dependent on the quality and diversity of the output from the fragment screen.

Screening technologies

Technologies used for fragment screening must be compatible with the smaller size, reduced complexity and consequently lower affinity of fragments. While HTS assays often identify compounds with strong affinity (10 μM is a typical lower limit), fragment hits typically have weaker affinities in the 100 μM – 10 mM range, and thus the screening methodology must provide 100-1000 fold higher sensitivity. Thus, while HTS routinely use biochemical assays, fragment screening generally utilizes more sensitive biophysical technologies, as demonstrated by polls conducted on the blog *Practical Fragments* (Figure 1.)[14]

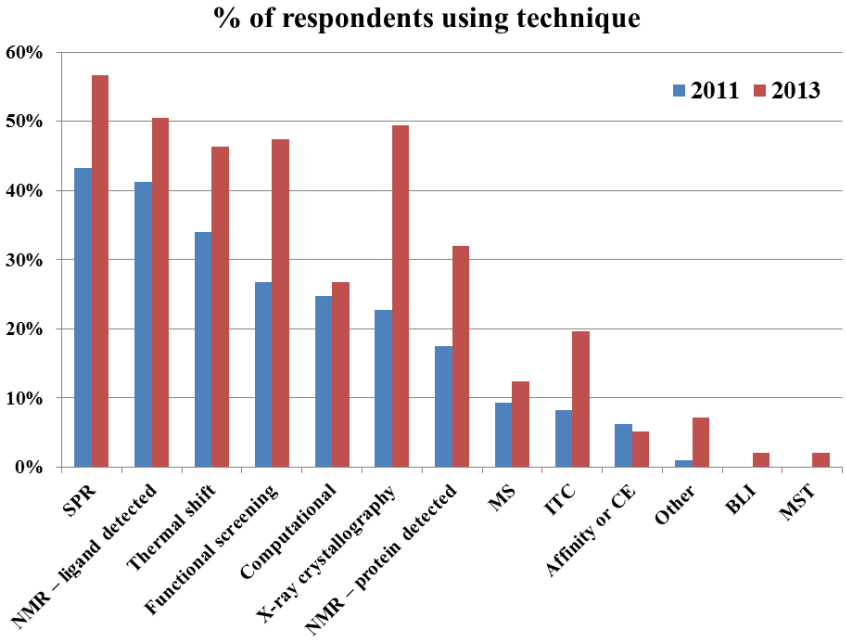


Figure 1. Poll results on the *Practical Fragments* blog on the use of fragment screening techniques. SPR – surface plasmon resonance based biosensor technology, MS – mass spectrometry, ITC – isothermal titration calorimetry, Affinity – (AC) affinity chromatography, CE – capillary electrophoresis, BLI – biolayer interferometry, MST - microscale thermophoresis. Results were collected from 97 responders split roughly evenly between academia and industry.

In 2011 SPR and ligand-detected NMR were clearly the most popular techniques followed by the thermal shift assay. The overall picture had not changed very much in 2013; however, the use of functional biochemical screening and X-ray crystallography was clearly increasing. The increase in functional biochemical screening may be due to improved diligence in avoiding false positives, the ease of running biochemical screens and/or the desire to use an orthogonal screen to a biophysical method. X-ray crystallography is less often used for

primary screening because of its lower throughput. The increasing use of X-ray crystallography is therefore more likely attributable to the importance of structural information for effective FBDD. It should be noted, however, that X-ray crystallography techniques are moving towards higher-throughput as exemplified by the CSIRO Collaborative Crystallization Centre in Melbourne, Australia. Similar reasoning applies to protein-detected NMR and ITC reaching 30% and 20%, respectively in 2013. These methods are also often used as follow-up assays (structural and thermodynamic characterization) rather than primary screening technologies due to the demands for protein labeling and/or quantity.

Those pursuing FBDD increasingly use orthogonal methods to assess hits: the average number of methods used increased from 2.4 in 2011 to 3.6 in 2013.[14] Of course, only so much can be read into a self-selected poll on a blog. However, similar results were observed when Swain analyzed 165 published fragment screening programs reporting 620 hits against 116 different targets. In particular, individual programs employ multiple techniques (Figure 2).[15]

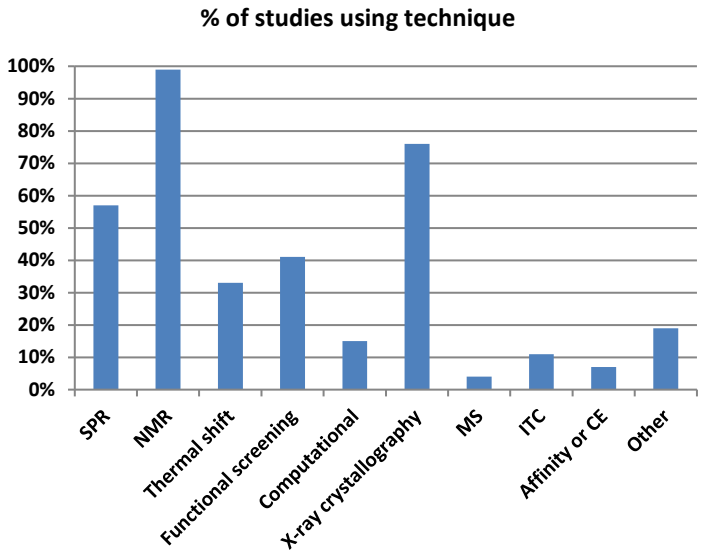


Figure 2. Fragment screening technologies used in published FBDD studies (adapted from ref. 15, November 2013)

Based on these analyses it seems that the major screening technologies in use for FBDD are SPR, ligand- and protein-detected NMR, X-ray crystallography, thermal shift and biochemical assays. Since all of these methods have been extensively reviewed previously[13,16] here we focus only on compound library related issues such as assay sensitivity (i.e., the ability to detect weak hits), specificity (i.e., the ability to discriminate between separate binding events), throughput and hit rate (Table 1). The propensity for false positives / false negatives in particular can be operator-dependent and even subjective and depends upon a definition of "hit." For example, biochemical screens will tend to have a higher false negative rate compared to NMR because of the lower sensitivity rather than a particular issue that can be solved by assay design. On the other hand, the false positive rate can be strongly influenced by assay format and compound selection.

Method	**Sensitivity limit**	**Specificity assessment**	**Throughput**	**Structural information**	**Propensity for false positive/ false negative**
Biochemical	high μM	+	high	none	high FP/FN
Ligand-NMR	low mM	-	medium	some	medium FP
Protein-NMR	low mM	+	low	high	low FP/FN
SPR	high μM	+	medium	none	medium FP
Thermal shift	high μM, low mM	-	high	none	high FP/FN

X-ray	mid mM	+	low	high	low FP/ high FN

Table 1. Sensitivity limit, feasibility of assessing specificity, throughput, structural information and false positive (FP) and false negative (FN) information on technologies most frequently used for fragment screening. The table gives an indication of typical output from properly configured experiments.

The use of particular libraries for different screening technologies is not strictly necessary as reflected in general fragment libraries commercialized by many vendors and screened in multiple FBDD programs. However, the different sensitivities and specificities of the multiple technologies do impact the overall design principles of libraries (Figure 3). The sensitivity and specificity of the given screening technology determines the assay concentration that should be used for the identification of active fragments. To reach this concentration the library should contain fragments with adequate solubility. Since high solubility is reliant on appropriate physicochemical properties, this requirement inevitably limits the distribution of MW and logP of the library screened. Consequently, the affinity of fragment hits that can be found is impacted by the screening method used because of the requirements of the initial library selection.

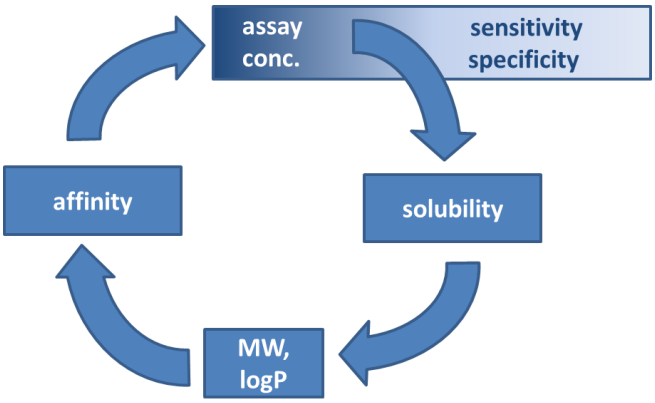


Figure 3. Relationship of assay sensitivity and specificity and physicochemical properties of the screening library. Sensitivity and specificity of the assay technology affect the solubility requirements for the library through the assay concentration needed for the detection of a binding event. These in turn influence the molecular weight, logP and likely affinity range of fragments from the library that sets limits for the assay concentration used to identify suitable hits.

However, varying attributes of the most common screening technologies (summarized in Table 1) do affect the kinds of hits, false positives, and false negatives that will emerge from screening a library.

Different levels of sensitivity will influence the optimal concentration of the compounds to be screened. NMR, X-ray and thermal shift assays are routinely performed with millimolar compound concentrations, while SPR and especially biochemical assays use high micromolar concentrations. In addition, some screening methods are more tolerant of lower solubility compounds than others. For example, X-ray crystallography requires compounds to have high occupancy in the protein in order for a hit to be detected and thus requires high solubility, whereas ligand-observed NMR can identify fragments that bind with much lower occupancy.

Thus, the sensitivity and specificity of the assay technologies can require different physicochemical properties of the fragment library. In addition, the false positive/false negative rates are influenced by the characteristics of the fragment library screened. Many false hits are due to a compound's interference with detection technology in a way that wrongly suggests interaction with the target protein.

Promiscuity is generally considered to be a deleterious property of drug leads except in cases where polypharmacology may be desired. However, in the context of fragment screening, the situation is more complicated. One source of promiscuity can arise from the fragment itself, whereby some types of molecules (e.g. promiscuous 2-aminothiazoles, or PrATs)[17] are frequent hitters in a broad range of assays. However, as some approved drugs contain 2-aminothiazoles, it is possible to evolve selectivity from such promiscuous starting points. The fact that an aminothiazole has many H-bond interaction opportunities suggests that substitution on one of the donor or acceptor atoms may well give the first step towards selectivity. Certain small halogenated fragments such as 4-bromopyrazole have also been reported as being promiscuous.[18] The opportunity for building in selectivity to promiscuous kinase fragment hits has been widely demonstrated[19] while Merck has provided a further example of this approach with the phosphodiesterase PDE10.[20]

It is important to distinguish between such promiscuous fragments that may bind selectively to multiple sites on different proteins from fragments that act pathologically. These "frequent hitters" include PAINS (Pan Assay INterference compounds), which often react covalently with proteins or interfere with the assay, [21,22] and aggregators, which are especially problematic at high concentrations.[23] Although the structural moieties in PAINS were identified in a biochemical assay using only one detection technology, these compounds have been reported to be active in many different assays and are often missed by reactivity filters.[21]

Finally, the throughput of the screening technology obviously impacts the size of the library screened. These factors together are reflected in the hit rates observed for fragment libraries screened against multiple targets using different screening technologies (Table 2).

Library	Number of fragments	Technology	# screens	Hit rate			Ref.
				Min	Max	Average	
AstraZeneca	600-40000	Biochem	7	0.2%	33%	7%	24
Genentech	Not stated	SPR	13	2.6%	14.9%	7.5%	25
GSK kinase*	1064	Biochem	30	2.2%	62.2%	26.2%	26
Novartis	2826	NMR	Not stated	3%	30%	2-8%	27
Pfizer	2592	NMR	13	2.8%	12.6%	6.9%	28
Vernalis	1063	NMR	12	0.4%	7.3%	3.2%	29

Table 2. Hit rates observed for fragment libraries screened against multiple targets using different screening technologies. *Note the GSK kinase library was specifically selected to have likely kinase binding motifs, hence the high hit rate.

It should also be noted that the definition of a screening hit is subjective and often influenced by the goals and constraints of the project as well as technical skills and expertise of relevant team members.

While using screening technologies in parallel can compensate for weaknesses in each assay used, selection of particular screening technologies is not always straightforward since performance depends on the assay conditions, the library and the nature of the target itself. In the next section we summarize case studies using different technologies, screening the same library against the same target. These case studies illustrate the value of orthogonal assays to help pinpoint true hits. However, some promising starting points may not appear active in all of the assays, so judicious follow-up is required.

Case studies with two or more screening technologies

- **HIV integrase:** Wielens and coworkers screened a 500 member fragment library against HIV integrase using STD-NMR and SPR.[30] NMR screening was performed with 50 cocktails of 10 compounds (each at a concentration of 1 mM) and identified 84 active fragments from which 62 were validated by protein-observed NMR experiments. In a parallel screening effort, the same library was tested as individual compounds by SPR at 500 μM concentration and yielded 16 hits validated by concentration-dependent SPR studies. Interestingly, no hits were confirmed by both the STD-NMR and SPR assays. Furthermore, the 6 hits co-crystallized from the 16 SPR hits did not overlap with the 15 hits co-crystallized from the 62 NMR hits. In this case the SPR hit rate was significantly lower (3%) than that from NMR (13%). In addition to different screening concentrations and buffers (pH 7.4 vs pH 8.5 for SPR and NMR screening, respectively) assay sensitivity might be one of the major reasons for the significantly different hit rates; the NMR screen was more sensitive to weak binders.
- **Checkpoint Kinase 2:** Montfort *et al.* investigated the performance of biochemical and thermal shift assays by screening 1869 fragments against Checkpoint Kinase 2.[31] Biochemical screening utilized an AlphaScreen assay performed at 300 μM compound concentration and resulted in 45 non-aggregating hits out of which 20 did not cause assay interference as validated in dose-response studies. The thermal shift assay was carried out at a considerably higher compound concentration of 2 mM and gave 63 hits, a significantly higher hit rate (3.4%) as compared to that of the AlphaScreen assay (1.1%). Comparing the hit lists, 12 fragments were identified by both assays and interestingly the authors found good correlation between ΔT_m and IC_{50} values. The

majority of hits, however, were found by only one technology. The biochemical and thermal shift assays respectively found 31 and 49 non-overlapping active fragments. Crystallization efforts were significantly more successful for overlapping hits since 8 of the 12 dual hits were confirmed by X-ray crystallography compared to only one fragment out of the 49 identified by the thermal shift assay alone. The need for different screening concentrations was due to the limited DMSO tolerance of the biochemical assay although it showed generally higher sensitivity towards less active hits.

