LINK TO ORIGINAL ARTICLE

Nature Reviews Drug Discovery | AOP, published online 12 July 2013; doi:10.1038/nrd3926-c1

The 'rule of three' for fragment-based drug discovery: where are we now?

Harren Jhoti, Glyn Williams, David C. Rees and Christopher W. Murray

During the past decade, fragment-based drug discovery (FBDD) has matured from a niche activity into an approach that is widely used across industry and academia, as highlighted in a recent news story (*Nature Rev. Drug Discov.* **12**, 5–7 (2013))[1]. Indeed, there is now even some evidence that leads generated by FBDD have better physicochemical properties than those generated from conventional screening strategies[2].

Ten years ago, we published the 'rule of three' (RO3)[3] — a set of guidelines describing desirable physicochemical properties for molecules in FBDD screening collections, which we believed at the time would be useful to those interested in applying the approach. Here, we discuss our views on the application of the RO3, as well as other facets of FBDD that are currently being debated.

Rule of three

The value of using the RO3 in FBDD is being (correctly) challenged[4] as many useful fragment hits do not adhere to its restrictions (see the Practical Fragments blog website). We have always believed that the RO3 concept has limitations and, like other rules related to desirable physicochemical properties, such as the 'rule of five' guidelines[5], it is simply a guideline that should not be overemphasized. Nevertheless, we believe that the RO3 has been useful in ensuring that fragment libraries really do consist of compounds with fragment-like properties, although reports of 'fragment' hits that actually resemble lead-like compounds — particularly considering their substantial molecular mass — do still appear.

The fundamental concept of fragment screening is to use simpler molecules so that the chemical space can be sampled much more efficiently than is possible when using molecules of greater complexity[6]. One consequence of screening smaller, simpler fragments is that their affinity is expected to be relatively low (>1 mM), given the limited numbers of potential interactions that they can make with the protein. Although weak in potency, fragment hits make high-quality interactions with the protein as they must overcome a substantial entropic barrier to binding, relative to their size. The detection of low-potency fragment hits presents substantial technical challenges, but to avoid these challenges and succumb to the temptation to screen larger, more complex molecules runs counter to the whole *raison d'être* of fragment screening.

We now have experience of over 30 fragment screens against a broad range of protein classes, and our results support the view that less complex molecules give a higher hit rate. Consequently, our fragment library now almost exclusively contains molecules with fewer than 17 non-hydrogen atoms (molecular mass <230 Da). As such, we believe the RO3 concept has assisted in limiting the molecular complexity in our fragment libraries.

Three-dimensional fragments

Several groups have reported initiatives to construct libraries of three-dimensional fragments[7,8] (see the 3D Fragment Consortium website), often designed to target protein–protein interfaces (PPIs), in the belief that fragment-based screening is unlikely to yield hits unless more 'three-dimensionality' can be incorporated into their fragments. Two assumptions seem to be implicit in this argument: first, that current fragment libraries are largely composed of 'flat structures'; and second, that pockets on PPIs are fundamentally different from other protein binding sites.

The first assumption may have arisen from an examination of the large number of fragment hits that have been reported against kinase ATP-binding pockets. Indeed, the vast majority of these fragments can be described as rather flat, which is not surprising as they mimic the adenine base of ATP. However, such hits can provide a false perception of the corresponding fragment library. For example, we have generated multiple fragment hits for kinase targets yet the majority of fragments in our library contain considerable three-dimensional characteristics.

The second assumption as to whether PPIs are fundamentally different to other types of targets depends on the level at which the protein architecture is probed. All proteins contain the same fundamental motifs: the amino acids. However, the apparent similarity (or difference) between two binding pockets will depend on the complexity of the molecules used to probe them. For example, a water molecule is unlikely to discriminate between a PPI and an enzyme binding site because it only requires two hydrogen-bonding interactions, which are usually available in both protein classes. However, as molecules become more complex (and lead-like), they need to form substantially more interactions in order to bind efficiently, which also means that one has to screen exponentially increasing numbers to find hits. This description of molecular recognition is well established and underpins FBDD. Fragments are — or should be — simple enough to probe the basic architecture of all proteins yet have sufficient complexity to allow them to be elaborated into lead compounds.

If one accepts that proteins present common recognition motifs to fragments, then we would expect some fragments to bind to different protein classes. Furthermore, this argues that fragment libraries (unless specifically targeted) should not display a bias for any specific protein class. This is exactly our experience, as a high percentage of the hits that we observe binding to PPIs have also been detected as hits against other proteins. We, as well as others, have reported that the fragment binding event often triggers a conformational change in the protein (including PPIs), resulting in the formation of new pockets that can be exploited during the fragment hit-to-lead process[9]. Therefore, although we would still encourage efforts to design more three-dimensional fragments to target PPIs, we would also caution against a tendency to increase fragment size.

Fragment detection

Another area of active debate centres on the lack of correlation of fragment hits obtained using different detection techniques. Given the inherently low affinity of fragment binding, the choice of detection technique and the experimental conditions must both be carefully considered. For example, we routinely use ligand-observed NMR and X-ray crystallography in our fragment screens and typically observe a correlation of 30–40% using these two techniques. We consider this correlation to be surprisingly good, given the differences in the experimental conditions

and the nature of the measurements. Among the many reasons for the discrepancy, only about 5% of a fragment needs to interact with the protein to be detected as an NMR hit, whereas in X-ray crystallography experiments a fragment needs to have at least 70% occupancy of the binding site to be defined as a hit. Consequently, NMR can detect hits with solubilities that are lower than their potencies, but these hits will be missed by X-ray crystallography. For this reason, great importance should be placed on accurately characterizing the solubility of a fragment library. These technical considerations have driven us to segment our fragment library into sublibraries that have appropriate physicochemical properties for different detection techniques.

