A decade of fragment-based drug design: strategic advances and lessons learned

Philip J. Hajduk and Jonathan Greer

Abstract | Since the early 1990s, several technological and scientific advances — such as combinatorial chemistry, high-throughput screening and the sequencing of the human genome — have been heralded as remedies to the problems facing the pharmaceutical industry. The use of these technologies in some form is now well established at most pharmaceutical companies; however, the return on investment in terms of marketed products has not met expectations. Fragment-based drug design is another tool for drug discovery that has emerged in the past decade. Here, we describe the development and evolution of fragment-based drug design, analyse the role that this approach can have in combination with other discovery technologies and highlight the impact that fragment-based methods have made in progressing new medicines into the clinic.

Forward and reverse genetics
Forward genetics approaches involve proceeding from phenotype to genotype by positional cloning or candidate-gene analysis. Reverse genetics approaches involve proceeding from genotype to phenotype through gene-manipulation techniques.

New chemical entity
A medication containing an active ingredient that has not been previously approved for marketing in any form.

Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA.
Correspondence to P.J.H.
e-mail: philip.hajduk@abbott.com
doi:10.1038/nrd2220
Published online
9 February 2007

Success in the target-based approach to drug discovery, which is currently predominant in pharmaceutical research, demands that two key issues are resolved. First, a biological target(s) with an activity that has a causative role in either the onset or progression of a human disease must be identified. Second, therapeutic agents that appropriately modulate the activity of this target(s) in humans with limited or no adverse effects must be developed.

To address these issues, several new technologies have been developed that promise to markedly improve the quality and efficiency of the drug discovery process. For the identification of new targets, various tools — including genome-wide forward and reverse genetics screens, small-interfering RNAs (siRNAs), and gene expression profiling — can be used to elucidate both the role of certain proteins in disease onset and progression and the biological consequences (both on-target and off-target) of a potential therapeutic agent[1–5]. To improve the chances of finding agents that are active against these targets, technologies such as combinatorial chemistry and ultra-high-throughput screening (HTS) approaches have considerably expanded the numbers of compounds that can be evaluated for their biological activity[6,7]. In addition, the use of structural information in virtual ligand screening[8,9] and structure-based drug design[10] can potentially reduce the numbers of compounds that need to be evaluated, and lead to new directions for synthetic optimization. Novel medicines have been developed, at least in part, on the basis of these approaches (for example, HIV protease inhibitors and imatinib (Gleevec; Novartis), among others). However, the investment in these technologies has not yet reversed the downward trend in the number of new chemical entities reaching the market[11]. So, although a strong case can be made that it is still too early to fully evaluate the productivity of some of these approaches (especially as 10–15 years is generally required for the discovery and development of new drugs)[6], there continues to be a need for new approaches to rapidly develop small-molecule agents against a wide range of therapeutic targets.

It is in this context of intense commercial pressure to increase discovery productivity that we must evaluate the advent of fragment-based drug design — a drug discovery strategy that was first demonstrated a decade ago. Paradoxically, fragment-based drug design is based on screening smaller numbers of compounds (typically several thousand) in the hopes of finding low-affinity fragments (with K_d values in the high micromolar to millimolar range). By contrast, conventional screening attempts to evaluate as many compounds as technologically possible (typically a million or more) in the hopes of finding relatively potent drug leads (with K_d values ideally less than 1 μM). In this article we will describe the initial development of fragment-based drug design and how it has been applied since its introduction. In addition, we will outline key strategic advances that enable this technology to be applied to an increasing number of therapeutic targets. We will also

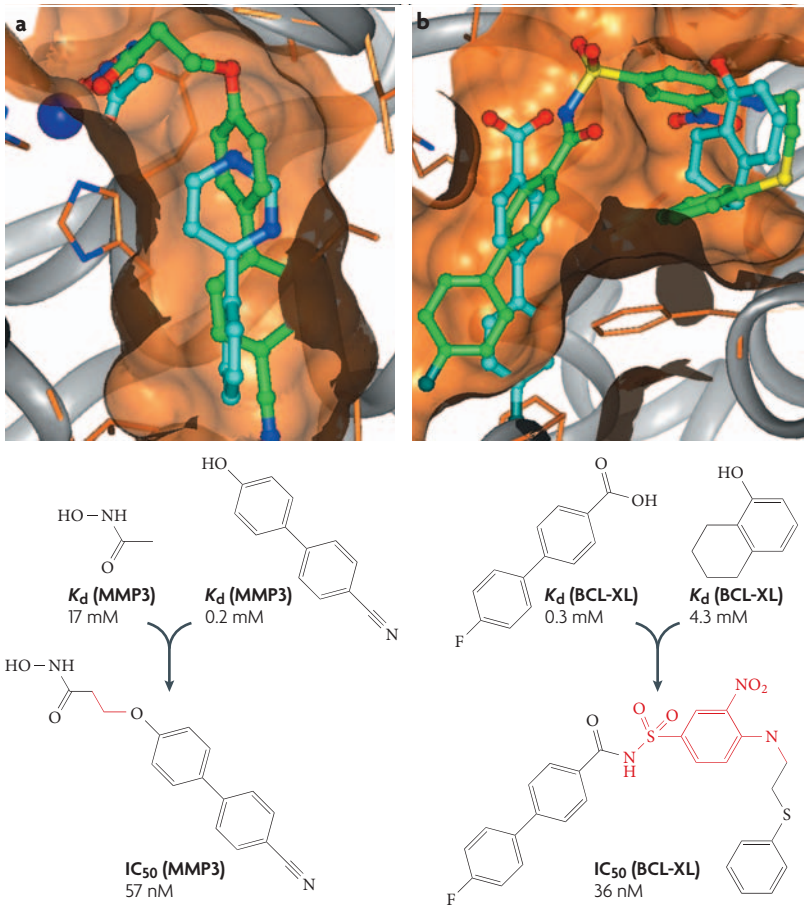

Figure 1 | **Applications of the 'SAR by NMR' method for fragment-based design.** The examples shown are matrix metalloproteinase 3 (MMP3) (**a**) and BCL-XL proteins (**b**). In each case at the top, the identified fragment leads are shown with cyan carbons, whereas the linked compounds are denoted with green carbon atoms. All structures were experimentally determined by NMR. The chemical structures (and *in vitro* potencies) of the fragment leads and subsequent high-affinity linked compounds are shown in the lower part of the figure. NMR, nuclear magnetic resonance; SAR, structure–activity relationships.

discuss lessons derived from critical comparative analyses of fragment-based screening and fragment optimization that can guide the optimal use of fragment-based drug design in a drug discovery setting, and consider its potential to increase R&D productivity.