- **p38α kinase:** The BioFocus fragment library was screened against p38α kinase using SPR and two biochemical assays including a mobility shift and fluorescence life time assays.[32] 266 fragments from the BioFocus library were screened by SPR in 200 μM and 1 mM compound concentrations and gave 102 primary actives. The mobility shift assay was performed in both 200 μM and 1 mM concentrations and identified 39 validated hits. In analyzing the hit lists the authors concluded that only 9 out of the 102 SPR hits were found by the mobility shift assay. In contrast, there was only one fragment out of the 10 identified by the biochemical assay that was missed by SPR. Hits identified by the mobility shift assay were also investigated by a fluorescence life time assay that confirmed 80% of them. These results indicate that the hits from the two biochemical assays correlate well but these are very different from those of the SPR assay.
- **MMP12 and Trypsin:** Boettcher and coworkers tested 352 fragments at 1 mM concentration using three biochemical (fluorescence life time**,** fluorescence intensity, mobility shift) and two biophysical tests (SPR and protein- detected NMR) against Matrix metalloprotease MMP12 and Trypsin.[33] The authors did not compare the hit

lists quantitatively but focused the comparison on false positive (FP) and false negative (FN) rates compared to NMR actives as reference (Table 3).

Technology	**MMP12**		**Trypsin**	
	FP	**FN**	**FP**	**FN**
FLT	13%	0%	13%	8%
FI	5%	7%	8%	13%
MSA	7%	15%	3%	9%
SPR	11%	7%	3%	27%

Table 3. False positive (FP) and false negative (FN) rates relative to NMR observed for different screening technologies (FLT – fluorescence life time, FI – fluorescence intensity, MSA – mobility shift assay) used for fragment screening against MMP12 and Trypsin

Comparison of FP and FN data revealed that biochemical assays performed similarly and SPR was comparable to them. The high false negative rate in the trypsin SPR assay was attributed to the weak response level due to low protein loading of the chip. The authors emphasized the importance of high solubility – influenced by the composition of the library - that should be determined by NMR rather than using a dynamic light scattering (DLS) assay.

- **HSP90**: Hubbard *et al*. screened 111 Vernalis fragments against HSP90 using NMR, SPR, thermal shift and biochemical assays.[34] The thermal shift assay (TSA) was conducted at 2 mM concentration; all the other tests were run at 500 μM. The authors analyzed the correlation between different hit lists (Table 4).

	SPR	Biochem	TSA
NMR	90%	74%	74%
SPR	-	73%	76%
Biochem		-	73%

Table 4. Overlap between the hit lists of different screening technologies used for the identification of HSP90 inhibitors

These results suggest that hits identified by NMR and SPR largely overlapped, but the hits of the third biophysical method (TSA) differed somewhat. This difference can be rationalized by the lower sensitivity of the thermal shift assay. Similarly, the correlation between biochemical and biophysical hits was also lower.

- **Endothiapepsin:** Schiebel *et al*. screened a 361 member fragment library against endothiapepsin using biochemical and two biophysical techniques (STD NMR and thermal shift assay) in 1mM, 300 μM and 2.5 mM concentrations, respectively.[35] The highest hit rate was found for biochemical screening (17%) followed by NMR (11%) and thermal shift assay (8%). It is important to note, however, that the authors neither tested assay interference nor aggregate formation. Hit lists obtained by the three screening methods were compared (Figure 4).

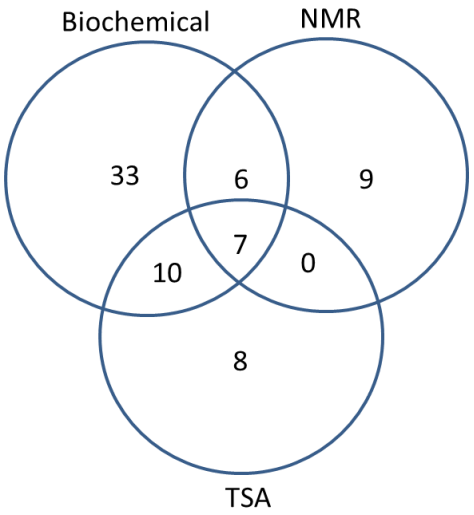


Figure 4. Venn diagram of biochemical, NMR and thermal shift hit lists obtained by screening against endothiapepsin

The different assays were performed under similar conditions including the buffers. Screening concentrations and also hit selection criteria, which are usually based on previous screening experience, were different. More importantly, however, the methodologies applied show very different sensitivity in detecting weak binders. These factors likely explain the limited overlap between the hits lists. The authors used the hit lists from all three assays to pick which compounds to explore further with crystallography. In contrast to the biochemical hits, all of the unique NMR and more than half of the unique thermal shift hits were crystallized successfully.

- **Multitarget studies:** In a large scale study a Novartis team reported screening a 1400 member fragment library in 35 campaigns using biophysical (NMR, SPR, TSA) and biochemical (high concentration screening, CE, TR-FRET) assays.[36] One of the major conclusions of this study is that different screening technologies provide hits with different physicochemical profiles on a given target, although there is a pronounced target dependency as well: NMR identified larger and more hydrophobic hits than TR-

FRET on one target but this was reversed on another. Although hit rates generally increased for larger and more lipophilic molecules, SPR hit rates were typically higher than TSA hit rates. This was partially rationalized by the non-specific binding of fragments to the chip surface. The lower hit rates obtained by thermal melting are likely a result of uncertainty as to how to handle results indicating thermal stabilization with lower T_m when a ligand binds. The complementary benefits of using both a biochemical and a biophysical assay are highlighted as being particularly effective in the detection of genuine fragment hits. (This has also been pointed out by the GSK group [37] in a similar meta-analysis which incorporates a broad discussion of the use of FBDD techniques.) The Novartis group does conclude that some fragments are "frequent hitters" although they consider them as "privileged fragments", because there is often an opportunity to build selectivity while evolving fragments.

Analyzing the active fragments from 35 campaigns on 20 targets it was interesting to see that 63% of the screened fragments never came up as hits. In contrast, the team identified privileged fragments that were found to be high value library members because they were active on more than one target. For more detailed analysis the authors selected two campaigns that used five different screening technologies and investigated all of their combinations. This analysis revealed that there is better complementarity between biochemical and biophysical tests than two biophysical methods. Therefore the authors suggested that a combination of biochemical and biophysical technologies would provide greater chemical coverage. It seems likely that the limited overlap of biochemical and biophysical methods is strongly related to the different sensitivity and specificity of these techniques, and that should be considered when designing screening libraries. On the other hand, higher assay concentrations

available for biophysical tests might generate higher numbers of false positives that could be reduced by the use of orthogonal techniques.

It is important to emphasize that a legitimate hit may be missed in orthogonal assays, in particular when one method is more sensitive than another. Demanding agreement in all assays will limit detection to the least sensitive method, which may not be desirable, since even extremely weak fragments with no functional activity can be progressed to potent leads.[38]

Fragment complexity and ligand efficiency

Hann *et al*. introduced the concept of Molecular Complexity and its application to FBDD as a formalism for helping to understand the challenges and opportunities of fragments.[11] This concept was inspired by examining protein-ligand interactions in crystal structures, which revealed that most interactions observed were productive. Unfavorable interactions either completely prevent binding or require ligands and proteins to adopt suboptimal (high energy) configurations. This is observed most obviously if water is considered to be the ultimate high concentration fragment, present at 55.5M! Tightly bound water molecules maximize their number of interactions (up to four) and water networks are built around these – see for example the combined use of theory and experiment to explore such networks in a GPCR.[39] The complexity model separates the probability of a maximally correct binding match of possible interacting pharmacophore points from the separate probability of whether the number and type of these interactions releases sufficient free energy to be measured in a biophysical experiment (Figure 5).

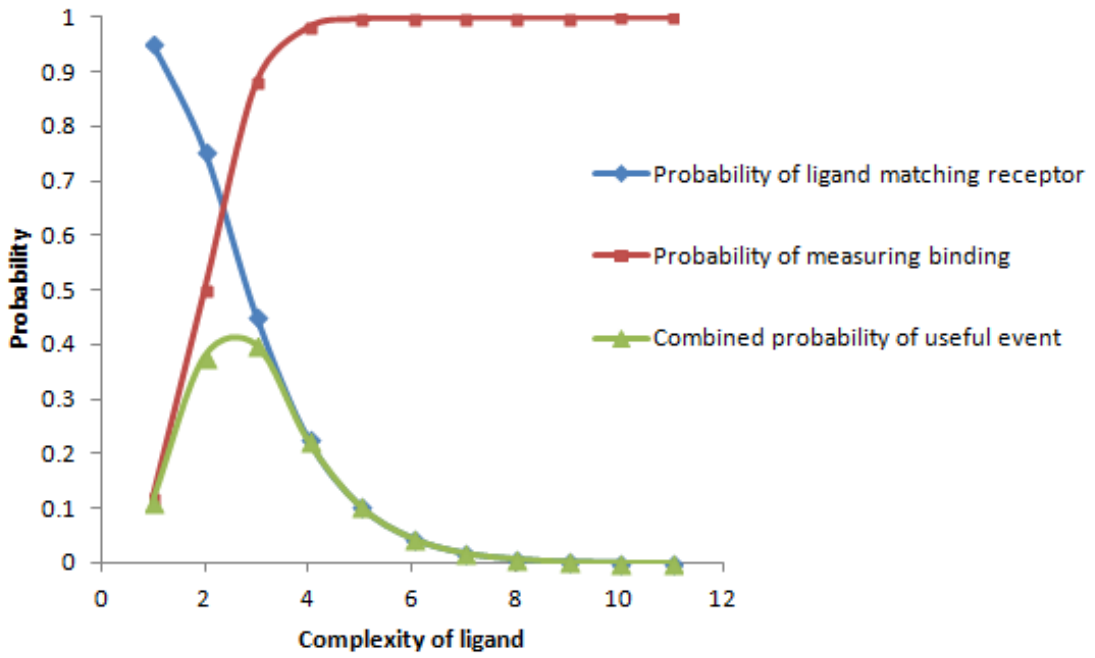


Figure 5. Probabilities of binding and detection as a function of molecular complexity. The blue curve represents the probability of a ligand matching the receptor as the complexity of the ligand increases. The red curve represents the probability of measuring the binding of a ligand to the receptor as the matches increase. The green curve represents the probability of a "useful event" which is defined as the product of the red and blue curves.

Hann *et al*. pointed out that there is a higher probability of a match if there are fewer interactions to get right, and that this is most easily achieved when fragments can make only a small number of interactions. Several groups have observed this relationship in their programs. The Astex group2 showed that the modal value of the Heavy Atom Count (HAC) of hits detected from many different fragment campaigns is 12 whereas the modal HAC value of the fragment library actually screened is 14. The Novartis group[27] observed HTS hit rates of 0.001% - 0.151% in the identification of ligands with an IC_{50} threshold in the micromolar range and fragment hit rates of 3% or more in NMR screening of fragments with an affinity

threshold in the millimolar range. Consequently, the complexity model has been widely adopted as a central mantra of the fragment approach and a number of publications have looked at both its extension[40] and further validation.[41,42]

One common question is how to relate the complexity model – with its abstract representation of molecular interaction points – to the reality of protein-ligand interactions. Because of the subtlety and actual complexity of real situations, it is best to use Molecular Complexity as a model and not to try to relate the matches to any specific type or strength of molecular interactions. However, what is apparent is that molecular interactions that have higher information content as represented by, for example, greater directionality or charge gradient, are intrinsically harder to position correctly. This is why it is so important to get these interactions right during the fragment screening stage (or during initial fragment optimization) while the molecule is still unencumbered with substituents.

