J Comput Aided Mol Des (2009) 23:453–458
DOI 10.1007/s10822-009-9292-1

WARR'S PIECE

Fragment-based drug discovery

Wendy A. Warr

Received: 1 June 2009 / Accepted: 4 June 2009 / Published online: 24 June 2009

Introduction

As long ago as the 1980s it was hypothesized that the total affinity of a ligand for a protein is a function of its constituent parts ("fragments"). Later, Roche's "needles" were an early example of fragments [1]. It was not until the mid-1990s that technologies for measuring the affinities of weakly binding ligands were developed; workers at Abbott (in "SAR by NMR") proved that meaningful SAR and stable binding modes could be observed even with mM ligands [2]. (Fragments typically bind with K_d between 100 μM and 10 mM.) At around that same time, X-ray crystallography was beginning to be used to map out "hot spots" in protein binding sites [3]. Soon afterward, X-ray crystallography for fragment-based screening was being used at Abbott and other companies. Both NMR and X-ray analyses provide structural information about the binding site of a hit.

Two concepts underlie the appeal of fragment-based drug discovery (FBDD): first, chemical space can be more efficiently probed by screening collections of small fragments rather than libraries of larger molecules, and second, the binding efficiency per atom is at least as high as for larger hit molecules. In FBDD several thousand compounds might be screened to find weak binders whereas in high throughput screening (HTS) a million or more compounds might be screened to find hits with K_d less than 1 μM. A million compounds cover only a tiny proportion of the available chemical space of 10^{60} compounds, whereas screening 10,000 smaller compounds covers a much higher proportion of chemical diversity space, since there are only about 14 million compounds with molecular weight below 160 Da. (Note, however, that the fraction of the total diversity space explored is still small.)

Less complex molecules should also show higher hit rates [4]. Reportedly [5], Novartis has observed FBDD hit rates 10–1,000 times higher than for HTS; deCODE genetics claimed hit rates of 2–8%. In this issue, Rod Hubbard of Vernalis reports an average hit rate of 3.2%. Of course the rate depends on the target, the library and the cutoff for a "hit": Hubbard gets a rate of perhaps 2–3% for protein–protein interactions and 8–10% for kinases (by NMR).

The second major concept of FBDD concerns binding efficiency. Kuntz et al. [6] showed that high molecular weight is not a necessity for high binding affinity and, since then, the ligand efficiency (LE) measure, binding energy divided by the number of heavy atoms [7], has been devised, but other measures such as Ligand-Lipophilicity Efficiency (LLE) [8], Group Efficiency (GE), the estimation of an individual group's contribution toward the overall free energy of binding [9], and others [10] have been proposed. LE is now being adopted not only for FBDD but also more generally by medicinal chemists. The "Astex Rule of 3" (three or fewer hydrogen bond acceptors, three or fewer hydrogen bond donors, and CLogP $\leq$ 3) has been suggested for selecting suitable fragments [11].

Once hits have been found they must be fully characterized before crystallography or SAR studies are carried out. There may be many false positives: one speaker at the recent Royal Society of Chemistry (RSC) meeting [5] reported that 450 hits gave more than 70% inhibition but only 135 gave a pIC_{50}. Incidentally, Hubbard's paper in

W. A. Warr (✉)
Wendy Warr & Associates, 6 Berwick Court, Holmes Chapel, Cheshire CW4 7HZ, UK
e-mail: wendy@warr.com

this issue suggests that there are few promiscuous fragments. Surface plasmon resonance (SPR) is a relatively recent technology for distinguishing good hits from ones which bind nonspecifically; the emergence of SPR as a robust technique was one of the noticeable differences between RSC's "Fragments 2009" and "Fragments 2007". NMR also shows if a fragment is bound in the active site. Many practitioners feel that determining the X-ray structure of the fragment and target is an essential next step but NMR and modeling may be used if an X-ray structure cannot be found. Structure-based drug design (SBDD) can then be used in optimization of the fragment.

One problem lies in ligand specificity: fragment-based approaches identify and characterize only "hot spots", i.e., the regions of a protein surface that are major contributors to the ligand binding free energy. Unfortunately many binding sites in the active site that are responsible for target specificity and/or selectivity are not included in these "hot spots". Smallness inevitably means selectivity will be less [4] but even so, Hubbard has shown that different kinases have different fragment hits, and different affinities and different binding modes [12]. Fragment screening finds the most efficient binder in the smallest core but optimization is still needed.

FBDD approaches enable the efficient development of novel leads as these technologies are design intensive rather than resource intensive, as compared to HTS. The low cost of market entry has led to the formation of many specialized companies (see Table 1, which does not include "instrument companies", protein crystallography companies, chemical library suppliers, and pharmaceutical companies). A few universities and research institutes have also adopted FBDD.