Fragment-based drug design

Working with fragments. The fragmentation of drug leads into smaller pieces, or even into discrete functional groups (for example, carboxylate, amine, aryl group and so on), has been used for some time to simplify the computational analysis of ligand binding and to map out different pharmacophoric elements required for high-affinity binding[12,13]. The concept of this approach is simple in that proper optimization of each unique interaction in the binding site and subsequent incorporation into a single molecular entity should produce a compound with a binding affinity that is the sum of the individual interactions. However, even the most advanced computational algorithms cannot accurately predict the affinity with which fragments might or might not bind to the protein surface. So, in order to effectively use fragments in drug design, an experimental method was required that could rapidly and reliably screen thousands of low-molecular-mass test compounds for weak (millimolar range) binding to the target protein. Two-dimensional, isotope-edited nuclear magnetic resonance (NMR) spectroscopy is well suited to this purpose as NMR chemical shifts are exquisitely sensitive to ligand binding and problems with compound interference can be solved by spectral editing[14].

Nevertheless, there was a significant amount of internal resistance to resourcing the experimental pursuit of fragment leads at Abbott, as it was commonly believed that such low-molecular-mass, low-affinity ligands, even if they could be detected, would not form a unique and stable complex with the protein that could be productively used in drug design. However, our initial results with FK506-binding protein (FKBP)[15] and matrix metalloproteinases (MMPs)[16] indicated otherwise, in that meaningful structure–activity relationships (SAR) and stable binding modes could be observed even with millimolar ligands for the protein — as long as the fragments were soluble at the test concentrations. It was partly in response to this internal debate that the method was coined 'SAR by NMR' to emphasize that SAR could be obtained by NMR even for weakly binding ligands.

Reduction to practice. Our first internal application of SAR by NMR was in the design of high-affinity inhibitors of the MMPs[16], a family of zinc-dependent endopeptidases that are implicated in various diseases, including arthritis and tumour metastasis[17]. We were initially interested in targeting MMP3 (stromelysin), and our attempts to identify non-peptide inhibitor leads against this protein using a conventional high-throughput activity screen failed. We performed a fragment screen against MMP3 and discovered that acetohydroxamate (a zinc-chelating moiety with a K_d value of 11 mM for the protein) could bind to the protein simultaneously with a number of biaryl compounds (with K_d values in the 20–100 μM range). The three-dimensional structure of a ternary complex (FIG. 1a) clearly revealed that these two fragments could be linked. In fact, one of the first linked compounds showed an IC_{50} value of 57 nM against stromelysin and bound to the protein as designed[16] (FIG. 1a). Lead optimization then began in earnest in order to improve the oral bioavailability of the series and to redirect potency against MMP2 and MMP9. These efforts culminated in ABT-518, which showed excellent oral antitumour efficacy in animal trials and was approved for Phase I clinical trials[18]. Our most recent application of SAR by NMR is in the development of inhibitors against the BCL-2 family of proteins[19]. BCL-2 family members have both pro- and anti-apoptotic activity, and many cancer cells overexpress the anti-apoptotic family members BCL-2 and BCL-XL to evade programmed cell death[20]. Similar to our efforts with MMP3, our initial attempts at conventional HTS against BCL-XL failed to yield productive leads. However, a fragment-based screen again revealed that small organic molecules (in this case biaryl carboxylates and tetrahydronaphthols;

K_d values
The equilibrium dissociation constant of a compound that reflects the concentration needed to reach half-maximal saturation of binding sites. K_d reflects the strength of binding of a compound to its specific binding site.

Pharmacophore
The ensemble of steric and electronic features that is necessary to ensure optimal interactions with a specific biological target structure and to trigger (or to block) its biological response.

FIG. 1b) could occupy proximal binding sites on the protein surface. The structure of the ternary complex clearly revealed that a single molecule could be designed that spanned both sites, and medicinal chemistry optimization[21] ultimately yielded ABT-737, which also shows potent antitumour effects in animal models[19].

Expansion of fragment-based screening. Over the past 10 years, we have used fragment-based drug design in the development of 14 highly potent (IC_{50} < 100 nM) inhibitor series against various protein targets (TABLE 1). During this time, the popularity of fragment-based screening has grown at a remarkable rate in both industry and academic institutions. Shown in FIG. 2 are the numbers of pharmaceutical companies and academic groups that have publicly documented (through publications or conference materials) the incorporation of experimentally driven fragment-based drug design into their discovery science. As of 2006, 23 companies use fragment-based screening as a part (if not the entirety) of their lead generation campaigns. Initially, two-dimensional, isotope-edited NMR spectra were used for detecting ligand binding, and several companies have adopted this approach, including AstraZeneca, Schering–Plough, and Aventis (now Sanofi–Aventis). However, many companies quickly developed alternative NMR-based approaches that obviated the need for isotope labelling and facilitated screening on larger numbers of targets[22–26]. Leaders in these approaches have been Novartis[27], Vertex[28] and Pharmacia (now Pfizer)[29,30].

In 2000, Abbott published its first report on the use of X-ray crystallography for the detection and utilization of fragment leads[31]. Since then, several companies (most notably Astex Therapeutics, SGX Pharmaceuticals and Plexxikon) have made crystallographic fragment screening a key component of their discovery efforts[32]. Tethering[33], developed at Sunesis, is yet another approach to identify fragment hits that has shown great promise for the development of potent drug leads. Overall, the industry has shown great versatility and resourcefulness in rapidly adopting fragment-based screening and developing innovative approaches to extend its applicability, leading to a host of new inhibitors for various protein targets. As shown in TABLE 1, 13 different institutions have reported the successful development of more than 49 potent (IC_{50} < 100 nM) inhibitors against diverse protein targets starting from weakly binding fragments.

Fragment-based screening: why it works

Accessing chemical diversity. Although low-molecular-mass, low-affinity fragments could be identified, a more important question is why one would pursue such an avenue in the first place — especially given the substantial resources concurrently devoted to building up combinatorial chemistry and HTS capabilities throughout the pharmaceutical industry. The answer is straightforward — especially when viewed from the perspective of trying to identify novel chemical matter for novel protein targets. Given that estimates for the size of the chemical universe are in the vicinity of 10^{60} compounds[34], screening 10^6 compounds (representing a fairly large corporate repository) barely scratches the surface of available chemical space. Another problem is the fact that substantial fractions of most corporate repositories are filled with compounds that have been optimized for historical targets, further reducing the chemical diversity in the library and decreasing the chances of finding novel leads. Indeed, in order to address this problem, many pharmaceutical companies have begun to invest large sums of money to increase both the sizes and the diversities of their chemical collections[35,36].