Ultimately this approach should facilitate the identification of the most efficient optimized ligand by prioritizing where and how researchers should focus their efforts as they optimize interactions between ligand and target. Molecular interactions with little information content (e.g. lipophilic moieties which have minimal directional requirements) can more easily "slide" into matching low information regions in the binding site. Aromatic groups can also adapt to different environments because of the intrinsic polarizability of such moieties which may help to explain their prevalence in fragment hits. A further benefit of getting the maximum reward from initial polar interactions in a fragment is that this is generally equated with an enthalpic signature to the overall binding free energy. Some evidence suggests that starting with such optimized interactions can help minimize the entropic contributions to the free energy (often equated with an unwelcome increase in lipophilicity) to build the desired level of potency.[4,43]

Ligand efficiency metrics have been widely adopted in medicinal chemistry to ensure a focus on an appropriate balance of physicochemical properties and potency.[7] This focus is particularly important while developing fragments into quality leads for final optimization. Two metrics that are particularly useful in this context are the heavy atom adjusted Lipophilic Ligand Efficiency (LLE_{AT})[44] and Group Efficiency (GE).[45] The metric LLE_{AT} (= 0.111 + 1.37(LLE/HAC)) (where LLE is the Lipophilic Ligand Efficiency (= pK_i – cLogP), sometimes referred to as LipE) links potency with both lipophilicity and size and is scaled so as to be comparable to Ligand Efficiency (LE), whereby an acceptable lower limit for good binding sites is considered to be 0.3. The importance of using size-adjusted lipophilic ligand efficiency is that the advantages gained by initially using a fragment approach can be easily squandered without continued focus on size.6 However, as with all such "rules", the metrics are guides rather than absolutes and due care should be taken when used in a filtering context. GE (equal to the free energy gain for the atoms added) helps focus on the contribution of different parts of a lead molecule to the overall binding affinity. It is similar to Ligand Efficiency but focuses only on the group of atoms that have been added to an initial molecule. A recent example of the use of GE is provided by the Abell group during the fragment based discovery of leads for Pts, a potential target for tuberculosis.[46] Each of these metrics focuses on affinity optimization rather than *in vivo* properties, but they do help control the physicochemical properties which are known to be key contributors to success *in vivo*.[47]

Size and shape considerations

The small size of fragments typically yields hits with low affinity and low specificity. These features can be correlated with the intrinsically low number of interactions fragments are able to form with proteins. Low specificity has two consequences: first, a fragment may be able to bind to various proteins, and second, a fragment may be able to bind to a single protein in several ways. In the first sense low specificity relates to the high hit rate of fragment screens and this is advantageous since specificity can be introduced later in fragment optimization. The ability of a fragment to bind to a protein in several ways can frustrate optimization schemes, which assume a consistent binding pose to establish structure-activity relationships. However, a fragment with multiple binding modes can still be useful in druggability studies. They can be incorporated into a well-established procedure to probe proteins with very small compounds, such as water and organic solvents both with experimental[48] and computational methods.[49] These compounds with few non-hydrogen atoms bind to proteins and form clusters whose positions coincide with those of known inhibitors. Moreover, the extension and relative position of these clusters carries significant information on protein druggability.[50]

A consensus is growing over the ideal size of fragments intended for lead optimization projects. Very small compounds are useful in binding site detection and druggablity studies but are less suited for fragment screening owing to their versatile binding modes to a single protein. According to a *Practical Fragments* poll, over 85% of responders voted for a minimum fragment size between 5 and 10 heavy atoms.[51] Clearly, the minimum size is a compromise between opposing requirements. One is that a better coverage of chemical space is available with smaller ligands as a result of lower molecular complexity; this is a fundamental advantage of FBDD. Other requirements are detectable affinity and specificity, and these favor larger compounds. As noted above, experiences with crystallographic

fragment screens show that the highest hit rates can be achieved with compounds containing 10-14 heavy atoms.[2]

The expansion of a hit to a lead-like compound is facilitated by the existence of a single or at least a prevailing binding mode that is preserved during fragment expansion. There have been numerous analyses of binding mode conservation of fragments either by cutting larger ligands to smaller pieces or the reverse, by analyzing the interactions of larger compounds as they were grown from fragments. Several studies obtained direct structural information of binding modes. Van Molle *et al*. investigated the deconstruction of lead-like inhibitors of the pVHL:HIF-1α interaction and found that the fragments (including the smallest one with 13 heavy atoms) maintained the binding pose observed in larger compounds.[52] Fry *et al*. also investigated the dissection of inhibitors of a protein-protein interaction and found that the smallest fragment with detectable binding contained 20 heavy atoms and exhibited the binding mode observed in the full inhibitor.[53] Even smaller fragments including phenylphosphate with only 11 heavy atoms were found to conserve both the binding site and the H-bonding network in the fragments and in the corresponding full length inhibitors in complexes with the phospho-tyrosine binding site of the SH2 domain of pp60src.[54] Andersen *et al*.[55] deconstructed a cyclopentapeptide chitinase inhibitor, argifin, into linear peptides and dimethylguanylurea. X-ray crystallographic analysis of the dissected compounds complexed with chitinase B1 showed that the conformations of linear peptides were similar to those in argifin and that dimethylguanylurea (9 heavy atoms) exhibited all significant interactions identified in argifin.

In contrast to these results where fragments obtained by ligand dissection preserved their binding mode, Barelier *et al*. found that fragments derived from Bcl-X_L inhibitors do not form

the same interactions as in the full inhibitor; rather they all bind within a single region of the protein site.[56] Babaoglu and Shoichet deconstructed a β-lactamase inhibitor into fragments with 8, 11 and 14 heavy atoms and observed that the fragments did not preserve the binding mode corresponding to the same moiety in the full inhibitor.[57] Satoh *et al*.[58] reported multiple binding modes of a ligand to the Nrf2 interaction site of Keap1 and suggested that, based on crystallographic and computational studies, the preferred binding mode is different in solution and in crystals, and the latter is affected by crystal packing.

In other studies of compound deconstructions, indirect evidence is used to deduce the existence of multiple binding modes. Barelier *et al*. investigated the deconstruction of substrates of six enzymes and observed no binding or significant loss in activity even when fragments included the key reactive groups.[59] Brandt *et al*. studied the binding of compounds obtained by the deconstruction of HIV-1 reverse transcriptase inhibitors bound to an allosteric site not present in the apo form of the enzyme.[60] Few of the fragment-sized compounds showed detectable affinity in surface plasmon resonance experiments and the authors suggest the absence of an efficient fragment binding hot spot at the site examined.

Fragments are often optimized by growing them into larger compounds and less frequently by linking fragments binding in proximal protein sites. These efforts are often based on the assumption of binding mode conservation although this is not always explicitly verified. Sometimes, fragment modifications make the question difficult to answer, as the original fragment hit within the new compound cannot be definitely identified. Edink *et al*. were able to optimize a fragment hit (17 heavy atoms) bound to an inducible subpocket of acetylcholine-binding protein into a larger inhibitor. Binding mode conservation was demonstrated by the X-ray structure of complexes formed by the fragment and by the

optimized compound.[61] Bauman *et al*. optimized a fragment (9 heavy atoms) bound in three different ways to influenza polymerase acidic protein N-terminal endonuclease domain. A fragment chelated to two metal ions was selected and optimized to a larger inhibitor in which the binding mode of the starting fragment was preserved.[62] Chen and Shoichet identified an unselective β-lactamase inhibitor (13 heavy atoms) that was optimized into a selective CTX-M inhibitor. Binding pose conservation was demonstrated by X-ray crystallographic analysis of the complexes formed by the fragment and the optimized compound.[63] Orita *et al*. analyzed 25 fragment optimizations where the complexes of the fragment and the optimized compound are both available. They showed that the average root mean square deviation of a core structure between fragment hit and lead is well below 1 Å and the hydrogen bonding pattern of the fragments is preserved in the optimized compounds.[64]

The cited examples clearly show that fragments typically having 10 to 20 heavy atoms are able to bind specifically to a variety of protein binding sites including polar pockets and protein-protein interfaces. The specific binding is proven by the conservation of the binding mode as the fragments are grown to larger inhibitors. However, there are many cases where the binding mode of a fragment does not agree with that of the corresponding moiety in a larger compound. An explanation of these differences was proposed by Kozakov *et al*.[65] based on the varying overlap of the bound fragments with the primary hot spot. They also propose a simple method based on their computational solvent mapping protocol[49] to identify primary hot spots as sites where the highest number of clusters of various probe compounds bind. A possible interpretation of this finding is that a site that binds various small probes contains a diverse set of proximal binding functions that interact collectively with fragments having appropriate pharmacophore features. This explains what was convincingly shown in several examples that fragments bound with appropriate overlap with the hot spot of consensus

clusters preserve their binding mode,[65] while those bound at the site of other clusters may change their binding mode upon structure expansion. It is also consistent with the fact that although fragments are versatile binders with the ability to bind to various protein targets, they can also be evolved to higher specificity for the more druggable targets. For these latter proteins a primary hot spot is available that is able to bind certain fragments specifically with an affinity not available for the same ligands at other protein sites.

Kozakov *et al*. argue that the intrinsic binding potential of the protein site has a decisive role in the conservation of the binding mode.[65] They suggest that secondary binding sites tend to have less potential to conserve binding modes for fragments and their larger derivatives. Whether this is indeed the case is an important question since targeting secondary binding sites by fragments both in enzymes[66] and in GPCRs[67,68] has been gaining increased attention. In addition to the binding potential of the protein site, structural features of the fragments are also critical for both binding and specificity. An analysis of experimental fragment-protein complexes in the Protein Data Bank[69] revealed that fragments form on average two optimal geometry hydrogen bonds with the protein hot spot.[43] Some typical examples are shown in Figure 6. Binding pharmacophores that have been identified from the analysis of molecular interactions can be useful in the design of fragment libraries. Ideally a variety of fragments that contain each of the binding pharmacophores should be included in the library.[70]

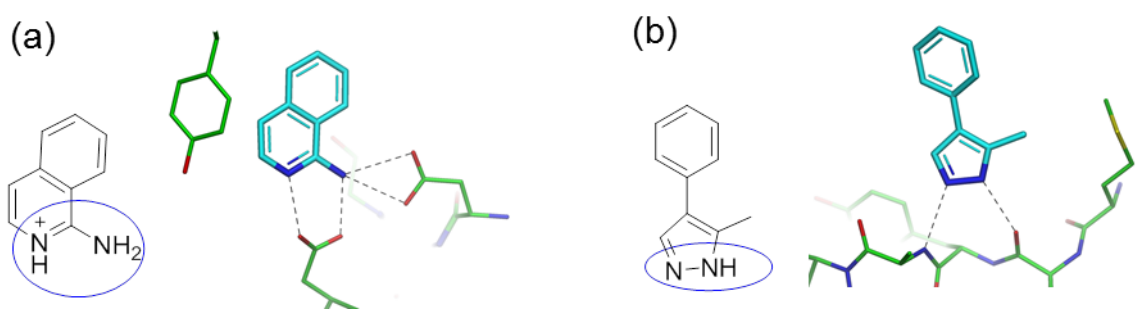


Figure 6. Crystal structures of fragments exhibiting key binding pharamcophores. Chemical structures are also given with the binding pharamcophore marked in blue. (a) amidine-like pharmacophore in beta secretase (PDB Code: 2OHK). (b) A donor acceptor pharmacophore in protein kinase A (PDB Code: 2UW3).

The presence of a few directional polar interactions in a predominantly apolar environment contributes to the strength and specificity of the binding and is in line with the properties of hot spots.[71,72,73,74] It was also found that less polar fragments tend to bind with lower affinity and more promiscuously.[29,56,75]

It seems that the presence of a strong hydrogen bond in the initial fragment indicates that a strong anchor point is likely to be maintained.[64] If the fragment does shift, it likely had multiple orientations in the first place, possibly disguised in poorly defined electron density. Alternatively, a stronger set of interactions may have been made in the optimization processes which overcame initial weak binding; this possibility can be particularly pertinent where synergy between interactions takes place, which is not possible in the smaller molecules. It is therefore important to continue to check ligand position and orientation throughout the fragment evolution process, particularly when unexpected SAR is encountered.

By contrast, potency driven by lipophilicity has been well documented to be a source of non-specific promiscuity and a good fragment hit can easily be squandered by inappropriate use of lipophilicity.6,[41,76,77] This promiscuous binding seems intrinsic to lipophilic interactions and has been interpreted in the molecular complexity model as resulting from the low information content of such interactions.[42]

A further topic relevant to binding specificity is the competitive or simultaneous binding of two or more fragments. This can happen when fragments are pooled together to improve the throughput of a screening technique. In ligand-observed NMR screening it is usually possible to identify multiple hits from a pool of fragments provided that the NMR spectra of the fragments do not overlap. In X-ray screening, it is important that fragments within a pool have shape diversity so that hits can be identified more easily. However, even a clear X-ray hit does not preclude the possibility that another fragment from the pool can bind at the same site. Further deconvolution experiments may therefore be required, and if the hit rate is high, it may be more efficient to screen compounds individually. The synergistic binding of two or more fragments may also interfere with the deconvolution of hits in X-ray screening of fragment cocktails. If the individual fragments do not bind strongly enough for detection they would be considered as false positives and yet the cooperative binding of the two fragments would provide valuable information to identify chemical starting points for developing potent inhibitors.[78] One benefit of having two fragments bound to proximal sites is the possibility of using fragment linking, [79] a theoretically attractive concept which unfortunately is not often successfully achieved. The challenge of fragment linking may be that at least one fragment will likely bind outside the principal hot spot, making the initial pose hard to maintain when this fragment is expanded.[65] Nevertheless, there are successful applications of fragment linking with significant affinity increases.[80,81] In these cited examples the significant affinity gain is achieved by linking fragments having not more than 15 heavy atoms. Moreover, the linking coefficient, defined as the difference between the affinity of the final ligand and the sum of fragment affinities is near to the estimated rigid body entropy loss of binding[82] suggesting that fragments can sometimes be linked without introducing strain or altering poses.

The appropriate shape of fragments garners much debate. Fragment libraries have traditionally been dominated by compounds with somewhat planar rings (such as pyrimidines), although some evidence suggests that non-aromatic compounds lead to improved quality of final clinical candidates.[83,84] However, this evidence has been challenged,[85] and it is fair to say that calculating physicochemical properties is no substitution for experimental measurements.

Nonetheless, the strategy of boosting "three dimensionality" of ligands is much discussed as a way to reduce promiscuity and improve solubility and thus ultimately improve the developability of ligands. This in turn has resulted in moves to measure and increase the 3D character of fragments.[86,87] However, this increased 3D character could result in an increase in the complexity of the fragments and thus lead to a reduction in the probability of any one of them matching the receptor. This can be partially abrogated by increasing the number of fragments screened, but this brings other problems in terms of the capacity of biophysical screening methods. Additional impetus to increase the presence of 3D fragments in libraries is the proposal that this leads to a broader coverage of biologically relevant chemical space and may be especially beneficial for challenging targets.[86] Nevertheless, initial studies suggest that hit rates for 3D fragments are lower than those for flatter compounds and this is consistent with the higher complexity of 3D fragments.[2] These considerations suggest that taking maximum advantage of FBDD hinges on a sensitive balance between complexity, size and diversity of fragments.