Screening methods

Recent developments in fragment-based screening technologies have resulted in higher throughput, increased cost effectiveness, and reductions in protein requirements. NMR is still the commonest way of finding leads. Unfortunately, both NMR and X-ray crystallography screening rely on high concentrations of both fragments and target, so the fragments must be soluble at high concentrations. SPR has the advantage of providing quantitative dynamics data on the binding interaction, such as binding constants, which are complementary to the structural information from X-ray and NMR screens. At the RSC meeting, Hubbard talked of making SPR work robustly: overbinding on the surface of the protein can lead to misleading SAR. Vernalis' SeeDs technology [12] adds an NMR competitive binding process to validate that binding has occurred. Some companies have also used thermal methods (isothermal titration calorimetry or protein thermal unfolding), mass spectrometry (MS) and high concentration bioassays. Daniel Erlanson, formerly of Sunesis and now at Carmot Therapeutics, reported [5] on MS and tethering with extenders (small molecules that bind to an active-site cysteine and contain a free thiol). Evotec applies a biochemical screening approach with protein NMR as a complement before X-ray crystallography: their FCS + fluorescence correlation spectroscopy is more sensitive than a traditional biochemical screen.

Orthogonal validation (using two or even three assay methods in parallel) is a fairly new trend. At the RSC meeting, deCODE reported one case where the oligomeric state of a protein had a significant effect on the screening results: there was poor correlation between NMR, SPR, and X-ray results. Screening of oligomeric proteins, proteins with multiple conformational states, or proteins in protein–protein interactions is challenging. Use of multiple methods will probably be necessary in such cases.

Targets

Most corporate collections contain molecules that have been optimized for historical targets. HTS sometimes fails to find hits that interfere with new targets, such as protein–protein interaction surfaces [13]. Hubbard says: "There's only a small chance that within even a large corporate collection you'd find compounds with the perfect constellation of shapes to fit a new target". FBDD is robust for novel target classes. For example, BCL targets are difficult; Abbott's work on Bcl-2 [14] and on Bcl-x_L [15] has encouraged people. Abbott's ABT737 is a rare example of a hit [14] salvaged from HTS. Targets can be "fast followers" or they can be novel, lacking in validation. Astex have been very successful partly because they have pursued targets that have some precedent.

Where FBDD is used to examine a novel target, the success rate is likely to be dramatically lower. Even though the work is a technical success, the program may fail. The chance of a novel target being relevant to disease is perhaps 1 in 7 or 1 in 10. The limiting factor in success here is the *biology* needed to move the lead into the clinic. In the case of Bcl-x, Abbott did not at first know what the target was supposed to do. Richard Law in this issue shows that FBDD methods can find chemotypes that bind to difficult targets such as BACE. Integral membrane proteins have so far proven to be extraordinarily difficult to produce and crystallize. ZoBio claim that their technology [16] even works for targets that cannot be obtained in large quantities and/or are insoluble (such as integral membrane proteins).

FBDD can also find new binding modes and allosteric sites. At the RSC meeting, Pfizer reported looking for