The fragment universe is many orders of magnitude smaller, with one estimate for the size of the chemical universe below 160 Da being ~14 million compounds[37]. So, screening a fragment library of 10,000 compounds captures substantially more chemical diversity space than a conventional high-throughput screen. An additional factor working in favour of fragment-based screening is that less complex molecules should show higher hit rates against protein targets, as theoretically formulated by Hann and co-workers[38]. As a result, even though a typical fragment screen will only explore much less than 1% of the available low-molecular-mass universe, the ability to find leads is substantially higher and further increases the value of the screen. This theoretical model has been recently validated by the Novartis group[39], in which the observed hit rates for fragment screens were 10–1,000 times higher than conventional high-throughput screens. This conclusion is consistent with observations throughout the fragment-screening community. It is important to note that the increased hit rates in fragment-based screens versus high-throughput screens are not simply due to the higher concentrations of compounds used in the fragment screens, as the fragments are substantially smaller than most compounds in typical HTS collections and are expected to bind more weakly. Rather, these higher hit rates reflect a greater exploration of chemical diversity space.

Profiling protein druggability. The combination of broader sampling of the potential chemical universe and increased hit rates for molecules of low complexity makes fragment-based screening a powerful tool for lead generation. In fact, a recent analysis[40] indicated that a sufficiently large fragment screen provides excellent coverage of clinically useful chemical space, as the ability (or inability) to find fragment leads against a protein target correlates well with our ultimate success in producing potent small-molecule modulators of protein activity (FIG. 3). This is in contrast to HTS of even very large chemical repositories, for which substantial increases in the library size can provide new leads for previously intractable targets[6]. So, a fragment screen provides a rapid and reliable means of interrogating a protein target for druggability before investing in further discovery research.

Comparison to HTS. In the past 10 years, Abbott has conducted fragment-based screens against more than 50 protein targets, and we can begin to investigate the hypothesis that fragment screens should capture larger

Two-dimensional, isotope-edited nuclear magnetic resonance (NMR) spectroscopy
NMR experiments that exploit nuclear coupling to correlate the chemical shifts of protons with other NMR-active nuclei, most often carbon-13 or nitrogen-15.

Structure–activity relationships
Correlations that are constructed between the features of chemical structure in a set of candidate compounds and parameters of biological activity, such as potency, selectivity and toxicity.

IC_{50} value
The half maximal inhibitory concentration. Represents the concentration of an inhibitor that is required for 50% inhibition of a biological or molecular process.

Druggability
The ability of a target to be modulated by a lead candidate that has the requisite physicochemical and absorption, distribution, metabolism and excretion properties for development as a drug candidate.

Table 1 | **Potent inhibitors (IC_{50} < 100 nM) derived from experimentally driven fragment-based screening and design**

Company	Target	Endpoint
Abbott	Matrix metalloproteinase (MMP)[18]	Phase I (ABT-518)
Abbott	B-cell CLL/lymphoma 2 (BCL-2), BCL-2-like 1 (BCL-XL)[19]	Preclinical development (ABT-737)
Abbott	FK506-binding protein (FKBP)[15]	Novel, potent inhibitors
Abbott	Leukocyte function-associated antigen-1 (LFA)[59]	Novel, potent inhibitors
Abbott	Protein tyrosine phosphatase-1B (PTP1B)[60]	Novel, potent inhibitors
Abbott	Dihydroneopterin aldolase (DHNA)[61]	Novel, potent inhibitors
Abbott	BCL-2 selective	Novel, potent inhibitors
Abbott	Heat shock protein-90 (HSP90)	Novel, potent inhibitors
Abbott	Survivin	Novel, potent inhibitors
Abbott	Poly (ADP-ribose) polymerase (PARP)	Novel, potent inhibitors
Abbott	Methionine aminopeptidase-2 (MetAP2)	Novel, potent inhibitors
Abbott	Casein kinase-2 (CK2)	Novel, potent inhibitors
Abbott	Kinase insert domain receptor (KDR)	Novel, potent inhibitors
Abbott	v-akt murine thymoma viral oncogene homolog-1 (AKT-1)	Novel, potent inhibitors
Astex Therapeutics	Aurora kinase[62]	Phase I (AT9283)
Astex Therapeutics	Cyclin-dependent kinase (CDK)[62]	Phase I (AT7519)
Astex Therapeutics	CDK[62]	Preclinical development (AT9311)
Astex Therapeutics	HSP90[62]	Preclinical development (AT13387)
Astex Therapeutics	Mitogen-activated protein kinase-14 (P38α)[63]	Novel, potent inhibitors
Aventis	Src SH2 domain[64,65]	Novel, potent inhibitors
Burnham Institute	Anthrax lethal factor[66]	Novel, potent inhibitors
Novartis	3α-hydroxysteroid dehydrogenase (3α-HSD)	Novel, potent inhibitors
Plexxikon	Peroxisome proliferator-activated receptor (PPAR) (Metabolic)	Phase II (PPM204)
Plexxikon	Oncogenic v-raf murine sarcoma viral oncogene homolog B1 (B-Raf)	Investigational new drug application (PLX4032)
Plexxikon	PPAR multiple sclerosis	Preclinical development
Plexxikon	FMS/KIT	Preclinical development
Plexxikon	Phosphodiesterase[67]	Preclinical development
Roche	DNA Gyrase[68]	Novel, potent inhibitors
SGX Pharmaceuticals	SYK[69]	Novel, potent inhibitors
SGX Pharmaceuticals	Aurora kinase[70]	Novel, potent inhibitors
Schering–Plough	β-site APP-cleaving enzyme-1 (BACE1)[71]	Preclinical development
Schering–Plough	Mouse double minute-2 (MDM2)	Novel, potent inhibitors
Schering–Plough	AKT-1[72]	Novel, potent inhibitors
Schering–Plough	Hepatitis C virus (HCV) polymerase	Novel, potent inhibitors
Sunesis	Aurora kinase[73]	Preclinical development (SNS-314)
Sunesis	Interleukin-2 (IL2)[74]	Novel, potent inhibitors
Sunesis	Caspase-3[75]	Novel, potent inhibitors
Sunesis	Caspase-1[76]	Novel, potent inhibitors
Triad	P38α[77]	Preclinical development
Triad	c-Jun N-terminal kinase-2 (JNK2)[77]	Novel, potent inhibitors
Vernalis	HSP90	Preclinical development
Vernalis	Phosphoinositide-dependent protein kinase-1 (PDK1)	Novel, potent inhibitors
Vernalis	Checkpoint kinase-1 (CHK1)	Novel, potent inhibitors
Vernalis	Aurora kinase	Novel, potent inhibitors
Vertex	Jun kinase-3 (JNK3)[78]	Novel, potent inhibitors
Vertex	Adipocyte lipid-binding protein-2 (aP2)[78]	Novel, potent inhibitors
Vertex	Regulatory erythroid kinase (REDK)	Novel, potent inhibitors