Target related aspects

It is a truth universally acknowledged that not all proteins can be targeted with small molecules. Those that can are referred to as "ligandable."[8] This term is preferable to the term "druggable," since a target may be ligandable even though it ultimately has no role in a disease. Indeed, since validating a therapeutic target often first requires identifying a small molecule modulator, researchers usually need to establish whether a target is ligandable before establishing whether it is ultimately druggable. Even within target classes that have been shown to be ligandable there is considerable variability in the molecular weight, polarity and lipophilcity of typical ligands with drug-like potencies.[88] Many attempts have been made to use computational methods to assess ligandability[89,90] but fragment screening has also proven to be an effective experimental means to assess ligandability. Multiple studies have shown that the hit rate from a fragment screen correlates well with the hit rate from high-throughput screening (HTS) and the ultimate success of lead optimization.[75,89,91]

Some targets, such as many enzymes, have a well-defined pocket that has evolved to bind small molecules, so it is not surprising that many fragments bind in these substrate binding sites, often making some of the same interactions as the natural substrates. For example, screens against kinases often yield fragments that bind in the adenosine binding site, and these often have high ligand efficiency values. On the other hand, protein-protein interfaces often consist of large, flat surfaces with few or no pockets, and fragment screens against such targets are usually less successful. Consequently, ligand efficiency values of ligands that disrupt protein-protein ligands tend to be lower (see below).[92]

Historically most FBDD programs have targeted enzymes or other soluble proteins, and fragment screening is much more challenging on, for example, membrane proteins. However, significant progress has been made in this latter area through the use of thermally stabilized

membrane proteins which are selected by effective use of mutagenesis studies.[93] After stabilization the protein is more amenable to standard assay techniques such as SPR, Thermal Shift and X-ray based FBDD experiments.

One point of debate in fragment library design is whether fragment libraries should be target-directed or general. For example, since kinases have evolved to bind ATP, it might make sense to design a kinase-directed library consisting of moieties that bind to the region responsible for binding the adenine part of ATP.[94] This approach has indeed been quite successful, as noted in Table 2, and many companies and commercial vendors have built custom libraries for specific target classes. A potential drawback is that the resulting fragments are less likely to be novel. However, this should not necessarily be a major consideration as much of the intellectual property is created during fragment to lead optimization and beyond. For example, both vemurafenib[95] and the clinical candidate 4-amino-N-[(1*S*)-1-(4-chlorophenyl)-3-hydroxypropyl]-1-(7H-pyrrolo[2,3-d]pyrimidin-4-yl)piperidine-4-carboxamide (AZD5363) [96] had their origins in a common 7-azaindole fragment, yet ultimately their structures and properties dramatically diverged. Indeed, the azaindole scaffold is such a fruitful source of fragments that an entire review was devoted to it.[97] A more serious objection to target-based libraries is that sometimes the most interesting hits are the least expected; it is hard to design for serendipity. For example, a screen against PAK1 identified a fragment with scant resemblance to adenine. This binds outside the ATP-binding site and led to allosteric molecules that are remarkably selective for the kinase.[98]

A related issue is selectivity: how selective should a fragment hit be? Researchers accustomed to working with HTS hits may instinctively prefer more specific fragment hits, but as the azaindole discussion above demonstrates this is not necessarily justified. Indeed, on the basis

of the molecular complexity argument, fragments should theoretically not be very selective or it would be impossible to get hits from a small library. That said, it is not uncommon to find fragments that show specificity. Bamborough and colleagues at GlaxoSmithKline demonstrated this experimentally by screening nearly 1000 fragments against 30 kinases.[26] These fragments were chosen to be likely to bind to the ATP hinge region of kinases, yet many of them proved to be quite selective. Even adenine, which should after all bind to all kinases, strongly inhibited less than half of them. However, when the researchers looked at larger molecules that contained these fragments, there was no correlation between the selectivity of the fragments and that of the more potent, elaborated molecules. Recent work from the Collins group showed similar results from fragment-to-lead programs, and also demonstrated that very small changes to a molecule could have drastic effects on selectivity.[99] In other words, it is possible to start with a non-selective fragment (such as 7-azaindole) and develop a selective inhibitor (such as vemurafenib). On the other hand, it is also all-too-possible to start with a selective fragment and end up with a non-selective inhibitor, especially by adding inappropriate levels of lipophilicity.

As hinted at above, one effective way to address selectivity among related targets is to sidestep the common active site entirely and focus on allosteric sites. Due to their ability to bind to small, sometimes cryptic binding sites, fragments are ideally suited for identifying allosteric sites. In fact, a crystallographic fragment screen from the Arnold group at Rutgers University revealed 16 different binding sites on HIV Reverse Transcriptase, though the function of many of these is uncertain.[100] Allosteric kinase inhibitors are well-precedented and often quite selective;[101] as noted above an allosteric PAK1 inhibitor derived from a fragment screen at Novartis appears to be >50-fold selective for the enzyme compared to 441 other kinases.[98] Novartis has also started clinical development on an allosteric ABL inhibitor.[102]

Allosteric modulators can also be effective against difficult targets. For example, caspases are a family of cysteine proteases with charged, extended binding sites, making it difficult to discover specific drug-like leads. A collaboration between Genentech and UCSF revealed fragments that bind to an allosteric site on procaspase-6 and stabilize this inactive zymogen form.[103] Another example of using allosteric modulators to prevent enzyme activation was published by researchers from Astex, who performed a fragment screen against the hepatitis C protein HCV NS3 and found fragments that bind between the protease and helicase domains. These were ultimately optimized to low nanomolar potency with cell-based activity.[104]

As mentioned above, protein-protein interactions (PPIs) tend to be more difficult than conventional targets, as reflected in lower hit rates from both high-throughput and fragment screens. Nonetheless, fragment screens have been successful in multiple cases.[105] One of the most prominent examples is work done by researchers at AbbVie in which they used SAR by NMR to identify fragment hits against the anti-cancer target BCL-x_L. After extensive medicinal chemistry, these were ultimately optimized to navitoclax.[106] Further modification of this scaffold led to venetoclax, [107] which is selective for the related protein BCL-2 and is reportedly on track for FDA approval in 2016. Fragment screening has also been successfully applied to discover inhibitors against another member of this protein family, MCL-1[108,109] as well as against RAD51.[110]

Due to the difficult nature of PPIs, higher concentrations of fragments are often necessary to find hits with even low ligand efficiency. Carefully done, however, this can be highly effective, as demonstrated by a recent study from researchers at Astex. An NMR screen of fragments at 10 mM revealed a very weak hit against the anti-apoptotic proteins cIAP1 and

XIAP. Unlike most previously reported molecules, this had comparable activity against both proteins, and structure-guided medicinal chemistry was successful at generating nanomolar inhibitors with cell-based activity.[38]

Another difficult target that has recently succumbed to fragments is Ras. Researchers at Genentech and Vanderbilt University independently used NMR screening to identify fragments that bind to a small surface-exposed pocket and block interactions with the exchange factor SOS.[111,112]

For particularly difficult targets, covalent bond formation can be effective at identifying low-affinity fragments; the bond can stabilize interactions that may be too weak to detect using other approaches. This was the idea behind Tethering, which relies on thermodynamically-driven disulfide exchange between a cysteine residue in the protein and libraries of disulfide-containing fragments.[113,114] The challenge with disulfide bonds is that they need to be replaced to be effective in cells, let alone animals. However, researchers at UCSF demonstrated that this replacement could be accomplished against a mutant form of K-Ras containing an activating cysteine residue.[115]

Covalent drugs have undergone a renaissance of sorts, and this has propagated through to libraries of designed covalent fragments. The Statsyuk group at Northwestern University has designed libraries of acrylate-containing fragments and shown that selective inhibitors can be found against different proteins.[116] Unlike in the case of Tethering, these fragments form irreversible bonds to the protein, so it is important that the intrinsic reactivities of the fragments do not vary too much. Somewhat sidestepping this issue, the Taunton group at UCSF has generated cyanoacrylamide-containing fragments, which can form reversible

covalent bonds with cysteine residues. These were used to develop potent, selective kinase inhibitors.[117] Computational methods can be successfully applied to identify covalent inhibitors, and a program developed by Shoichet and colleagues at UCSF is freely available.[118]

Synthesis related aspects

When a fragment screening library is used to provide hits to start a drug discovery program it is important that the hits obtained are "optimizable". That is, there is a high likelihood that when a fragment binds to the target protein there will be accessible positions available to further exploit the hit chemically. This concept of optimizability has been built into fragment screening sets in a number of different ways. One approach is to ensure that a fragment has a defined synthetic handle.[16] One challenge here is that polar moieties such as acids or amines can often provide the key interaction in a lower molecular weight fragment and any synthetic modification could change the nature of the group (e.g. by amide formation).[119] Schuffenhauer *et al*.[27] overcame this problem by developing a reaction dictionary that allowed the searching of fragments that mask the reactive functionality. Lau et. al.[28] used experienced medicinal chemists to judge the chemical expandability of fragments in the development of a fragment screening library at Pfizer. It is also possible to provide an *in silico* assessment of "optimizability"[37] to guide in the triaging of thousands of potential fragments. A RECAP like[120] approach was used to identify the "core" for each fragment by trimming back to potential reactive groups and a substructure search in internal and external data sources was able to identify the number of substitution positions on the fragment, allowing ranking of fragments. Cox *et al.* have described the design of a poised fragment library where fragments

derived from simple synthetic reactions are included in the library so as to facilitate the rapid synthesis of analogs of fragment hits.[121]

Historically, fragment libraries have used substructures derived from drugs as one of the sources of ideas.[122] Pihan *et al.*[123] have provided a set of commercially available fragments based on substructures of drugs. However, Morley *et al.*[86] have analyzed a number of fragment libraries using principal moments of inertia (PMI) plots and suggest that they have limited shape diversity when compared to fragments derived from compounds tested in humans. They describe the 3D Fragment Consortium which has a goal to provide additional compounds with greater three dimensionality. Hung *et al.* have exploited diversity-orientated synthesis as a method for generating 3D fragments.[124] Natural products have been a rich source of drugs and tend to have a much higher sp^3 content than synthetic compounds.[125,126] Over *et al.*[127] describe an analysis of natural products to determine a set of natural product-derived fragments rich in sp^3 carbon atoms. The library was validated by screening against p38 kinase and identifying an atypical non-aromatic kinase fragment binding in an allosteric site. Vu *et al.*[128] describe the creation of a natural product library of fragment sized compounds, identifying compounds binding to *Plasmodium falciparum* 2'-deoxyuridine 5'-triphosphate nucleotidohydrolase (PfdUTPase). However, care must also be taken when designing natural product fragment libraries, and many of the other principles described in this Perspective should still be applied, of particular concern being the increase in molecular complexity and synthetic tractability associated with increasing the sp^3 content. Furthermore, some natural products contain many groups that are considered PAINS[21] and in isolation would be very problematic in terms of selectivity and mode of action, for example quinones and catechols.

Many commercial vendors offer fragment libraries derived from their own synthetic compounds, marketed drugs and/or natural products[129] and a number of CROs also claim unique fragment libraries as part of their offering, some of which are highlighted in Table 5. A brief summary of the design principles behind each set is also given where available. Greater detail can be found in the references and links in the table. Most of these libraries are designed within the bounds of the Rule-of-Three9 and many vendors now have tens of thousands of fragment like compounds available. Fluorine-containing libraries are available for ^{19}F NMR and bromine-containing libraries for exploring halogen bonding effects[130] and facilitating crystallographic screening. There are also a number of specialized libraries for specific target classes and covalent binding. A full assessment of these offerings is beyond the scope of this review, and indeed comparison is difficult without knowing the purpose or target of the campaign. However, the guidelines given throughout this review should prove useful in guiding the reader to select fragments from commercial sources. These include querying the available measured solubility and purity, applying substructure filters and selection algorithms appropriate to the target(s), and assessing the compounds for undesirable features and "optimizability" either by visual inspection or algorithmically. Chris Swain has analyzed the physicochemical properties of some of these libraries,[131] though it is important to note that vendors do periodically change the compositions of their libraries.

Despite the availability of fragment sets from commercial sources, there is a need for continuing development of novel fragments to tackle new targets. The impact of synthetic methodology on the properties of the resulting molecules has been recognized in the context of lead discovery,[132] emphasising the role of the synthetic chemist in addressing these issues. In a recent essay,[70] Murray and Rees have called on the chemistry community to increase research in the development of novel synthetic chemistry methodology, particularly with

regard to fragment-sized compounds . In particular they highlight the need to more thoroughly explore and describe the synthetic routes to all potential growth vectors prior to screening because much time can be spent in the initial phases of optimization developing appropriate chemistry. The further development of approaches to C-H functionalization[133] and use of high-pressure continuous flow methods to give novel heterocycles[134] are some examples of recent chemistry developments that can be exploited in the synthesis of novel fragments.