Table 1 Selected FBDD companies

Company	Web site	Expertise
Astex	http://www.astex-therapeutics.com/	Pyramid: proprietary, integrated, fragment-based drug discovery (high-throughput biophysical and computational techniques)
Beactica	http://www.beactica.com/	Sprint: integrated fragment-based platform that considers molecular properties in combination with an interaction analysis system
BioLeap	http://www.bioleap.com/	Computational fragment-based drug design platform
BioSolveIT	http://www.biosolveit.de/	Tools for virtual fragment- based drug design and library generation: CoLibri, ReCore, FlexNovo, FTrees-FS
Carmot therapeutics	http://www.carmot.us/	Chemotype evolution using a tethering with extenders technology licensed from Sunesis
Crystax pharmaceuticals	http://www.crystax.com/	Core areas of expertise are: X-ray crystallography, NMR, fragment screening and hit optimization by biophysical analysis and computational chemistry
deCODE Genetics	http://www.decode.com/	Fragments of Life: proprietary technology for fragment-based lead discovery and optimization
Evotec	http://www.evotec.com/en/	Proprietary EVOlution platform: fragment library, fragment screening technologies (biochemical and NMR-based) and optimization of fragments
Graffinity pharmaceuticals	http://www.graffinity.com/	Chemical microarrays, fragment screening, SPR label-free imaging of protein-ligand interactions, computational chemistry and lead discovery
Ibis biosciences (recently acquired by Abbott)	http://www.ibisbiosciences.com/	SAR by MS, against structured RNA targets
IOTA pharmaceuticals	http://www.iotapharma.com/	Provides fragment-based drug design services (molecular design, compound screening and lead discovery services) to partners in the pharmaceutical industry
Locus pharmaceuticals	http://www.locusdiscovery.com/	CORASYS: proprietary virtual fragment-based drug design platform
MEDIT	http://medit-pharma.com/	MED-SuMo computational fragment-based drug design protocol
Plexxikon	http://www.plexxikon.com	Scaffold-like screening library and proprietary Scaffold-Based Drug Discovery Platform
Proteros biostructures	http://proteros.com/	X-ray protein structure analysis
Proteros fragments	http://www.proteros-fragments.com/	Customized leads from screening of fragment library and structure-guided lead generation
Pyxis discovery	http://www.pyxis-discovery.com/	Services offered include fragment-based lead discovery
Sareum (recently acquired by Galapagos)	http://www.sareum.co.uk/	Fragment-based lead identification and optimization; specialists in structure-based drug discovery
SGX pharmaceuticals (acquired by Eli Lilly in 2008)	http://www.sgxpharma.com/	FAST: proprietary fragment-based, protein structure guided drug discovery technology
Structure based design	http://strbd.com/service.php	Crystal structure determination and fragment- based drug discovery (crystallographic screen)
Sunesis (preclinical research ceased 2008)	http://sunesis.com/	Tethering technology to screen drug fragments
Vernalis	http://www.vernalis.com	Research capability based on structure-based drug discovery technologies. SeeDs: structural exploitation of experimental drug startpoints
Zenobia therapeutics (formed by former ActiveSite employees)	http://www.zenobiatherapeutics.com/	Fragment-based lead discovery; expertise in protein biochemistry and crystallization emphasized
ZoBio	http://www.zobio.com/	TINS: Target Immobilized NMR Screening

compounds that bind at allosteric sites and finding a hit that induces a novel binding pocket for ITK. DeCODE, working on HCV helicase, non-structural protein 3 (NS3), found a fragment bound in the NS3 site and in another site on the surface. This could be synergistic for inhibition; it could be that a dual inhibitor will have a reduced rate of resistance.

Hajduk and Greer state that a fragment screen provides a rapid and reliable means of interrogating a protein target

for druggability before investing in further discovery research [17]. Results from Carlson et al. [18] support the mathematical models of Hajduk and his colleagues, and they have implications for predicting druggability, but they show that LE should not be applied too rigidly across all protein targets, since enzymes and nonenzymes differ. Druggability score (DScore) calculated by Schrödinger's SiteMap has allowed Hubbard (this issue) to differentiate targets of low and high hit rates.

Big pharma

The discipline has matured and large pharmaceutical companies that used to be critical are now bringing fragment-based approaches to the fore, but for successful implementation of FBDD, researchers from different disciplines must collaborate. Big pharmaceutical companies have the required skill sets but bringing them together has not necessarily been easy for cultural and practical reasons. Different streams of line management and different priorities may mean that even pharmaceutical companies enthusiastic about FBDD are not organized well to carry it out. Compounds with low potency may not appeal to the medicinal chemist; commitment to timely structural biology may be difficult to achieve; and the blockbuster culture is well entrenched, so that a high proportion of the company's targets (e.g., GPCRs and ion channels) may not be suitable for FBDD. A lot of larger companies are countering these problems by partnering with smaller biotechs [13].

A new trend is the increasing use of fragments as a window on the corporate collection. It is not surprising that the complementary use of HTS and FBDD is a common theme in the larger companies: "high-concentration screening" and "reduced complexity screening" are frequently reported. The compound collections used are a hybrid of a true fragment library and a typical HTS collection. A publication (admittedly by a FBDD enthusiast) holds that as the size of a fragment increases, the size of the library needed increases, and as complexity grows, the chances of an H-bonding mismatch or steric clash between the fragment and the target protein rapidly increase, reducing the chance of finding hits. A larger library is thus needed to achieve the required hit rate and space coverage, but it is harder to design a diverse, leadlike library [9]. It is also claimed that the smallest hits in such a library will be undetectable.

At the RSC meeting a Pfizer speaker discussed the complementary nature of FBDD and HTS. The availability of a high concentration assay allows SAR generation. Fragment screening data informs HTS series selection. Fragment screening identifies optimal cores well, while HTS identifies selectivity features well. Compared with HTS, fragment screening can deliver more hits against larger numbers of protein targets. Artifacts in HTS and false positives can lead to low confirmation rates. HTS can be the quickest route to generating initial leads but fragment screening is generally more reliable [17], because low molecular weight compounds tend to be more soluble and methods of detection are simpler and more robust.