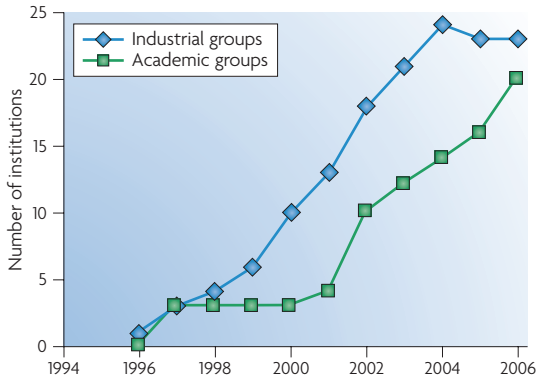

Figure 2 | **Trends in the application of fragment-based screening.** The number of industrial (blue diamonds) and academic groups (green squares) that have used fragment-based screening as part of their lead generation strategies (based on publications and/or conference materials) as a function of time.

numbers of chemically diverse hits than conventional screens of large corporate repositories. Shown in FIG. 4 are data for screening and subsequent discovery research on a set of 45 Abbott targets that underwent both fragment screening and conventional HTS. Fragment-based screens were successful in identifying interesting, chemically tractable hits for 76% of these targets (FIG. 4a), compared with 53% for HTS ($p < 0.01$), validating the premise that fragment screening can deliver more hits against larger numbers of protein targets. Importantly, lead optimization (FIG. 4b) was initiated on fragment leads for 42% of these targets, ultimately leading to potent ($IC_{50} < 100$ nM) inhibitors for 31% of the proteins (FIG. 4c). This can be compared with lead optimization on HTS hits for 33% of these same targets, yielding potent inhibitors for 26% of the proteins.

It is significant to note the highly complementary nature of the two lead generation approaches, in which lead optimization was pursued on compounds from both HTS and fragment-based screening for 20% of the targets, resulting in potent inhibitors from both sources for 15% of the targets. In addition, potent leads (FIG. 4c) were obtained for an additional 13% (6 out of 45) of the target set by using both HTS and fragment-based screening as compared with using either screening technology alone. In some cases, the fragment leads were optimized in parallel to yield completely novel, alternative chemotypes, whereas in other cases the fragment leads were incorporated directly into the optimization of the HTS leads.

Another important consideration when comparing HTS and fragment-based screening is the quality of the hits that result from the initial screen. Often, HTS yields hundreds or even thousands of compounds that modulate the biochemical response of an assay with apparent IC_{50} values on the order of 10 μM or less. As many assays used in HTS can be complex, multi-component systems, many of the hits could actually disrupt or inhibit an assay component other than the target of interest. In addition, false positives can result from the method of detection or mechanisms such as compound aggregation[41] or reactivity[42]. This can lead to very low confirmation rates for HTS hits[43] that can complicate or even confound the lead triage process. By contrast, fragment-based screening is less prone to such artefacts as the low-molecular-mass compounds tend to be more soluble and the methods of detection are simpler and more robust. So, although HTS might be the quickest route to generating initial leads for many targets (provided that sufficiently robust secondary assays are in place), fragment-based screening is generally more reliable and provides higher quality chemical matter. This was certainly true in the case of BCL-XL, described above, in which the ultimate clinical candidate was derived from a fragment hit despite the application of multiple HTS assay formats[44].

Harnessing the power of structure-based drug design. It can be observed from FIG. 4b that leads from HTS will often be pursued over fragment leads for the same target (given that leads exclusively from HTS were pursued for 11% of the targets for which fragment leads were also available). This can be expected as the HTS leads are often 2–3 orders of magnitude more potent than the fragment leads and offer, at least at first glance, more attractive starting points for optimization. Of course, these scenarios should be carefully evaluated as the fragment leads could offer more benefit from the view of ligand efficiency[45,46]. Although converting a millimolar ligand into a nanomolar drug lead might seem like a daunting task, the strategic use of structure-based drug design can greatly facilitate this process. As shown in FIG. 5, the ability to obtain NMR or X-ray crystal structures on fragment leads has had a dramatic influence on the success of fragment-based drug design. In fact, the ability to produce potent inhibitors ($IC_{50} < 100$ nM) after initiating lead optimization on fragment leads nearly triples with the aid of structure-based design, increasing from 33% to 93% ($p < 0.01$). This remarkably high rate of success with the aid of structure-based drug design

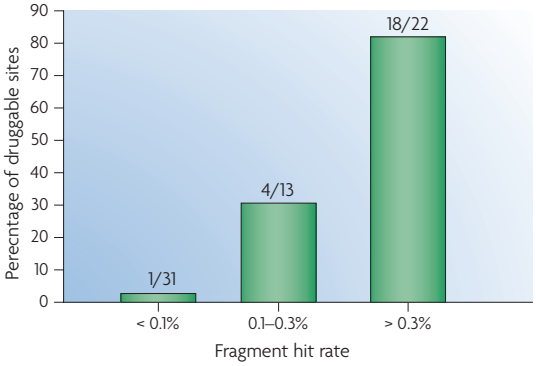

Figure 3 | **Percentage of protein-binding sites that can be targeted with small, drug-like molecules as a function of hit rate from NMR-based fragment screening.** A total of 66 binding sites on 58 protein targets were used in the analysis. A binding site was defined as druggable if potent ($IC_{50} < 300$ nM), non-covalent, small-molecule inhibitors have been reported or internally identified[40]. NMR, nuclear magnetic resonance.