Vendor	**Website**	**Library**	**# cmpds**	**Commercially available**
3D Fragment Library Consortium	Ref. 86	Increased shape diversity through greater three dimensionality	>500[135]	N
ACB Blocks	www.acbblocks.com	^{19}F NMR-oriented, RO3 compliant, predicted to be soluble. Purity > 96%	1280	Y
Analyticon	www.ac-discovery.com	FRGx: Fragments from Nature. RO3 compliant, high solubility, purity > 95%.	213	Y

AnCoreX	www.ancorex.com	MetaKel™ (metal chelating. MW < 300) TCI-Frag™ (Targeted covalent inhibitor fragment screening. Mildly reactive functionalities, RO3 compliant)	>500 >100	N
Asinex	www.asinex.com	Fragment library	22,524	Y
Beactica	www.beactica.com	SPRINT. Validated for SPR. 2000 purchased fragments.	1946	N
Beryllium	www.be4.com	Fragments of Life™.[138]	1,500	N
BioBlocks	www.bioblocks.com	Proprietary fragment library based on exploration of 3D shapes.	~1000	N

Charles River	www.criver.com	Core fragment library, 500 member kinase focussed set and 500 member ^{19}F labelled.	1,500	N
ChemBridge	www.chembridge.com	ChemBridge Fragment Library, RO3 compliant with predicted solubility. Minimum purity 90% by ^{1}H NMR.	>7,000	Y
ChemDiv	www.chemdiv.com	3D designed fragment library	4,283	Y
Enamine	www.enamine.com	RO3 compliant Golden Fragment Library (diverse subset of full library), "simple" fragment	18,108 1,794 126,597	Y

		library: RO3 compliant <=20 heavy atoms from screening collection. Other smaller sets of fluorinated, brominated (for X-ray), covalent sp^3 rich and PPI fragments.		
InFarmatik	www.infarmatik.com	Consolidated library from different concepts, In3D,GPCR, Kinase	1700	Y
IOTA	www.iotapharma.com	Diverse, mainly RO3 compliant.	1,500	Y
Integrex	www.integrexresearch.com	Diversity in shape and chemical structure, RO3 allowing one violation.	1,500	Y

Key Organics	www.keyorganics.net	Fragment library 2nd Generation with assured aqueous solubility, RO3 compliant. Fragments from Nature, RO3 compliant, assured solubility and high Fsp^3 content. CNS fragment library. More stringent filters (e.g. mw<240) Fluorine Bromine	26,000 1166 183 700 1,950 1,656	Y
Life Chemicals	www.lifechemicals.com	General RO3 compliant(and subsets of predicted soluble, Fluorinated,	31,000 14,000	Y

		Brominated and Fsp^3 enriched, Covalent and PPI focused)		
Maybridge	www.maybridge.com	RO3 compliant Diversity Fragment library with assured solubility in DMSO and PBS buffer. 1,000 fragment subset available. Fragment collection, filtered by purity, mw<350 and substructures.	2,500 >30,000	Y
Otava	www.otavachemicals.com	General RO3 compliant, predicted to be soluble. Assured	12,486 1,000	Y

		solubility in DMSO and PBS Fluorine Metal chelator Halogen-enriched with Bromine for X-ray studies.	 1,217 1,023 618	
Prestwick Chemical	www.prestwickchemical.com	Prestwick Fragment Library mainly derived from drug fragments, Ro3 compliant.	910	Y
Selcia	www.selcia.com	Selcia Fragment Library, RO3 compliant, predicted and measured solubility with purity > 95%.	1,400	Y
TimTec	www.timtec.net	Fragment-based library structurally diverse with	3,200	Y

		predicted high solubility.		
Vitas-M	www.vitasmlab.com	RO3 compliant.	18,932	Y
Zenobia	www.zenobiafragments.com	Fragment library from different design paradigms, cores from drugs, higher Fsp^3 flexible cores etc.	968	Y
ZoBio	http://www.zobio.com/	RO3 compliant diverse commercially available and a smaller proprietary library.	2,300	N

Table 5. Selected list of vendors and CRO fragment libraries, highlighting the range of libraries on offer in terms of design principles and numbers.

Management of fragment libraries

Management of fragment libraries is prosaic yet critical. The first step is to make sure that every fragment that enters the library passes acceptable purity criteria, typically 90-95%. At a minimum this should be established by HPLC-MS, but preferably this will be confirmed by

NMR. If NMR will be used as part of the screening process, it will be important to obtain spectra of the compounds anyway, so purity can be established with minimal extra effort. Recognizing the importance of this, some commercial vendors have started including NMR spectra for their fragment libraries. However, this should not substitute for in-house confirmation: an analysis by Darren Begley, then at Emerald Biosciences, found that 16% of samples across a set of >10 vendors failed QC, with failure rates as high as 33% in some cases.[136]

In addition to the working library, it is essential to have access to dry stocks of fragments for confirmation experiments. When Pfizer built its fragment library, the requirement was that at least 200 mg of material was available for in-house compounds.[28] When purchasing compounds from a commercial vendor, it is prudent to check availability. It is all too common (and frustrating) to find that a fragment is no longer available months or years later when it shows up as a screening hit.

DMSO is almost always the solvent of choice for fragment screening libraries. If NMR will be used at some point, it makes sense to dissolve the fragments in deuterated DMSO, particularly since the added cost is minimal. Compounds are typically dissolved at a concentration of at least 50 mM to minimize the amount of solvent introduced into the assay: a final concentration of 1 mM fragment would give 2% DMSO. Even higher concentration stocks are advantageous in reducing the amount of organic solvent in the final assay. For example, researchers at Monash University prepared their fragment libraries at 200 mM.[137] However, this comes with the risk that fewer fragments will be soluble.

Besides DMSO, the only other solvent that has been used to any extent is methanol; due to its volatility it can be used to add larger amounts of fragments to a crystallization plate without changing the final buffer composition.[138] However, this approach does not appear to have been used widely. The low surface tension of methanol makes it harder to precisely aliquot, and fragments may not easily redissolve in buffer once they have dried onto a plate.

Compound stability is also not to be taken for granted. Most fragment libraries are prepared in DMSO, which is slightly oxidizing, and some compounds can rapidly degrade in it.[139] Even water is not necessarily benign: seemingly stable moieties such as benzoxazoles can hydrolyze, even when stored as solids.[140] In another case a library compound cyclized in DMSO to form an active species, but this in turn proved to be unstable in water, complicating assay interpretation.[13] Thus, best practice is to periodically (often annually) reevaluate fragment libraries to ensure their integrity.[141]

There is some debate as to how best to store fragment stock solutions. A study of 7200 compounds stored in DMSO at room temperature at Procter and Gamble found that while only 8% had degraded after 3 months, this had increased to 17% after 6 months and a whopping 48% after 1 year.[142] Storing DMSO solutions frozen would seem to be a solution, but the resulting freeze-thaw cycles are themselves problematic, particularly as they increase the likelihood of introducing atmospheric water into the DMSO, which is highly hygroscopic. Compound solubility in DMSO has been reported to decrease dramatically in the presence of even small amounts of water.[143] Consistent with this, a study of 232 compounds subjected to repeated freeze-thaw cycles showed precipitation and decreased concentrations compared to compounds that had been stored frozen or even at room temperature.[142] It is important to note

that the molecules used in these studies were not fragments, and it is possible that larger molecules have higher odds of instability and precipitation.

Although pure DMSO freezes at around +19 ˚C, compound stocks at 100 mM or more often remain liquid even at low temperatures. Thus, a common strategy is to store fragment stocks at low temperatures to slow decomposition but minimize freeze-thaw cycles. In a poll taken at *Practical Fragments* in early 2014, more than half of respondents reported storing fragment libraries at +4 ˚C or – 20 ˚C, while nearly a third kept their libraries at room temperature (Figure 7).[144] Keeping libraries under inert gas in low humidity is also good practice.

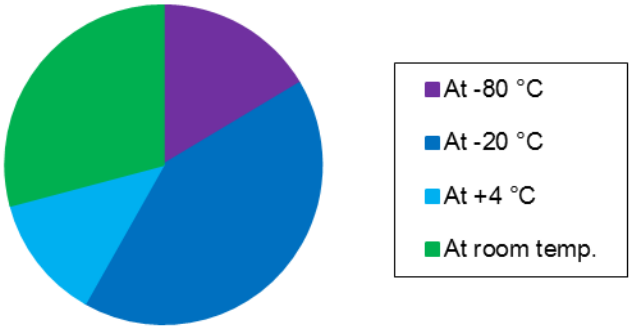


Figure 7. Results of a poll taken on Practical Fragments in January and February 2014 asking how fragment libraries are stored; there were 79 responses.[144]

Perspectives, conclusion

Fragment-based drug discovery is well-established in industry and has resulted in more than 30 drugs entering clinical trials,[145] with one – vemurafenib – already approved. FBDD also

has key attractions for academia. Notably, it is able to tackle difficult or novel targets for which no chemical matter may be found in existing HTS collections. Moreover, establishing and screening a fragment library, which can be as small as a few hundred compounds, is far less expensive and challenging than working with a standard HTS collection of 500,000+ compounds.

However, precisely because fragment libraries are small, the outcome of FBDD programs is very much dependent on the initial fragment hits, so it is essential that these are carefully selected. Design of fragment libraries requires the consideration of multiple factors that often contradict each other, complicating the process.

At a minimum, fragment libraries should be purged of compounds known to be generically reactive or that interfere with common assays, including PAINS such as quinones and redox cyclers such as toxoflavin. Certain types of compounds that repeatedly show up as screening hits but rarely yield structural information, such as PrATs, should probably also be avoided. On the other hand, "non-selectively specific fragments" that bind to multiple proteins through well-characterized interactions, such as 7-azaindole and 4-bromopyrazole, may be useful additions to a fragment library.

Simple filters based exclusively on physicochemical properties have a role – there is no place for a massively lipophilic 500 Da molecule in a fragment library – but as with all of drug discovery these should be seen as guidelines rather than hard and fast rules. On the other hand, high solubility is critical, and this should be measured experimentally.

Other factors are harder to prescribe. Molecular complexity impacts the affinity and specificity against targets, but is more difficult to measure. In general smaller, simpler fragments are likely to be more useful, while fragments with multiple stereocenters or more three-dimensional shapes may provide lower hit rates.[37]

Synthetic considerations are important too: it can be frustrating to discover a fragment binding perfectly in a target only to find that growing in the desired direction requires developing unprecedented chemistry. A good practice is to ensure that multiple analogs of any fragment included in the library are available, either commercially or in-house. Given the limited diversity of available protein binding sites, together with the small relative size of fragments, the size and diversity of the library – assuming a reasonable minimum – are probably less important.

The overall quality of the fragment library depends on physical factors too, of which purity, stability and storage conditions are clearly the most important. Periodic assessment of the fragment library for degradation, and confirmation of any hits with fresh sample, are essential practices.

The success or failure of discovery programs is very much influenced by the screening hits. Since the quality of hits is determined by the screening technology and the library screened, updated design principles of fragment libraries will hopefully contribute to the increased success rate of FBDD programs. Table 6 contains guidelines we would recommend when constructing a fragment library or when considering the addition of new fragments to an existing library. As discussed throughout, these guidelines should be considered in the context of both the biological target and the fragment screening technique.

Building on more than a decade of screening fragment libraries, we hope that this paper will promote wider adoption of fragment-based approaches by helping to improve the quality of fragment libraries and associated screening technologies.

Property	**Guideline**
Library size	500 – 3000 fragments (smaller libraries are more appropriate for lower throughput screening methods (e.g., x-ray) and larger libraries are more appropriate for higher throughput methods (e.g, SPR))
Physico-chemical properties	Molecular weight: ~140-230; Non-hydrogen atoms: 9-16; (Fragments at the lower end of the size range are more appropriate for lower throughput, higher sensitivity screening methods) Lipophilicity (cLogP): ~0.0 to 2.0; Number of freely rotatable bonds: 0-3; Number of chiral centres: 0-1, sometimes 2. Always use racemates. Properties commensurate with biophysical screening at high concentrations, e.g., aqueous solubility (preferably ≥5mM in 5% DMSO, or other screening co-solvents); stability (>24h in solution); avoid compounds/functional groups known to be associated

	with high reactivity, aggregation in solution, or false positives.
Molecular recognition	Diverse, usually polar groups for binding to a protein (a single pharmacophore). An aspiration is to express any given binding pharmacophore in a variety of diverse scaffolds (chemotypes).
Shape	Variety of 2- and 3-dimensional shapes for each scaffold and pharmacophore;
Availability of analogs	Fragments should contain multiple synthetically accessible vectors for fragment growth in 3 dimensions to access new binding interactions. Fragments should be synthesizable in <4 steps from commercial available starting materials. Under some circumstances it may be desirable to choose fragments that are commercially available and/or have many commercially available analogs
Diversity	Diversity metrics should be used to prevent the inclusion of close analogs in the library. Care should be taken to ensure the diversity metrics do not favor the selection of larger fragments over simpler examples and that appropriate descriptors and similarity cutoffs are used.

Specialized considerations	Depending on techniques to be used in screening, there may be special considerations in library design. For example aromatic C-H bonds or ^{19}F may assist NMR screening.

Table 6. Recommended guidelines for construction of a fragment library.

Funding Sources:

GMK and GGF were supported by the National Brain Research Program (KTIA_NAP_13-1-2013-0001). Michael Hann and Stephen Pickett are employees of GlaxoSmithKline. Chris Murray is an employee of Astex Pharmaceuticals. Daniel Erlanson is a co-founder and employee of Carmot Therapeutics.

Acknowledgement

The authors would like to thank the numerous colleagues and members of the FBDD community who provided data and observations for this Perspective, particularly to David Rees and Glyn Williams for their helpful comments. GMK and GGF are grateful for the support of the National Brain Research Program (KTIA_NAP_13-1-2013-0001).

Corresponding Author Information:

Prof. György Miklós Keserű

Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok krt. 2, H-1117 Budapest, Hungary

Phone: +36-1-3826821, E-mail: gy.keseru@ttk.mta.hu

Biographies:

György M Keserű

György obtained his Ph.D. at Budapest, Hungary and joined Sanofi-Aventis CHINOIN heading a chemistry research lab. He moved to Gedeon Richter in 1999 as the Head of Computer-aided Drug Discovery. Since 2007 he was appointed as the Head of Discovery Chemistry at Gedeon Richter and contributed to the discovery of the antipsychotic Vraylar® (cariprazine) that has been approved by the US FDA in 2015. From 2013 he served as a director general of the Research Center for Natural Sciences (RCNS) at the Hungarian Academy of Sciences. Now he is heading the Medicinal Chemistry Research Group at RCNS. His research interests include medicinal chemistry and drug design. He has published over 180 papers and more than 15 books and book chapters.

Daniel A. Erlanson

Daniel A. Erlanson is a co-founder of Carmot Therapeutics, Inc. Prior to Carmot, Dr. Erlanson spent a decade developing technologies and leading medicinal chemistry efforts in oncology and in metabolic and inflammatory diseases at Sunesis Pharmaceuticals, which he joined at the company's inception. Before Sunesis, he was an NIH postdoctoral fellow with James A. Wells at Genentech. Dr. Erlanson earned his PhD in chemistry from Harvard University in the laboratory of Gregory L. Verdine and his BA in chemistry from Carleton College. As well as co-editing two books on fragment-based drug discovery, Dr. Erlanson is an inventor or author on more than forty patents and scientific publications. He is also editor of a blog devoted to fragment-based drug discovery, Practical Fragments (http://practicalfragments.blogspot.com/).