Publications describing the design and characterization of fragment libraries are becoming more common. GSK's fragment libraries total about 8,000 molecules; a reduced complexity set has 5,870 molecules, designed by automatic methods but all of them visually inspected [5]. There is a trend for larger pharmas to build larger libraries, although the size of the library depends on the assay and the size of the compounds. Earlier practitioners preferred a molecular weight of 110–250 Da. In this journal issue Hubbard claims that a library of 1,200 fragments of that size would give 40–100 hits, which was quite enough. Big pharma can afford to do things on a larger scale: GSK reportedly "shot" about 4,600 crystals in 2008.

Optimization

Once a hit has been found, it is optimized into a lead by one of three approaches: linking, growing, or merging. Linking seems like a good idea (additive delta-G) but it can be hard to find second site fragments and two pK_d can be lost in conformational strain of the linker. There have been successes (as in Abbott's early SAR by NMR work) but it is often impossible to develop the appropriate chemistry to link fragments while maintaining the orientation and position needed for key interactions. It is usually more successful to grow fragments by structure guided medicinal chemistry or grow by using the fragment binding motif to search for similar compounds that can be purchased (a method often called "SAR by catalog"). One leading academic, however, is cautious about "SAR by catalog". One of his current research interests is analyzing libraries for purchase, as a result of which he is also synthesizing focused libraries for second round screening, because his analysis shows that many vendors sell compounds that are too diverse or esoteric, and that are not soluble enough for FBDD.

There are challenges to merging; not many people have taken this approach. You need to mobilize considerable chemistry. Growth strategies are quicker but intellectual property around the core may limit the growing strategy. Hubbard (in this issue) says that a particularly attractive approach is to use the structure of fragments (and other hit compounds) binding to the target site to design new compounds that combine key interaction features. This is a sort of merging, or scaffold hopping.

Efficiency and potency are the criteria used in optimization but obviously potency alone is not enough: you also need to consider lipophilicity, polarity, charge, stability etc. It is not easy to maintain LE during fragment growth and during optimization, the binding mode of a lead may change. The static receptor hypothesis is also dangerous.

According to Hajduk and Greer [17] the ability to produce potent inhibitors after initiating lead optimization on fragment leads nearly triples with the aid of SBDD. The optimization of fragment leads can be predicted with high accuracy: there is a linear relationship between potency and molecular mass along the path of ideal optimization. Hubbard's take-home message is that FBDD does not make drug design easier but it does give you more options. Optimization is still difficult and in some cases it is an intractable problem. Research continues on in situ fragment assembly techniques such as click chemistry, dynamic combinatorial library design, and tethering with extenders.

Astex expect a 1 mM fragment hit to be transformed to a 1 μM lead in 6 months, and to a 100 nM lead in 9–12 months [13]. They have one compound in clinical trials (AT7519) that took only 18 months from first synthesis to dosing in patients. Vernalis has reported that structure-guided lead optimization takes about 15–18 months. A GSK scientist has said [5] that next year he may be able to talk about a candidate found in just over 2 years. Will FBDD really have a big impact on time to market? Probably not. FBDD might find hits fast but optimization is no faster than with any other method of hit detection and might even be slower. What *has* caused speed-up of optimization is SBDD.

The future

There are still many opportunities for new development, such as improving the novelty, structural diversity and physicochemical properties of fragment libraries. It might also be possible to push detection methods to find 1–10 mM hits. Can we identify new types of interactions: protein–protein interactions, novel templates and new binding modes? There is plenty of scope for increasing the efficiency of fragment optimization. There will be retrospective analysis of projects. One challenge is deciding which fragments to progress, other than using the subjective decisions of a medicinal chemist. The thermodynamics of binding need investigation. Further developments in modeling and computational chemistry [19] can be expected, including methods for predicting binding modes, and examination of entropy and desolvation. The sheer number of fragments detected nowadays is challenging for the chemist. Tools for assessing synthetic accessibility may help. We also need progress with optimizing fragments in the absence of a crystal structure; greatly improved methods will be needed if detailed SAR information is to be gathered in the 100 μM to 1 mM potency range. Further progress in structural biology (e.g., understanding the active sites of receptors, rather than enzymes) would also help.