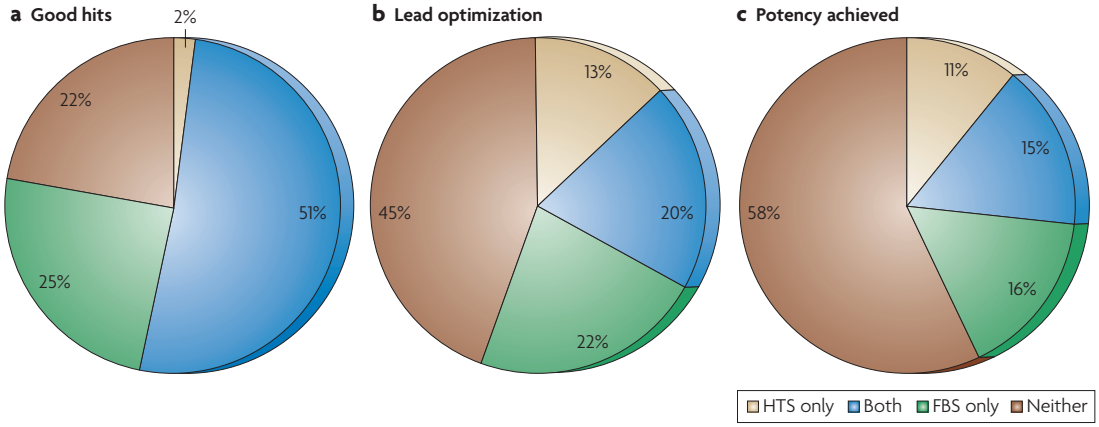

Figure 4 | **Comparing high-throughput screening and fragment-based screening.** This figure shows the percentage of a set of 45 protein targets that underwent both high-throughput screening (HTS) and fragment-based screening (FBS) and reached critical discovery milestones: (**a**) Chemically tractable hits; (**b**) Lead optimization initiated around hits from the specified lead source; (**c**) Potency (defined as IC_{50} < 100 nM) achieved from the specified lead source.

compares well with our ability to translate HTS hits into potent leads, for which a success rate of 88% has been historically achieved after lead optimization begins. So, fragment-based screening coupled with structure-based drug design provides a powerful avenue for maximizing the exploration of chemical diversity space and generating novel, potent inhibitors for various protein targets.

Of course, potency is not the sole criterion for producing a successful drug candidate. Also shown in FIG. 5 is that less than a quarter of the potent ligands that result from an initial fragment screen are viewed as suitable for extensive preclinical development (which is comparable to the fraction of leads developed from HTS hits; data not shown). Issues such as bioavailability, toxicity and target validity continue to have a large effect on attrition rates in drug development.

Fragment selection and optimization. The process of lead optimization typically results in an increase in both the size and lipophilicity of the original hit[47]. Given that both of these properties directly influence oral bioavailability[48], there is much debate in the industry as to what constitutes a 'good' hit — that is, one that will remain rule-of-five compliant after optimization. This is very difficult to assess with hits derived from HTS, as a significant fraction of the molecules could be interacting sub-optimally with the receptor. Although there has been a general trend towards starting with smaller, less lipophilic (and correspondingly weaker binding) compounds[47], there is still no clear consensus as to what constitutes an acceptable lead in terms of potency, molecular mass and hydrophobicity.

The same is true for fragment screening. For example, what is the probability that a fragment lead with a molecular mass of 250 Da and a K_d value of 1 mM will yield a low nanomolar inhibitor that is orally bioavailable? A recent retrospective analysis[49] indicates that, unlike leads from HTS, the optimization of fragment leads can be predicted with high accuracy. In this analysis, 18 highly optimized inhibitors were systematically reduced in size until a fragment-like lead could be identified. As shown in FIG. 6 for a subset of five inhibitors, a remarkably linear relationship exists between potency and molecular mass along this path of ideal optimization. In fact, these relationships are nearly co-linear, with an average value for the slope of 64 over all the series. So, during fragment optimization, an increase of 1 pK_d unit can be expected for every 64 mass units added to the compound. Although these results have significant implications for molecular recognition in general, their most immediate use is the ability to predict the final mass of an optimized inhibitor given a fragment lead of known potency and molecular mass. This places well defined limits on the acceptable size and potency of fragments leads that should be considered for use in fragment-based drug design, and also enables a critical and quantitative assessment of lead identification and optimization in general.

A paradigm shift for drug discovery

Smaller might be better. The growth and success of fragment-based drug design has necessitated a paradigm shift for small-molecule drug discovery — especially for the medicinal chemists whose job it is to produce high-affinity drugs. Whereas before, most chemists would not consider a compound's activity 'interesting' unless it was in the range of 1 μM, chemists at many pharmaceutical companies are now routinely beginning synthetic programmes around small-molecule leads with affinities as weak as 1 mM. Absolute affinity of binding is not as important as relative efficiency of binding[45,46] — especially with the availability of structural information that can facilitate rapid improvements in potency. Of course, leaders in the pharmaceutical industry have independently realized that chasing after potency at the expense of other physicochemical properties (such as lipophilicity, polarity, charge, stability, and so on) carries serious risks of failure owing to inadequate pharmacokinetic properties of the resulting compound[50]. This has spawned a whole new movement (the 'lead-like'

Rule-of-five
The 'rule of five' identifies several key properties that should be considered for compounds with oral delivery in mind. These properties are molecular mass < 500 Da, cLogP < 5, number of hydrogen-bond donors < 5 and number of hydrogen-bond acceptors < 10.

movement) away from the use of large lipophilic compounds as leads towards smaller compounds that will have reasonable chances of possessing good pharmacokinetic properties after the optimization process is complete[47,51]. The successful applications of fragment-based drug design have provided ample support that the use of fragments could, in many cases, be the most direct route to the best achievable balance between potency and pharmacokinetic properties.

Science and serendipity. Although the concept of fragment-based drug design is simple and elegant, there remains the hard work of making a drug and all the pitfalls and surprises along the way. For example, the initial concept of SAR by NMR entailed the linking of two or more fragments that bound to neighbouring pockets on a protein surface. However, simply identifying multiple ligands will not guarantee success. Although the linking process was straightforward for the MMP example described above (FIG. 1a), the process of incorporating two fragments leads into a single ligand for BCL-XL (FIG. 1b) was much more difficult, requiring multiple synthetic strategies in order to identify the proper compound[21]. In addition, it is often the case that multiple fragments that simultaneously occupy the binding pocket cannot be found, prompting the development of an array of strategies for using fragments in drug design[14]. So, the fragment-based drug designer must be flexible and always allow for some degree of serendipity in the design process, as unexpected results often provide new opportunities for further research.