György G. Ferenczy

György G. Ferenczy received his PhD in computational chemistry from the Eötvös University of Budapest. Following postdoctoral research at the University of Oxford, UK and at the University of Nancy, France, Dr Ferenczy worked as a computational chemist and as a group leader first at Gedeon Richter (Budapest) and later at Sanofi (Budapest and Strasbourg). From 2012 he is a senior research fellow at the Semmelweis University and from 2013 at the RCNS of the Hungarian Academy of Sciences. His research interest includes the development and application of computational tools for extended biochemical systems and studying molecular interactions relevant to drug discovery.

Michael M. Hann

After completing his PhD in organic chemistry in 1980, Mike has worked in Pharma R&D initially as a medicinal chemist and then as a computational chemist. He joined Glaxo in 1986 and has been responsible for initially building and leading the computational chemistry department, and more recently leading the biophysics and protein crystallography activities including fragments theory and practice in lead identification. His current role is in looking at new ways to enhance discovery approaches, reducing attrition and promoting scientific excellence across the GSK R&D sites. A particular current interest is in understanding drug distribution at cellular and subcellular resolution. Mike is also an Adjunct Professor in the chemistry department at Imperial College.

Christopher W. Murray

Chris Murray received his BA (1986) in chemistry and PhD in theoretical chemistry (1989) from the University of Cambridge. Currently he is SVP of Discovery Technology at Astex where he has contributed to the design and exploitation of fragment libraries. He has

extensive experience of structure-based drug design and has helped discover a number of compounds that have demonstrated efficacy in human clinical trials.

Stephen D. Pickett

Stephen Pickett studied for a degree in Chemistry at Keble College, Oxford and his PhD , under Professor Sir John Meurig Thomas at the Royal Institution of Great Britain. Following postdoctoral studies at the Imperial Cancer Research Fund with Professor Michael Sternberg, he worked in Pharma R&D at Rhone-Poulenc Rorer and Roche before taking up his current position as a Group Leader in Computational Chemistry and Informatics at GlaxoSmithKline. Dr Pickett's current interests include the computational chemistry and informatics aspects of fragment based drug discovery, high throughput screening, screening collection design, QSAR and virtual screening. He is the author of over 50 peer-reviewed scientific articles and reviews and is a named contributor on six patent applications.

Abbreviations Used:

AC – affinity chromatography

BLI – biolayer interferometry

CE – capillary electrophoresis

DLS – dynamic light scattering

FN – false negative

FP – false positive

FI – fluorescence intensity

FLT – fluorescence life time

FBDD – Fragment-based drug discovery

GE – Group Efficiency

HAC – Heavy Atom Count

HCS – high concentration screening

HTS – high throughput screening

ITC – isothermal titration calorimetry

LLE – Lipophillic Ligand Efficiency

MS – mass spectrometry

MSA – mobility shift assay

MST - microscale thermophoresis

PAINS – Pan Assay Interference Compounds

PfdUTPase – *Plasmodium falciparum* 2'-deoxyuridine 5'-triphosphate nucleotidohydrolase

PrATs – promiscuous 2-aminothiazoles

PPIs – protein-protein interactions

QC – quality control

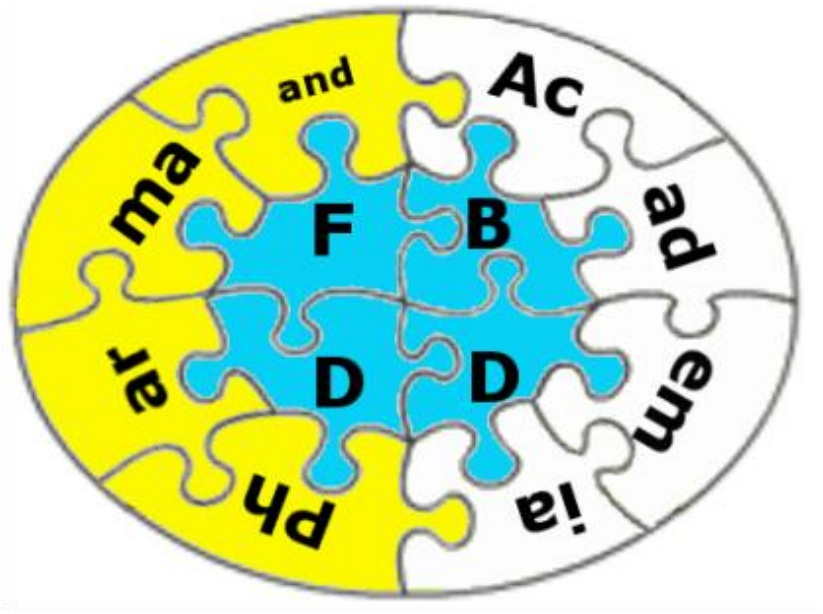
RO3 – Rule of Three

SBDD – Structure-based drug design

SPR – surface plasmon resonance based biosensor technology

TSA – thermal shift assay

Table of Contents graphic



References

1Hann, M.M.; Oprea, T.L. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255-263.

2 Hall, R.J.; Mortenson, P.N.; Murray, C.W. Efficient exploration of chemical space by fragment-based screening. *Progr. Biophys & Mol. Biol.*, **2014**, *116*, 82-91.

3 Murray, C.W., Rees, D.C. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1*, 187-192.

4 Ferenczy, G.G.; Keserű, G.M. On the enthalpic preference of fragment binding. *Med. Chem. Commun.*, **2016**, *7*, 332-337.

5 Ferenczy G.G.; Keserű. G.M. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. *J. Med. Chem.* **2013**, *56*, 2478-2486.

6 Hann, M.M.; Keserű, G.M. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Disc.* **2012**, *11*, 355-365.

7 Hopkins. A.L.; Keserű, G.M.; Leeson, P.D.; Rees, D.C.; Reynolds, C.H., The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **2014**, *13*, 105-121.

8 Edfeldt, F.N., Folmer, R.H.; Breeze, A.L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discov. Today* **2011**, *16*, 284-287.

9 Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A"Rule of Three" for fragment-based lead discovery? *Drug Discov. Today* **2003**, *8*, 867-868.

10 Jhoti, H.; Williams, G.; Rees, D.C.; Murray, C.W. The 'rule of three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.* **2013**, *12*, 644-645.

11 Hann, M.M.; Leach, A.R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856-864.

12 Palmer, N.; Peakman, T,M,; Norton, D.; Rees, D.C. Design and synthesis of dihydroisoquinolones for fragment-based drug discovery (FBDD). *Org. Biomol. Chem.*, **2016**,*14*, 1599-1610.

13 Davis, B.J.; Erlanson, D.A., Learning from our mistakes: the 'unknown knowns' in fragment screening. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2844-2852.

14 Practical Fragments Blog. http://practicalfragments.blogspot.hu/2014/01/poll-results-affiliation-fragment.html (accessed Jan 20, 2016).

15 Optibrium Home Page. http://www.optibrium.com/downloads/Swain_Fragments.pdf (accessed Jan 20, 2016).

16 Siegal, G.; Eiso A.B.; Schultz, J. Integration of fragment screening and library design, *Drug Discov. Today*, **2007**, *12*, 1032-1039.

17 Devine, S.M. Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold. *J. Med. Chem.* **2015**, *58*, 1205–1214.

18 Bauman, J.D.; Harrison, J.J.E.K.; Arnold, E. Rapid experimental SAD phasing and hot-spot identification with halogenated fragments. *IUCr J.* **2016**, *3*, 51-60.

19 Tsai, J.; Lee, J.T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N.K.; Sproesser, K.; Li, L.; Smalley, K.S.; Fong, D.; Zhu, Y.L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S.H.; Schlessinger, J.; Zhang, K.Y.; West, B.L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P.N.; Hirth, P.; Artis, D.R.; Herlyn, M.; Bollag, G. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Nat. Acad. Sci.* **2008**, *105*, 3041–3046.

20 Shipe, W.D.; Sharik, S.S.; Barrow, J.C.; McGaughey, G.B.; Theberge, C.R.; Uslaner, J.M.; Yan, Y.; Renger, J.J.; Smith, S.M.; Coleman, P.J.; Cox, C.D. Discovery and optimization of a

series of pyrimidine-based phosphodiesterase 10A (PDE10A) inhibitors through fragment screening, structure-based design, and parallel synthesis. *J. Med. Chem.* **2015**, *58*, 7888–7894.

21 Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, *53*, 2719– 2740.

22 Erlanson, D.A. Learning from PAINful lessons. *J. Med. Chem.* **2015**, *58*, 2088–2090.

23 Irwin, J.J.; Duan, D.; Torosyan, H.; Doak, A.K.; Ziebart, K.T.; Sterling, T.; Tumanian, G.; Shoichet, B.K. An aggregation advisor for ligand discovery. *J. Med. Chem.* **2015**, *58*, 7076–7087.

24 Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J. Fragment screening: an introduction. *Molecular Biosystems* **2006**, *2*, 430-446.

25 Giannetti, A. M. From experimental design to validated hits: a comprehensive walk-through of fragment lead identification using surface plasmon resonance. In Methods in Enzymology, Lawrence, C. K., Ed. Academic Press: New York, 2011; Vol. Volume 493, pp 169-218.

26 Bamborough, P.; Brown, M. J.; Christopher, J. A.; Chung, C.-w.; Mellor, G. W. Selectivity of kinase inhibitor fragments. *J. Med. Chem.* **2011**, *54*, 5131-5143.

27 Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A. L.; Jahnke, W.; Blommers, M.; Selzer, P.; Jacoby, E. Library design for fragment based screening. *Curr. Top. Med. Chem.* **2005**, *5*, 751-762.

28 Lau, W. F.; Withka, J. M.; Hepworth, D.; Magee, T. V.; Du, Y. J.; Bakken, G. A.; Miller, M. D.; Hendsch, Z. S.; Thanabal, V.; Kolodziej, S. A.; Xing, L.; Hu, Q.; Narasimhan, L. S.; Love, R.; Charlton, M. E.; Hughes, S.; van Hoorn, W. P.; Mills, J. E. Design of a multi-

purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J. Comput. Aided Mol. Des.* **2011**, *25*, 621-36.

29 Chen, I. J.; Hubbard, R. E. Lessons for fragment library design: analysis of output from multiple screening campaigns. *J. Comput. Aided Mol. Des.* **2009**, *23*, 603-620.

30 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. Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification? *J. Biomol. Screen.* **2013**, *18*, 147-159.

31 Silva-Santisteban, M.C.; Westwood, I.M.; Boxall, K.; Brown, N.; Peacock, S.; McAndrew, C.; Barrie, E.; Richards, M.; Mirza, A.; Oliver, A.W.; Burke, R.; Hoelder, S.; Jones, K.; Aherne, G.W.; Blagg, J.; Collins, I.; Garrett, M.D.; van Montfort, R.L. Fragment-based screening maps inhibitor interactions in the ATP-binding site of checkpoint kinase 2. *PLoS One*, **2013**, *12*, e6568.

32 Pollack, S.J.; Beyer, K.S.; Lock, C.; Müller, I.; Sheppard, D.; Lipkin, M.; Hardick, D.; Blurton, P.; Leonard, P.M.; Hubbard, P.A.; Todd, D.; Richardson, C.M.; Ahrens, T.; Baader, M.; Hafenbradl, D.O.; Hilyard, K.; Bürli, R.W. A comparative study of fragment screening methods on the p38α kinase: new methods, new insights. *J. Comput. Aided Mol. Des.* **2011**, *25*, 677-687.

33 Boettcher, A.; Ruedisser, S.; Erbel, P.; Vinzenz, D.; Schiering, N.; Hassiepen, U.; Rigollier, P.; Mayr, L.M.; Woelcke, J. Fragment-based screening by biochemical assays: Systematic feasibility studies with trypsin and MMP12. *J. Biomol. Screen.* **2010**, *15*, 1029-1041.

34 Meiby, E.; Simmonite, H.; le Strat, L.; Davis, B.; Matassova, N.; Moore, J.D.; Mrosek, M.; Murray, J.; Hubbard, R.E.; Ohlson, S. Fragment screening by weak affinity chromatography:

comparison with established techniques for screening against HSP90. *Anal. Chem.* **2013**, *85*, 6756-6766.

35 Schiebel, J.; Radeva, N.; Köster, H.; Metz, A.; Krotzky, T.; Kuhnert, M.; Diederich, W.E.; Heine, A.; Neumann, L.; Atmanene, C.; Roecklin, D.; Vivat-Hannah, V.; Renaud, J.P.; Meinecke, R.; Schlinck, N.; Sitte, A.; Popp, F.; Zeeb, M.; Klebe G. One question, multiple answers: biochemical and biophysical screening methods retrieve deviating fragment hit lists. *ChemMedChem.* **2015**, *10*, 1511-21.

36 Kutchukian, P.S.; Wassermann, A.M.; Lindvall, M.K.; Wright, S.K.; Ottl, J.; Jacob, J.; Scheufler, C.; Marzinzik, A.; Brooijmans, N.; Glick, M. Large scale meta-analysis of fragment-based screening campaigns: privileged fragments and complementary technologies. *J. Biomol. Screen.* **2015**, *20*, 588-96.

37 Wall, I.D.; Hann, M.M.; Leach, A.R.; Pickett, S. D. Current status and future direction of fragment-based drug discovery: a computational chemistry perspective. in Fragment-Based Drug Discovery; edited by Steven Howard, Chris Abell; RSC: London, 2015, pp 73-100.

38 Chessari, G.; Buck, I.M.; Day, J.E.; Day, P.J.; Iqbal, A.; Johnson, C.N.; Lewis, E.J.; Martins, V.; Miller, D.; Reader, M.; Rees, D.C.; Rich, S.J.; Tamanini, E.; Vitorino, M.; Ward, G.A.; Williams, P.A.; Williams, G.; Wilsher, N.E.; Woolford, A.J. Fragment-based drug discovery targeting inhibitor of apoptosis proteins: discovery of a non-alanine lead series with dual activity against cIAP1 and XIAP. *J. Med. Chem.* **2015**, *58*, 6574-6588.