Most of the publications and presentations on FBDD are remarkably positive but many are produced by advocates of the technology or by authors with a business bias. I therefore made efforts to find a skeptic who could give me a completely opposite viewpoint. Interestingly, I could not find one. What I could find was scientists who believe that FBDD is just one of many technologies that may help with drug discovery and development. Mark Murcko of Vertex, for example, says that no single technology in isolation is ever enough because drug discovery and development is horribly complex and can fail for highly diverse reasons. He also rightly pointed out that whenever we look at two X-ray structures of two different ligands bound in somewhat different ways to the same active site, we ask, "What if we combined this part of this molecule with that part of that molecule?" This happens all the time and it could be called "fragment-based design". Another correspondent reminds me that medicinal chemists in big pharma also fragment "big, fat and ugly" hits (by visual inspection) to initiate FBDD thinking.

A number of "success stories" have been published. In 2007 Hajduk and Greer [17] tabulated about 50 potent inhibitors attributable to FBDD, several of them in clinical development, developed by 13 organizations. More recently, 23 published FBDD studies have been analyzed [20], but as yet there is no "FBDD drug" on the market. In an article entitled "the timetable of invention", Baines [21] talks of proof of relevance, proof of concept, and proof of technology each taking 7 years. He thus expects a new technology to take two decades to become established, with "an initial harbinger of success appearing around year 14 when the one product example that 'makes it', and thus proves proof of concept in man, is launched". Many FBDD enthusiasts are looking for vindication by 2011, the earliest that a drug developed from FBDD techniques could be marketed, assuming the typical 15 year drug discovery pipeline. Coincidentally the sequel to this year's RSC meeting will be "Fragments 2011". It will be interesting to see whether FBDD is well established as a part of standard medicinal chemistry best practice by then.

Acknowledgments I am grateful to Rod Hubbard of Vernalis and Mark Murko of Vertex for giving balanced answers to my many challenging questions and to David Bailey of IOTA Pharmaceuticals [22] for helping me with Table 1. I would also like to thank several people who have given me ideas but wished to remain anonymous.

References

1. Hilpert K, Ackermann J, Banner DW, Gast A, Gubernator K, Hadvary P, Labler L, Mueller K, Schmid G, Tschopp TB, van de Waterbeemd H (1994) J Med Chem 37:3889
2. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW (1996) Science 274:1531
3. Allen KN, Bellamacina CR, Ding X, Jeffery CJ, Mattos C, Petsko GA, Ringe D (1996) J Phys Chem 100(7):2605
4. Hann MM, Leach AR, Harper G (2001) J Chem Inf Comput Sci 41:856
5. Fragments 2009 conference organized by the Royal Society of Chemistry Biological and Medicinal Chemistry Group at AstraZeneca, Cheshire, England, March 4–5, 2009
6. Kuntz ID, Chen K, Sharp K, Kollman P (1999) PNAS 96:9997
7. Hopkins AL, Groom CR, Alex A (2004) Drug Discovery Today 9:430
8. Leeson PD, Springthorpe B (2007) Nat Rev Drug Discovery 6:881
9. Congreve M, Chessari G, Tisi D, Woodhead AJ (2008) J Med Chem 51:3661
10. Abad-Zapatero C, Metz JT (2005) Drug Discovery Today 10(7):464
11. Congreve M, Carr R, Murray C, Jhoti H (2003) Drug Discovery Today 8(19):867
12. Hubbard RE, Chen L, Davis B, Drysdale MJ (2007) Curr Top Med Chem 7:1568
13. Everts S (2008) Chem Eng News 86(29):15
14. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH (2005) Nature 435(7042):677
15. Petros AM, Dinges J, Augeri DJ, Baumeister SA, Betebenner DA, Bures MG, Elmore SW, Hajduk PJ, Joseph MK, Landis SK, Nettesheim DG, Rosenberg SH, Shen W, Thomas S, Wang X, Zanze I, Zhang H, Fesik SW (2006) J Med Chem 49:656
16. Vanwetswinkel S, Heetebrij RJ, van Duynhoven J, Hollander JG, Fillipov DV, Hajduk PJ, Siegal G (2005) Chem Biol 12(2):207
17. Hajduk PJ, Greer J (2007) Nat Rev Drug Discovery 6:211
18. Carlson HA, Smith RD, Khazanov NA, Kirchhoff PD, Dunbar JB, Benson ML (2008) J Med Chem 51:6432
19. Hubbard RE, Chen I, Davis B (2007) Curr Opin Drug Discovery Dev 10(3):289
20. de Kloe GE, Bailey D, Leurs R, de Esch IJP (2009) Drug Discovery Today doi:10.1016/j.drudis.2009.03.009 (in press)
21. Baines W (2006) Drug Discovery World. Summer:28
22. The Drug Design Resource sponsored by IOTA Pharmaceuticals http://drugdesignresource.com/index.php. Accessed April 2009