Progressing to the clinic. The science of fragment-based drug design is exciting, but the industry is eager to see this new technology make an impact on human disease. TABLE 1 lists 48 examples of advanced programmes that began with fragment-based screening, many of which are still being actively pursued. Plexxikon has perhaps the most advanced drug candidate derived from fragment-based design, with PPM204 now in Phase II trials for the treatment of metabolic disease. Astex Therapeutics has two compounds that are currently in clinical trials that were derived from fragment-based leads, boasting a mere 18 months from the first synthesis of AT7519 to dosing in patients[52]. In addition, a number of other companies (including Abbott, Schering–Plough, Sunesis and Vernalis) have compounds in advanced preclinical development (TABLE 1). These examples demonstrate that clinically useful compounds can be generated using a fragment-approach (even when other discovery technologies fail), and indicate that fragment-based drug design could deliver on its promise to increase the success rate and decrease the cost of modern drug discovery.

Future directions

Fragment-based drug design has now become an established paradigm at many pharmaceutical companies. As highlighted in this review, fragment-based screening, either alone or in combination with conventional HTS, can successfully deliver clinically useful drug leads against various protein targets. Obtaining structural information on the initial fragment leads complexed to the protein target is a key factor for success and is the primary screening tool at many companies that employ X-ray crystallographic screening.

However, this advantage is also a major limitation to the number and types of target that are amenable to fragment-based approaches in that X-ray crystal structures and/or NMR structural data cannot be obtained for many protein targets. These experimental screening approaches also require large quantities of purified proteins (typically 10–1,000 milligrams), which places

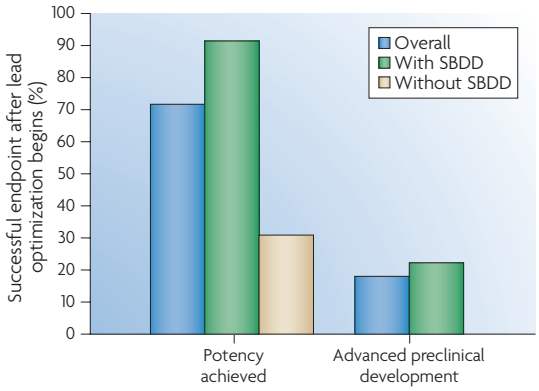

Figure 5 | **Importance of structural information to the success of fragment-based drug design.** This figure illustrates the influence of structure-based drug design (SBDD) on the ability to obtain potent (IC_{50} < 100 nM) inhibitors or compounds suitable for advanced preclinical development after a lead optimization programme was initiated. The analysis includes synthetic optimization programmes on 20 leads from fragment-based screening (14 of which received X-ray- or nuclear magnetic resonance-based structural support).

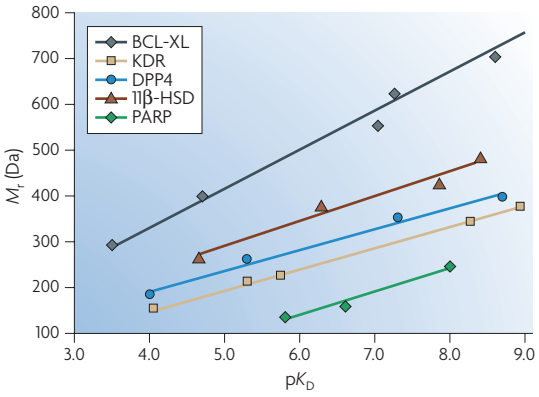

Figure 6 | **Relationship between potency and molecular mass during fragment optimization.** Plots of pK_d values (defined as the negative base-10 logarithm of the K_d value expressed in molar units) versus molecular mass (M_r) for compounds identified in the deconstruction of highly optimized leads for 5 protein targets[49]. Best-fit linear trends are shown for each series and coloured according to the legend. 11β-HSD, 11β-hydroxysteroid dehydrogenase; DPP4, dipeptidyl peptidase 4; KDR, kinase insert domain receptor; PARP, poly (ADP-ribose) polymerase.

another constraint on the types of target that can be screened. These factors underscore the remarkable success that has been achieved with fragment-based drug design against particular target families (especially protein kinases) that can be readily prepared in large amounts and subjected to X-ray structural analysis. However, this will not be the case for many targets. At Abbott, fragment-based methods have only been used on ~30% of all discovery targets owing, in large part, to a lack of recombinant protein in a suitable form for fragment screening.

So, the future growth of fragment-based drug design might depend significantly on our ability to heterologously express recalcitrant proteins in sufficient quantity and quality such that screening and structural studies can be initiated. This is especially true for integral membrane proteins, which have so far proven to be extraordinarily difficult to produce and crystallize. Alternatively, new methods for fragment screening that overcome the need for large quantities of proteins and that are amenable to being applied with membrane proteins could expand the target universe for fragment-based drug design. This could entail alternative applications of existing methodologies[53], novel hardware designs[54] or entirely new concepts, such as the rapidly evolving field of dynamic combinatorial chemistry[55,56]. There has also been tremendous progress in the computational analysis of fragment binding[57], which promises to reduce the dependence on experimental screening and the structure determination of protein–ligand complexes. We hope that these and other advances will continue to enable and expand the application of fragment-based drug design in the development of new drugs.