39 Bortolato, A.; Tehan, B.G.; Bodnarchuk, M.S.; Essex, J.W.; Mason, J.S. Water network perturbation in ligand binding: adenosine A2A antagonists as a case study. *J. Chem. Inf. Model.* **2013**, *53*, 1700−1713.

40 Nilar, S.H.; Ma, N.L.; Keller, T.H. The importance of molecular complexity in the design of screening libraries *J. Comput. Aided Mol. Des.* **2013**, *27*, 783-792.

41 Leach, A.R.; Hann, M.M. Molecular complexity and fragment-based drug discovery: ten years on. *Curr. Opin. Chem. Biol.* **2011**, *15*, 489-496.

42 Hann, M.M.; Leach, A.R. Coping with complexity in molecular design. in De Novo Molecular Design, Edited by Schneider, Gisbert. Wiley-VCH Verlag GmbH & Co: Weinheim, 2014, pp 57-77.

43 Ferenczy, G.G.; Keserű, G.M. Thermodynamics of fragment binding. *J. Chem. Inf. Model.* **2012**, *52*, 1039–1045.

44 Mortenson, P.N.; Murray, C.W. Assessing the lipophilicity of fragments and early hits. *J. Comput. Aided Mol. Des.* **2011**, *25*, 663–667.

45 Verdonk, M.L.; Rees, D.C. Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem* **2008**, *3*, 1179-1180.

46 Hung, A.W.; Silvestre, H.L.; Wen, S.; George, G.P.; Boland, J.; Blundell, T.L.; Ciulli, A.; Abell, C. Optimization of inhibitors of Mycobacterium tuberculosis pantothenate synthetase based on group efficiency analysis. *ChemMedChem.* **2016**, *11*, 38-42.

47 Young, R. J. Physical properties in drug design. in Tactics in Contemporary Drug Design. Meanwell, N.A. (ed) pp 1-68, Springer, Berlin, Gemany, 2015.

48 Mattos, C.; Bellamacina, C.R.; Peisach, E.; Pereira, A.; Vitkup, D.; Petsko, G.A.; Ringe, D. Multiple solvent crystal structures: probing binding sites, plasticity and hydration. *J. Mol. Biol.* **2006**, *357*, 1471-1482.

49 Hall, D.R.; Kozakov, D.; Vajda, S. Analysis of protein binding sites by computational solvent mapping. *Methods Mol. Biol.* **2012**, *819*, 13-27.

50 Kozakov, D.; Hall, D.; Napoleon, R. L.; Yueh, C.; Whitty, A.; Vajda, S. New frontiers in druggability. *J. Med. Chem.* **2015**, *58*, 9063-9088.

51 Practical Fragments Blog. http://practicalfragments.blogspot.hu/2013/06/poll-results-how-small-are-your.html (accessed Jan 20, 2016).

52 Van Molle, I.; Thomann, A.; Buckley, D. L.; So, E. C., Lang, S.; Crews, C. M.; Ciulli, A. Dissecting fragment-based lead discovery at the von Hippel-Lindau Protein: hypoxia inducible factor 1α protein-protein interface. *Chem. Biol.* **2012**, *19*, 1300–1312.

53 Fry, D. C.; Wartchow, C.; Graves, B.; Janson, C.; Lukacs, C.; Kammlott, U.; Belunis, C.; Palme, S.; Klein, C.; Vu, B. Deconstruction of a nutlin: dissecting the binding determinants of a potent protein–protein interaction inhibitor. *ACS Med. Chem. Lett.* **2013**, *4*, 660– 665.

54 Lange, G.; Lesuisse, D.; Deprez, P.; Schoot, B.; Loenze, P.; Bénard, D.; Marquette, J.-P.; Broto, P.; Sarubbi, E.; and Mandine, E. Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of pp60Src are identical to those for high affinity binding of full length inhibitors *J. Med. Chem.* **2003**, *46*, 5184– 5195.

55 Andersen, O. A.; Nathubhai, A.; Dixon, M. J.; Eggleston, I. M.; van Aalten, D. M. Structure-based dissection of the natural product cyclopentapeptide chitinase inhibitor argifin. *Chem. Biol.* **2008**, *15*, 295–301.

56 Barelier, S.; Pons, J.; Marcillat, O.; Lancelin, J.-M.; Krimm, I. Fragment-based deconstruction of Bcl-XL inhibitors *J. Med. Chem.* **2010**, *53*, 2577– 2588.

57 Babaoglu, K. and Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2*, 720– 723.

58 Satoh, M.; Saburi, H.; Tanaka, T.; Matsuura, Y.; Naitow, H.; Shimozono, R.; Yamamoto, N.; Inoue, H.; Nakamura, N.; Yoshizawa, Y.; Aoki, T.; Tanimura, R.; Kunishima, N. Multiple binding modes of a small molecule to human Keap1 revealed by X-ray crystallography and molecular dynamics simulation. *FEBS Open Bio.* **2015**, *5*, 557-70.

59 Barelier, S.; Cummings, J. A.; Rauwerdink, A. M.; Hitchcock, D. S.; Farelli, J. D.; Almo, S. C.; Raushel, F. M.; Allen, K. N.; Shoichet, B. K. Substrate deconstruction and the nonadditivity of enzyme recognition. *J. Am. Chem. Soc.* **2014**, *136*, 7374-7382.

60 Brandt, P.; Geitmann, M.; Danielson, U. H. Deconstruction of non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1 for exploration of the optimization landscape of fragments. *J. Med. Chem.* **2011**, *54*, 709– 718.

61 Edink, E.; Rucktooa, P.; Retra, K.; Akdemir, A.; Nahar, T.; Zuiderveld, O.; van Elk, R.; Janssen, E.; van Nierop, P.; van Muijlwijk-Koezen, J.; Smit, A. B.; Sixma, T. K.; Leurs, R.; de Esch, I. J. Fragment growing induces conformational changes in acetylcholine-binding protein: a structural and thermodynamic analysis. *J. Am. Chem. Soc.* **2011**, *133*, 5363–5371.

62 Bauman, J. D.; Patel, D.; Baker, S. F.; Vijayan, R. S.; Xiang, A.; Parhi, A. K.; Martínez-Sobrido, L.; Lavoie, E. J.; Das, K.; Arnold, E. Crystallographic fragment screening and structure-based optimization yields a new class of influenza endonuclease inhibitors. *ACS Chem. Biol.* **2013**, *8*, 2501–2508.

63 Chen, Y.; Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2009**, *5*, 358-364.

64 Orita, M.; Ohno, K.; Warizaya, M. Amano, Y.; Niim, T. Lead generation and examples: opinion regarding how to follow up hits. *Methods in Enzymology* **2011**, *493*, 383-419.

65 Kozakov, D.; Hall, D. R.; Jehle, S.; Luo, L. ; Ochiana, S. O.; Jones, E. V.; Pollastri, M.; Allen, K. N.; Whitty, A.; Vajda, S. Ligand deconstruction: Why some fragment binding positions are conserved and others are not. *Proc. Natl. Acad. Sci. USA*, **2015**, *112*, E2585-E2594.

66 Ludlow, R. F.; Verdonk, M. L.; Saini, H. K.; Tickle, I. J.; Jhoti, H. Detection of secondary binding sites in proteins using fragment screening. *Proc. Natl. Acad. Sci. USA*, **2015**, *112*, 15910–15915.

67 Vass, M.; Ágai-Csongor, É.; Horti, F.; Keserű, G. M. Multiple fragment docking and linking in primary and secondary pockets of dopamine receptors. *ACS Med. Chem. Lett.*, **2014**, *5*, 1010–1014.

68 Christopher, J. A.;, Aves, S.J.; Bennett, K. A.; Doré, A. S.; Errey, J.C.; Jazayeri, A.; Marshall, F. H.; Okrasa, K.; Serrano-Vega, M. J.; Tehan, B. G.; Wiggin, G.R.; Congreve, M. Fragment and structure-based drug discovery for a class C GPCR: Discovery of the mGlu5 negative allosteric modulator HTL14242 (3-Chloro-5-[6-(5-fluoropyridin-2-yl)pyrimidin-4-yl]benzonitrile). *J. Med. Chem.*, **2015**, *58*, 6653–6664.

69 Home page of the Protein Data Bank, http://www.rcsb.org/pdb/home/home.do (accesssed on 22 January, 2016)

70 Murray, C.W., Rees, D.C. Opportunity knocks: organic chemistry for fragment-based drug discovery (FBDD). *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 488-492.

71 Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, *280*, 1−9.

72 Fraser, C. M.; Fernández, A.; Scott, L. R. Dehydron analysis: quantifying the effect of hydrophobic groups on the strength and stability of hydrogen bonds. *Adv. Exp. Med. Biol.* **2010**, *680*, 473−479.

73 Young, T.; Abel, R.; Byungchan, K.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 808−813.

74 Muley, L.; Baum, B.; Smolinski, M.; Freindorf, M.; Heine, A.; Klebe, G.; Hangauer, D. G. Enhancement of hydrophobic interactions and hydrogen bond strength by cooperativity: synthesis, modeling, and molecular dynamics simulations of a congeneric series of thrombin inhibitors. *J. Med. Chem.* **2010**, *53*, 2126−2135.

75 Barelier, S., and Krimm, I. Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr. Opin. Chem. Biol.* **2011**, *15*, 469−474.

76 Hann, M.M. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Commun.*, **2011**, *2*, 349-355.

77 Hopkins, A.L.; Mason, J.S.; Overington, J.P. Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* **2006**, *16*, 127-136.

78 Nair, P. C.; Malde, A. K.; Drinkwater, N.; Mark, A. E. Missing fragments: detecting cooperative binding in fragment-based drug design. *ACS Med Chem Lett.* **2012**, *3*, 322-326.

79 Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R.Fragment-based lead discovery. *Nat. Rev. Drug Dicov.* **2004**, *3*, 660– 672.

80 Borsi, V.; Calderone, V.; Fragai, M.; Luchinat, C.; Sarti, N. Entropic contribution to the linking coefficient in fragment based drug design: a case study. *J. Med. Chem.*, **2010**, *53*, 4285–4289.

81 Nazaré, M.; Matter, H.; Will, D. W.; Wagner, M.; Urmann, M.; Czech, J.; Schreuder, H.; Baue, A.; Ritter, K.; Wehner, V. Fragment deconstruction of small, potent factor Xa inhibitors: exploring the superadditivity energetic of fragment linking in protein-ligand complexes. *Angew. Chem., Int. Ed. Engl.* **2012**, *51*, 905– 911.

82 Murray, C. W.; Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 741–753.

83 Lovering, F.; Bikker, J.; Humblet, C. Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **2009**, *52*, 6752–6756.

84 Ritchie, T. J.; Macdonald, S. J. F. The impact of aromatic ring count on compound developability—are too many aromatic rings a liability in drug design? *Drug Discov. Today* **2009**, *14*, 1011–1020.

85 Kenny, P.W.; Montanari, C.A. Inflation of correlation in the pursuit of drug-likeness. *J. Comput. Aided Mol. Des.* **2013**, *27*, 1-13.

86 Morley, A.D.; Pugliese, A.; Birchall, K.; Bower, J.; Brennan, P.; Brown, N.; Chapman, T.; Drysdale, M.; Gilbert, I.H.; Hoelder, S.; Jordan, A.; Ley, S.V.; Merritt, A.; Miller, D.; Swarbrick, M.E.; Wyatt, P.G. Fragment-based hit identification: thinking in 3D. *Drug Discov. Today* **2013**, *18*, 1221-1227.

87 Pascolutti, M.; Campitelli, M.; Nguyen, B.; Pham, N.; Gorse, A.D.; Quinn, R.J. Capturing Nature's diversity. *PLoS One.* **2015**, *10*, e0120942.

88 Morphy, R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *J. Med. Chem.* **2006**, *49*, 2969-2978.

89 Olsson, T.S.; Williams, M.A.; Pitt, W.R.; Ladbury, J.E. The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. *J. Mol. Biol.* **2008**, *348*, 1002-1017.

90 Kozakov, D.; Grove, L.E.; Hall, D.R.; Bohnuud, T.; Mottarella, S.E.; Luo, L.; Xia, B.; Beglov, D.; Vajda S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* **2015**, *10*, 733-755.

91 Hajduk, P.J.; Huth, J.R.; Fesik, S.W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.*, **2005**, *48*, 2518–2525.

92 Wells, J.A.; McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature*. **2007**, *450*, 1001-1009.

93 Andrews, S.P.; Brown, G.A.; Christopher, J.A. Structure-based and fragment-based GPCR drug discovery. *ChemMedChem* **2014**, *9*, 256–275.

94 Akritopoulou-Zanze, I.; Hajduk, P.J. Kinase-targeted libraries: the design and synthesis of novel, potent, and selective kinase inhibitors. *Drug Discov. Today.* **2009**, *14*, 291-297.

95 Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P.N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E.A.; Wong, B.; Tsang, G.; West, B.L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K.Y.; Artis, D.R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P.S.; Lee, R.J.; Grippo, J.; Puzanov, I.; Kim, K.B.; Ribas, A.; McArthur, G.A.; Sosman, J.A.; Chapman, P.B.; Flaherty, K.T.; Xu, X.; Nathanson, K.L.; Nolop, K. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature.* **2010**, *467*, 596-599.

96 Addie, M.; Ballard, P.; Buttar, D.; Crafter, C.; Currie, G.; Davies, B.R.; Debreczeni, J.; Dry, H.; Dudley, P.; Greenwood, R.; Johnson, P.D.; Kettle, J.G.; Lane, C.; Lamont, G.; Leach, A.; Luke, R.W.; Morris, J.; Ogilvie, D.; Page, K.; Pass, M.; Pearson, S.; Ruston, L. Discovery of 4-amino-N-[(1S)-1-(4-chlorophenyl)-3-hydroxypropyl]-1-(7H-pyrrolo[2,3-d]pyrimidin-4-yl)piperidine-4-carboxamide (AZD5363), an orally bioavailable, potent inhibitor of Akt kinases. *J. Med. Chem.* **2013**, *56*, 2059-2073.