Although better cross-validation would provide increased confidence in fragment hits, there is a danger that this may also result in the systematic selection of more potent hits, and this strategy implicitly places a reliance on the least sensitive technique. This is of particular concern as the most potent fragment is often not the best starting point for hit-to-lead chemistry. Furthermore, some techniques that are currently used for fragment screening, such as high-concentration screening with bioassays, may not be optimal owing to their known tendency to generate false positives.

Our strategy for avoiding false positives involves emphasizing X-ray crystallography. Although there is no question that fragment screening using X-ray crystallography will result, like most other techniques, in some false negatives, the near-absence of false positives using this approach, coupled with its ability to detect the most weakly binding fragments, means that it remains our most successful method. Of course, not all proteins can be crystallized, and other biophysical techniques may be considerably easier to implement. However, when available, the quality of information generated from a protein fragment crystal structure remains substantially higher than other techniques such as surface plasmon resonance (SPR), calorimetry, mass spectrometry, bioassays and even NMR. Specifically, it allows rapid evaluation of the binding modes of different fragment hits in order to assess which ones should be progressed into hit-to-lead chemistry.

Fragment hits to leads

Another challenge that is often highlighted is how to transform fragment hits into useful lead compounds. At this stage, multiple criteria should be assessed to decide which fragments to progress. These include: physicochemical properties such as the calculated octanol–water partition coefficient (cLogP); efficiency metrics such ligand efficiency (LE) and lipophilic ligand efficiency (LLE); the experimentally determined binding mode with a consideration of the growth vectors and potential for new interactions; and synthetic accessibility. However, it would be a mistake to assume that problems with converting fragments into lead compounds are disconnected from deficiencies in the fragment library itself. By analogy, high-throughput screening (HTS) often delivers micromolar hits, but it is seldom argued that difficulties in converting HTS hits into leads are disconnected from the quality of the HTS library. Contrary to some perceptions, not all fragment libraries are alike.

We typically choose four to six fragment hits and explore relatively small changes in order to optimize the binding interactions of each fragment. We have found this to be more effective than trying to make large changes to the molecule. One of the reasons why we have favoured this fragment-growing approach over fragment-linking is that it allows better control of the physicochemical properties of the eventual lead molecule. During the fragment hit-to-lead process, we typically synthesize 50–100 compounds to increase the binding affinity from millimolar to nanomolar, and we routinely generate multiple lead series for each target. This process relies on extensive X-ray crystallographic data; indeed, we believe that fragment optimization in the absence of such structural information might prove to be an extremely challenging endeavour.

Concluding remarks

We trust that our comments, some of which are deliberately provocative, on these many facets of FBDD will generate active discussion and might assist in improving the success of this approach for the broader drug discovery community. Over the past decade, we — like many other groups — have addressed considerable technical challenges in implementing FBDD. As our experience grows, our rules, including the RO3, evolve. Finally, we would encourage senior management to invest appropriately in this area of lead discovery otherwise there will be a substantial risk of generating poor-quality drug candidates, which will lead to more costly failures in later-stage development.

Harren Jhoti, Glyn Williams, David C. Rees and Christopher W. Murray are at Astex Pharmaceuticals, 436 Cambridge Science Park, Cambridge CB4 0QA, UK.
Correspondence to H.J.
e-mail:harren.jhoti@astx.com

doi:10.1038/nrd3926-c1
Published online 12 July 2013

1. Baker, M. Fragment-based lead discovery grows up. *Nature Rev. Drug Discov.* **12**, 5–7 (2013).
2. Leeson, P. D. & St-Gallay, S. A. The influence of the 'organisational factor' on compound quality in drug discovery. *Nature Rev. Drug Discov.* **10**, 749–765 (2011).
3. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A 'Rule of Three' for fragment-based lead discovery. *Drug Discov. Today* **8**, 876–877 (2003).
This paper suggests that fragments are defined using a 'rule of three' concept, in which the molecular weight of a fragment is <300, the cLogP is ≤3, the number of hydrogen bond donors is ≤3 and the number of hydrogen bond acceptors is ≤3. The latter two criteria have not been widely adopted. Indeed, there are ambiguities in how donors and acceptors are defined.
4. Koester, H. *et al.* Small non-rule of 3 compatible fragment library provides high hit rate of endothiapepsin crystal structures with various fragment chemotypes. *J. Med. Chem.* **54**, 7784–7796 (2011).
5. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.* **46**, 3–26 (2001).
6. 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.* **41**, 856–864 (2001).
7. Over, B. *et al.* Natural-product-derived fragments for fragment-based ligand discovery. *Nature Chem.* **5**, 21–28 (2013).
8. Hung, A. W. *et al.* Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl Acad. Sci. USA* **108**, 6799–6804 (2011).
9. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).

Competing interests statement
The authors declare competing financial interests: see Web version for details.

FURTHER INFORMATION
3D Fragment Consortium: http://www.3dfrag.org
Practical Fragments blog: http://practicalfragments.blogspot.co.uk

ALL LINKS ARE ACTIVE IN THE ONLINE PDF