1. Butcher, E. C. Can cell systems biology rescue drug discovery? *Nature Rev. Drug Discov.* **4**, 461–467 (2005).
2. van der Greef, J. & McBurney, R. N. Innovation: Rescuing drug discovery: *in vivo* systems pathology and systems pharmacology. *Nature Rev. Drug Discov.* **4**, 961–967 (2005).
3. Hardy, L. W. & Peet, N. P. The multiple orthogonal tools approach to define molecular causation in the validation of druggable targets. *Drug Discov. Today* **9**, 117–126 (2004).
4. Betz, U. A., Farquhar, R. & Ziegelbauer, K. Genomics: success or failure to deliver drug targets? *Curr. Opin. Chem. Biol.* **9**, 387–391 (2005).
5. Zambrowicz, B. P. & Sands, A. T. Knockouts model the 100 best-selling drugs — will they model the next 100? *Nature Rev. Drug Discov.* **2**, 38–51 (2003).
6. Macarron, R. Critical review of the role of HTS in drug discovery. *Drug Discov. Today* **11**, 277–279 (2006).
An important analysis of the current and future impact of high-throughput screening on the drug discovery process.
7. Silverman, L., Campbell, R. & Broach, J. R. New assay technologies for high-throughput screening. *Curr. Op. Chem. Biol.* **2**, 397–403 (1998).
8. Oprea, T. I. & Matter, H. Integrating virtual screening in lead discovery. *Curr. Opin. Chem. Biol.* **8**, 349–358 (2004).
9. Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **432**, 862–865 (2004).
10. Scapin, G. Structural biology and drug discovery. *Curr. Pharm. Des.* **12**, 2087–2097 (2006).
11. Brown, D. & Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discov. Today* **8**, 1067–1077 (2003).
12. Bohm, H. J. Site-directed structure generation by fragment-joining. *Persp. Drug. Disc. Design* **3**, 21–33 (1995).
13. Miranker, A. & Karplus, M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **11**, 29–34 (1991).
14. Hajduk, P. J., Meadows, R. P. & Fesik, S. W. NMR-based screening in drug discovery. *Q. Rev. Biophys.* **32**, 211–240 (1999).
15. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
A seminal paper describing the SAR by NMR method.
16. Hajduk, P. J. *et al.* Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. *J. Am. Chem. Soc.* **119**, 5818–5827 (1997).
Application of the SAR by NMR concept to matrix metalloproteinases, ultimately resulting in the first compound designed using a fragment-based approach to reach the clinic (see reference 18 below).
17. Beckett, R. P., Davidson, A. H., Drummond, A. H., Huxley, P. & Whittaker, M. Recent advances in matrix metalloproteinase inhibitor research. *Drug. Disc. Today* **1**, 16–26 (1996).
18. Wada, C. K. *et al.* Phenoxyphenyl sulfone *N*-formylhydroxylamines (Retorhydroxamates) as potent, selective, orally bioavailable matrix metalloproteinase inhibitors. *J. Med. Chem.* **45**, 219–232 (2002).
19. Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677–681 (2005).
The most recent example of the SAR by NMR approach applied to a protein–protein interaction target, ultimately resulting in a compound that is currently in Phase I clinical trials for the treatment of cancer.
20. Kirkin, V., Joos, S. & Zornig, M. The role of Bcl-2 family members in tumorigenesis. *Biochim. Biophys. Acta* **1644**, 229–249 (2004).
21. Petros, A. M. *et al.* Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. *J. Med. Chem.* **49**, 656–663 (2006).
22. Lepre, C. A., Moore, J. M. & Peng, J. W. Theory and applications of NMR-based screening in pharmaceutical research. *Chem. Rev.* **104**, 3641–3676 (2004).
23. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed. Engl.* **42**, 864–890 (2003).
An excellent overview of the numerous applications of NMR in fragment-based screening.
24. Dalvit, C. *et al.* A general NMR method for rapid, efficient, and reliable biochemical screening. *J. Am. Chem. Soc.* **125**, 14620–14625 (2003).
25. London, R. E. Theoretical analysis of the inter-ligand overhauser effect: a new approach for mapping structural relationships of macromolecular ligands. *J. Magn. Reson.* **141**, 301–311 (1999).
26. Becattini, B. *et al.* Structure-activity relationships by interligand NOE-based design and synthesis of antiapoptotic compounds targeting Bid. *Proc. Natl Acad. Sci. USA* **103**, 12602–12606 (2006).
27. Lin, M., Shapiro, M. J. & Wareing, J. R. Diffusion-edited NMR-affinity NMR for direct observation of molecular interactions. *J. Am. Chem. Soc.* **119**, 5249–5250 (1997).
28. Fejzo, J. *et al.* The SHAPES strategy: An NMR-based approach for lead generation in drug discovery. *Chem. Biol.* **6**, 755–769 (1999).
29. Stockman, B. J. NMR spectroscopy as a tool for structure-based drug design. *Prog. Nucl. Magn. Reson. Spectrosc.* **33**, 109–151 (1998).
30. Dalvit, C. *et al.* NMR-based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. *J. Med. Chem.* **45**, 2610–2614 (2002).
31. Nienaber, V. L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnol.* **18**, 1105–1108 (2000).
32. Carr, R. A., Congreve, M., Murray, C. W. & Rees, D. C. Fragment-based lead discovery: leads by design. *Drug Discov. Today* **10**, 987–992 (2005).
A concise description of the theory, advantages and applications of fragment-based drug design.
33. Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
34. Martin, Y. C. Challenges and prospects for computational aids to molecular diversity. *Perspect. Drug Discov. Des.* **7–8**, 159–172 (1997).
35. Jacoby, E. *et al.* Key aspects of the Novartis compound collection enhancement project for the compilation of a comprehensive chemogenomics drug discovery screening collection. *Curr. Top. Med. Chem.* **5**, 397–411 (2005).
36. Davis, A. M., Keeling, D. J., Steele, J., Tomkinson, N. P. & Tinker, A. C. Components of successful lead generation. *Curr. Top. Med. Chem.* **5**, 421–439 (2005).
37. Fink, T., Bruggesser, H. & Reymond, J. L. Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew. Chem. Int. Ed. Engl.* **44**, 1504–1508 (2005).
38. 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).
The authors provide a sound theoretical basis for the success of fragment-based screening based on molecular complexity and basic principles of molecular recognition.
39. Schuffenhauer, A. *et al.* Library design for fragment based screening. *Curr. Top. Med. Chem.* **5**, 751–762 (2005).
40. Hajduk, P. J., Huth, J. R. & Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* **48**, 2518–2525 (2005).
An analysis that establishes the utility of using fragment-based screening not only for lead identification, but also for characterizing the druggability of protein targets with small-molecule ligands. A computational analysis of protein druggability is also described.
41. McGovern, S. L., Caselli, E., Grigorieff, N. & Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **45**, 1712–1722 (2002).
42. Huth, J. R. *et al.* ALARM NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens. *J. Am. Chem. Soc.* **127**, 217–224 (2005).
43. Hajduk, P. J. & Burns, D. J. Integration of NMR and high-throughput screening. *Comb. Chem. High Throughput Screen.* **5**, 613–621 (2002).
44. Qian, J. *et al.* Discovery of novel inhibitors of Bcl-xL using multiple high-throughput screening platforms. *Anal. Biochem.* **328**, 131–138 (2004).
45. Abad-Zapatero, C. & Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discov. Today* **10**, 464–469 (2005).
46. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9**, 430–431 (2004).
This work (together with reference 45) provides some benchmarks for evaluating the usefulness of drug leads based on relationships between potency and mass.