97 Mérour, J.Y.; Buron, F.; Plé, K.; Bonnet, P.; Routier, S. The azaindole framework in the design of kinase inhibitors. *Molecules.* **2014**, *19*, 19935-19979.

98 Karpov, A.S.; Amiri, P.; Bellamacina, C.; Bellance, M.H.; Breitenstein, W.; Daniel, D.; Denay, R.; Fabbro, D.; Fernandez, C.; Galuba, I.; Guerro-Lagasse, S.; Gutmann, S.; Hinh, L.; Jahnke, W.; Klopp, J.; Lai, A.; Lindvall, M.K.; Ma, S.; Möbitz, H.; Pecchi, S.; Rummel, G.;

Shoemaker, K.; Trappe, J.; Voliva, C.; Cowan-Jacob, S.W.; Marzinzik, A.L. Optimization of a dibenzodiazepine hit to a potent and selective allosteric PAK1 inhibitor. *ACS Med. Chem. Lett.* **2015**, *6*, 776-781.

99 Allen, C.E.; Welford, A.J.; Matthews, T.P.; Caldwell, J.J.; Collins, I. Fragment growing to retain or alter the selectivity of anchored kinase hinge-binding fragments. *Med. Chem. Commun.*, **2014**, *5*, 180-185.

100 Bauman, J.D.; Patel, D.; Dharia, C.; Fromer, M.W.; Ahmed, S.; Frenkel, Y.; Vijayan, R.S.; Eck, J.T.; Ho, W.C.; Das, K.; Shatkin, A.J.; Arnold, E. Detecting allosteric sites of HIV-1 reverse transcriptase by X-ray crystallographic fragment screening. *J. Med. Chem.* **2013**, *56*, 2738-2746.

101 Palmieri, L.; Rastelli, G . αC helix displacement as a general approach for allosteric modulation of protein kinases. *Drug Discov. Today*. **2013**, *18*, 407-414.

102 American Society of Hematology Home Page. https://ash.confex.com/ash/2014/webprogram/Paper76344.html (accessed Jan 20, 2016)

103 Murray, J.; Giannetti, A.M.; Steffek, M.; Gibbons, P.; Hearn, B.R.; Cohen, F.; Tam, C.; Pozniak, C.; Bravo, B.; Lewcock, J.; Jaishankar, P.; Ly, C.Q.; Zhao, X.; Tang, Y.; Chugha, P.; Arkin, M.R.; Flygare, J.; Renslo, A.R. Tailoring small molecules for an allosteric site on procaspase-6. *ChemMedChem.* **2014**, *9*, 73-77.

104 Saalau-Bethell, S.M.; Woodhead, A.J.; Chessari, G.; Carr, M.G.; Coyle, J.; Graham, B.; Hiscock, S.D.; Murray, C.W.; Pathuri, P.; Rich, S.J.; Richardson, C.J.; Williams, P.A.; Jhoti, H. Discovery of an allosteric mechanism for the regulation of HCV NS3 protein function. *Nat. Chem. Biol.* **2012**, *8*, 920-925.

105Magee, T.V. Progress in discovery of small-smolecule modulators of protein-pprotein interactions via fragment screening *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2461-2468.

106 Park, C.M.; Bruncko, M.; Adickes, J.; Bauch, J.; Ding, H.; Kunzer, A.; Marsh, K.C.; Nimmer, P.; Shoemaker, A.R.; Song, X.; Tahir, S.K.; Tse, C.; Wang, X.; Wendt, M.D.; Yang, X.; Zhang, H.; Fesik, S.W.; Rosenberg, S.H.; Elmore, S.W. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. *J. Med. Chem.* **2008**, *51*, 6902-6915.

107 Souers, A.J.; Leverson, J.D.; Boghaert, E.R.; Ackler, S.L.; Catron, N.D.; Chen, J.; Dayton, B.D.; Ding, H.; Enschede, S.H.; Fairbrother, W.J.; Huang, D.C.; Hymowitz, S.G.; Jin, S.; Khaw, S.L.; Kovar, P.J.; Lam, L.T.; Lee, J.; Maecker, H.L.; Marsh, K.C.; Mason, K.D.; Mitten, M.J.; Nimmer, P.M.; Oleksijew, A.; Park, C.H.; Park, C.M.; Phillips, D.C.; Roberts, A.W.; Sampath, D.; Seymour, J.F.; Smith, M.L.; Sullivan, G.M.; Tahir, S.K.; Tse, C.; Wendt, M.D.; Xiao, Y.; Xue, J.C.; Zhang, H.; Humerickhouse, R.A.; Rosenberg, S.H.; Elmore, S.W. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **2013**,*19*, 202-208.

108 Burke, J.P.; Bian, Z.; Shaw, S.; Zhao, B.; Goodwin, C.M.; Belmar, J.; Browning, C.F.; Vigil, D.; Friberg, A.; Camper, D.V.; Rossanese, O.W.; Lee, T.; Olejniczak, E.T.; Fesik, S.W. Discovery of tricyclic indoles that potently inhibit Mcl-1 using fragment-based methods and structure-based design. *J. Med. Chem.* **2015**, *58*, 3794-3805.

109 Petros, A.M.; Swann, S.L.; Song, D.; Swinger, K.; Park, C.; Zhang, H.; Wendt, M.D.; Kunzer, A.R.; Souers, A.J.; Sun, C. Fragment-based discovery of potent inhibitors of the anti-apoptotic MCL-1 protein. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 1484-1488.

110 Scott, D.E.; Coyne, A.G.; Venkitaraman, A.; Blundell, T.L.; Abell, C.; Hyvönen, M. Small-molecule inhibitors that target protein-protein interactions in the RAD51 family of recombinases. *ChemMedChem.* **2015**, *10*, 296-303.

111 Maurer, T.; Garrenton, L.S.; Oh, A.; Pitts, K.; Anderson, D.J.; Skelton, N.J.; Fauber, B.P.; Pan, B.; Malek, S.; Stokoe, D.; Ludlam, M.J.; Bowman, K.K.; Wu, J.; Giannetti, A.M.; Starovasnik, M.A.; Mellman, I.; Jackson, P.K.; Rudolph, J.; Wang, W.; Fang, G. Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc. Natl. Acad. Sci. USA,* **2012**, *109*, 5299-5304.

112 Sun, Q.; Burke, J.P.; Phan, J.; Burns, M.C.; Olejniczak, E.T.; Waterson, A.G.; Lee, T.; Rossanese, O.W.; Fesik, S.W. Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 6140-6143.

113 Erlanson DA, Braisted AC, Raphael DR, Randal M, Stroud RM, Gordon EM, Wells JA. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 9367-72.

114 Erlanson, D.A.; Wells, J.A.; Braisted, A.C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199-223.

115 Ostrem, J.M.; Peters, U.; Sos, M.L.; Wells, J.A.; Shokat, K.M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*. **2013**, *503*, 548-551.

116 Kathman, S.G.; Xu, Z.; Statsyuk, A.V. A fragment-based method to discover irreversible covalent inhibitors of cysteine proteases. *J. Med. Chem.* **2014**, *57*, 4969-4974.

117 Miller, R.M.; Paavilainen, V.O.; Krishnan, S.; Serafimova, I.M.; Taunton, J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J. Am. Chem. Soc.* **2013**, *135*, 5298-5301.

118 London, N.; Miller, R.M.; Krishnan, S.; Uchida, K.; Irwin, J.J.; Eidam, O.; Gibold, L.; Cimermančič, P.; Bonnet, R.; Shoichet, B.K.; Taunton, J. Covalent docking of large libraries for the discovery of chemical probes. *Nat. Chem. Biol.* **2014**, *10*, 1066-1072.

119 Blomberg, N.; Cosgrove, D.A.; Kenny, P.W.; Kolmodin, K. Design of compound libraries for fragment screening. *J. Comput. Aided. Mol. Des.* **2009**, *23*, 513-525.

120 Lewell, X. Q.; Judd, D. B.; Watson, S. P.; Hann, M. M. RECAP retrosynthetic combinatorial analysis procedure: A powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511-522.

121 Cox, O.B.; Krojer, T.; Collins, P.; Monteiro, O.; Talon, R.; Bradley, A.; Fedorov, O.; Amin, J.; Marsden, B.D.; Spencer, J.; von Delft, F.; Brennan, P.E. A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.*, **2016**, *7*, 2322-2330.

122 Fejzo, J.; Lepre, C.A.; Peng, J.W.; Bemis, G.W.; Ajay; Murcko, M.A.; Moore, J.M. The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem. Biol.* **1999**, *6*, 755-769.

123 Pihan, E.; Colliandre, L.; Guichou, J.F.; Douguet, D. E-Drugs3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. *Bioinformatics*, **2012**, *28*, 1540-1541.

124 Hung, A.W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J.A.; Clemons, P.A.; Young, D.W. Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl. Acad. Sci. USA*, **2011** *108*, 6799-6804.

125 Feher, M.; Schmidt, J.M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218-227.

126 Ertl, P.; Roggo, S.; Schuffenhauer, A. Natural product-likeness score and its application for prioritization of compound libraries. *J. Chem. Inf. Model.* **2008**, *48*, 68-74.

127 Over, B.; Wetzel, S.; Grütter, C.; Nakai, Y.; Renner, S.; Rauh, D.; Waldmann, H. Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* **2013**, *5*, 21-28.

128 Vu, H.; Roullier, C.; Campitelli, M.; Trenholme, K.R.; Gardiner, D.L.; Andrews, K.T.; Skinner-Adams, T.; Crowther, G.J.; Van Voorhis, W.C.; Quinn, R.J. Plasmodium gametocyte inhibition identified from a natural-product-based fragment library. *ACS Chem. Biol.* **2013**, *8*, 2654-2659.

129 Home page of Cambridge Medchem Consulting, http://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_collections.html (accessed 22 January, 2016)

130 Wilcken, R.; Zimmermann; M.O., Lange, A.; Joerger, A.C.; Boeckler, F.M. Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J. Med. Chem.* **2013**, *56*, 1363-1388.

131 Cambridge MedChem Consulting Home Page. http://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_collection_profiles.html (accessed Apr 4, 2016).

132 Nadin, A.; Hattotuwagama, C.; Churcher, I. Lead-oriented synthesis: A new opportunity for synthetic chemistry, *Angewandte Chemie Int. Ed.* **2012**, *51*, 1114-1122.

133 Cernak, T.; Dykstra, K.D.; Tyagarajan, S.; Vachal, P.; Krska, S.W. The medicinal chemist's toolbox for late stage functionalization of drug-like molecules. *Chem. Soc. Rev.* **2016**, *45*, 546-576.

134 Lengyel, L.C.; Sipos, G.; Sipocz, T.; Vago, T.; Dorman, G.; Gerencser, J.; Makara, G.; Darvas, F. Synthesis of condensed heterocycles by the Gould–Jacobs reaction in a novel three-mode pyrolysis reactor. *Org. Process Res. Dev.*, **2015**, *19*, 399–409.

135 Personal communication from Justin Bower, 4 April 2016.

136 Practical Fragments Blog. http://practicalfragments.blogspot.com/2014/10/caveat-emptor.html (accessed Jan 20, 2016).

137 Doak, B.C.; Morton, C.J.; Simpson, J.S.; Scanlon, M.J. Design and evaluation of the performance of an NMR screening fragment library. *Austral. J. Chem.* **2013**, *66*, 1465-1472.

138 Davies, D.R.; Mamat, B.; Magnusson, O.T.; Christensen, J.; Haraldsson, M.H.; Mishra, R.; Pease, B.; Hansen, E.; Singh, J.; Zembower, D.; Kim, H.; Kiselyov, A.S.; Burgin, A.B.; Gurney, M.E.; Stewart, L.J. Discovery of leukotriene A4 hydrolase inhibitors using metabolomics biased fragment crystallography. *J. Med. Chem.* **2009**, *52*, 4694-4715.

139 Procházková, E.; Jansa, P.; Březinová, A.; Cechová, L.; Mertlíková-Kaiserová, H.; Holý, A.; Dračínský, M. Compound instability in dimethyl sulphoxide, case studies with 5-aminopyrimidines and the implications for compound storage and screening. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6405-6409.

140 McElhinny, C.J. Jr.; Lewin, A.H.; Mascarella, S.W.; Runyon, S.; Brieaddy, L.; Carroll, F.I. Hydrolytic instability of the important orexin 1 receptor antagonist SB-334867: possible confounding effects on in vivo and in vitro studies. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6661-6664.

141 Baurin, N.; Aboul-Ela, F.; Barril, X.; Davis, B.; Drysdale, M.; Dymock, B.; Finch, H.; Fromont, C.; Richardson, C.; Simmonite, H.; Hubbard, R.E. Design and characterization of libraries of molecular fragments for use in NMR screening against protein targets. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157-2166.

142 Kozikowski, B.A.; Burt, T.M.; Tirey, D.A.; Williams, L.E.; Kuzmak, B.R.; Stanton, D.T.; Morand, K.L.; Nelson, S.L. The effect of room-temperature storage on the stability of compounds in DMSO. *J. Biomol. Screen.* **2003**, *8*, 205-209.

143 Oldenburg, K., Pooler, D., Scudder, K., Lipinski, C., Kelly, M. High throughput sonication: evaluation for compound solubilization. *Comb. Chem. High Throughput Screen.* **2005**, *8*, 499-512.

144 Practical Fragments Blog. http://practicalfragments.blogspot.com/2014/02/poll-results-how-do-you-store-your.html (accessed Jan 20, 2016).

145 Practical Fragments Blog. http://practicalfragments.blogspot.com/2015/01/fragments-in-clinic-2015-edition.html (accessed Jan 20, 2016).

Table of Contents graphic.