47. Oprea, T. I., Davis, A. M., Teague, S. J. & Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci*. **41**, 1308–1315 (2001).
An important analysis of the process of lead optimization, in which quantitative measures of the changes in mass, hydrophobicity and other physicochemical properties that occur during synthetic optimization are given.
48. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **44**, 235–249 (2000).
A landmark paper correlating oral bioavailability with various physicochemical properties such as mass and hydrophobicity. The guidelines described in this paper are now used throughout the pharmaceutical industry.
49. Hajduk, P. J. Fragment-based drug design: how big is too big? *J. Med. Chem*. **49**, 6972–6976 (2006).
A retrospective analysis of the process of fragment optimization, in which it is demonstrated that the ultimate ligand efficiency of an optimized inhibitor is dictated by the fragment core, thereby enabling a quantitative evaluation of lead selection and optimization.
50. Lipinski, C., Lombardo, F., Dominy, B. & Feeney, P. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev*. **23**, 3–25 (1997).
51. Wunberg, T. *et al*. Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits. *Drug Discov. Today* **11**, 175–180 (2006).
52. Wolfson, W. Fragmentary solutions Astex Therapeutics puts the pieces together. *Chem. Biol*. **13**, 799–801 (2006).
53. Barker, J., Courtney, S., Hesterkamp, T., Ullmann, D. & Whittaker, M. Fragment screening by biochemical assay. *Expert Opin. Drug Discov*. **1**, 225–236 (2006).
54. Vanwetswinkel, S. *et al*. TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery. *Chem. Biol*. **12**, 207–216 (2005).
55. Lehn, J. M. & Eliseev, A. V. Dynamic combinatorial chemistry. *Science* **291**, 2331–2332 (2001).
56. Otto, S., Furlan, R. L. & Sanders, J. K. Recent developments in dynamic combinatorial chemistry. *Curr. Opin. Chem. Biol*. **6**, 321–327 (2002).
57. Vajda, S. & Guarnieri, F. Characterization of protein–ligand interaction sites using experimental and computational methods. *Curr. Opin. Drug Discov. Devel*. **9**, 354–362 (2006).
58. Hajduk, P. J. *et al*. Design of adenosine kinase inhibitors from the NMR-based screening of fragments. *J. Med. Chem*. **43**, 4781–4786 (2000).
59. Liu, G. *et al*. Novel *p*-arylthio cinnamides as antagonists of leukocyte function-associated antigen-1/intracellular adhesion molecule-1 interaction. 2. Mechanism of inhibition and structure-based improvement of pharmaceutical properties. *J. Med. Chem*. **44**, 1202–1210 (2001).
60. Szczepankiewicz, B. G. *et al*. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J. Am. Chem. Soc*. **125**, 4087–4096 (2003).
61. Sanders, W. J. *et al*. Discovery of potent inhibitors of dihydroneopterin aldolase using CrystaLEAD high-throughput X-ray crystallographic screening and structure-directed lead optimization. *J. Med. Chem*. **47**, 1709–1718 (2004).
62. Astex Therapeutics. Clincal candidates: AT13387. *Astex Therapeutics web site* [online], <http://www.astex-therapeutics.com/products/clinicalcandidates.php#AT13387>.
63. Gill, A. L. *et al*. Identification of novel p38α MAP kinase inhibitors using fragment-based lead generation. *J. Med. Chem*. **48**, 414–426 (2005).
64. Lange, G. *et al*. Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of (pp60)Src are identical to those for high affinity binding of full length inhibitors. *J. Med. Chem*. **46**, 5184–5195 (2003).
65. Lesuisse, D. *et al*. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: application to the discovery of nanomolar inhibitors of Src SH2. *J. Med. Chem*. **45**, 2379–2387 (2002).
66. Forino, M. *et al*. Efficient synthetic inhibitors of anthrax lethal factor. *Proc. Natl Acad. Sci. USA* **102**, 9499–9504 (2005).
67. Card, G. L. *et al*. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nature Biotech*. **23**, 201–207 (2005).
68. Oblak, M. *et al*. In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. *Bioorg. Med. Chem. Lett*. **15**, 5207–5210 (2005).
69. Blaney, J., Nienaber, V. L. & Burley, S. in *Fragment-Based Approaches in Drug Discovery* (eds Jahnke, W. & Erlanson, D. A.) 215–248 (Wiley–VCH, Weinheim, Germany, 2006).
70. Warner, S. L. *et al*. Identification of a lead small-molecule inhibitor of the Aurora kinases using a structure-assisted, fragment-based approach. *Mol. Cancer Ther*. **5**, 1764–1773 (2006).
71. Zhu, Z. *et al*. Heterocyclic aspartyl protease inhibitors. US Patent 20060111370 (2006).
72. Chan, T-Y. *et al*. Kinase inhibitors. European Patent WO2006081230 (2006).
73. Sunesis. Pipeline programs: SNS-314. *Sunesis web site* [online], <http://www.sunesis.com/pipeline/programs.php>.
74. Braisted, A. C. *et al*. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J. Am. Chem. Soc*. **125**, 3714–3715 (2003).
75. Choong, I. C. *et al*. Identification of potent and selective small-molecule inhibitors of caspase-3 through the use of extended tethering and structure-based drug design. *J. Med. Chem*. **45**, 5005–5022 (2002).
76. O'Brien, T. *et al*. Structural analysis of caspase-1 inhibitors derived from Tethering. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun*. **61**, 451–458 (2005).
77. Sem, D. in *Fragment-Based Approaches in Drug Discovery* (eds Jahnke, W. & Erlanson, D. A.) 149–177 (Wiley–VCH, Weinheim, Germany, 2006).
One of a series of chapters in a book dedicated to the theory, implementation and application of fragment-based drug design.
78. Moore, J. *et al*. Leveraging structural approaches: applications of NMR-based screening and X-ray crystallography for inhibitor design. *J. Synchrotron Radiat*. **11**, 97–100 (2004).

Acknowledgements
The authors would like to thank R. Artis (Plexxikon), M. Congreve (Astex), D. Erlanson (Sunesis), R. Hubbard (Vernalis), W. Jahnke (Novartis), C. Lepre (Vertex), and D. Wyss (Schering–Plough) for help in gathering the data for Table 1.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Gene:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
BCL-2 | BCL-XL | MMP2 | MMP3 | MMP9
Access to this links box is available online.