ELSEVIER

Expanding medicinal chemistry space

Andy Barker, Jason G. Kettle, Thorsten Nowak and J. Elizabeth Pease

AstraZeneca R&D, Oncology iMed, Alderley Park, Macclesfield SK10 4TG, UK

Clinically useful drugs target a relatively small number of proteins that lie within a clearly defined and chemically accessible space. However, many high value biological targets lie outside this chemical space, and an ability to access such 'intractable' targets not amenable to traditional small molecule intervention would expand treatment options and be a major boost for patients and the pharmaceutical industry. To date, success has been limited but new technologies and approaches are beginning to emerge that could provide novel lead generation capabilities that enable access to new drug target classes. We review these new approaches and their ability to provide the novel leads needed to tackle a new generation of biological targets.

Introduction

The mainstay of lead generation for pharmaceutical companies over the past two decades has been, and continues to be, HTS. Continued improvements in screening technologies combined with growing company collections have resulted in many leads for targets that have subsequently been developed into drugs. These compound collections have been productive sources of hits for many drug classes including, for example, kinases and certain classes of G-protein-coupled receptors (GPCRs). They have however been much less successful in providing leads for several important and biologically attractive target classes, the so called intractable/undruggable targets, for example inhibitors of protein–protein interactions and phosphatases. Figure 1 shows the proportion of HTS campaigns that led to successful lead identification programmes at AstraZeneca in the period 2004–2008, together with reasons for devalidation failure of the target or failure to find useful hits. It is clear that, for historically tractable target classes such as ion channels and nuclear hormone receptors, HTS continues to be an effective lead generation strategy. However, success rates remain low for many novel target classes. Even within target classes that often yield to broad screening strategies there are subclasses for which such approaches remain unproductive (for example Class B GPCRs). For a class such as the kinases, where HTS has frequently delivered useful leads, success remains at a modest 50% and, although not all these failures can be attributed to a lack of leads, this broad picture illustrates the challenges faced.

Reflecting, as they do, the isolated history of previous drug discovery projects, there is a growing awareness of the limited structural diversity in compound collections [1]. This, coupled with the limitations of traditional biochemical screening assays in identifying hits, has sparked the development of new lead generation approaches to develop chemistries that provide access to new biological space, thus enabling hits to be identified against many new and highly desirable targets.

Bioactive chemical space

Chemical space is vast. It has been estimated that there are potentially 10^{60} organic molecules with a molecular weight below 500 Da [2]. By contrast, biological space is understood to be relatively modest with approximately 30 000 disease-modifying genes, although as little as 10% might be implicated in human disease states [3]. Uniting synthetically accessible chemical space with disease-relevant biological space is at the heart of all drug discovery efforts, and experience shows that this challenge is a significant one. Simply, either these spaces do not overlap completely or there are parts of the overlapping spaces that have not been populated with appropriate small molecules [4]. The first explanation results in the thesis that some parts of the genome are 'undruggable' and the second account is that current synthetic or design techniques limit access to appropriate, medicinally relevant chemical space. The evidence for the existence of undruggable

Corresponding author:. Kettle, J.G. (jason.kettle@astrazeneca.com)

 1359-6446/06/$ - see front matter http://dx.doi.org/10.1016/j.drudis.2012.10.008

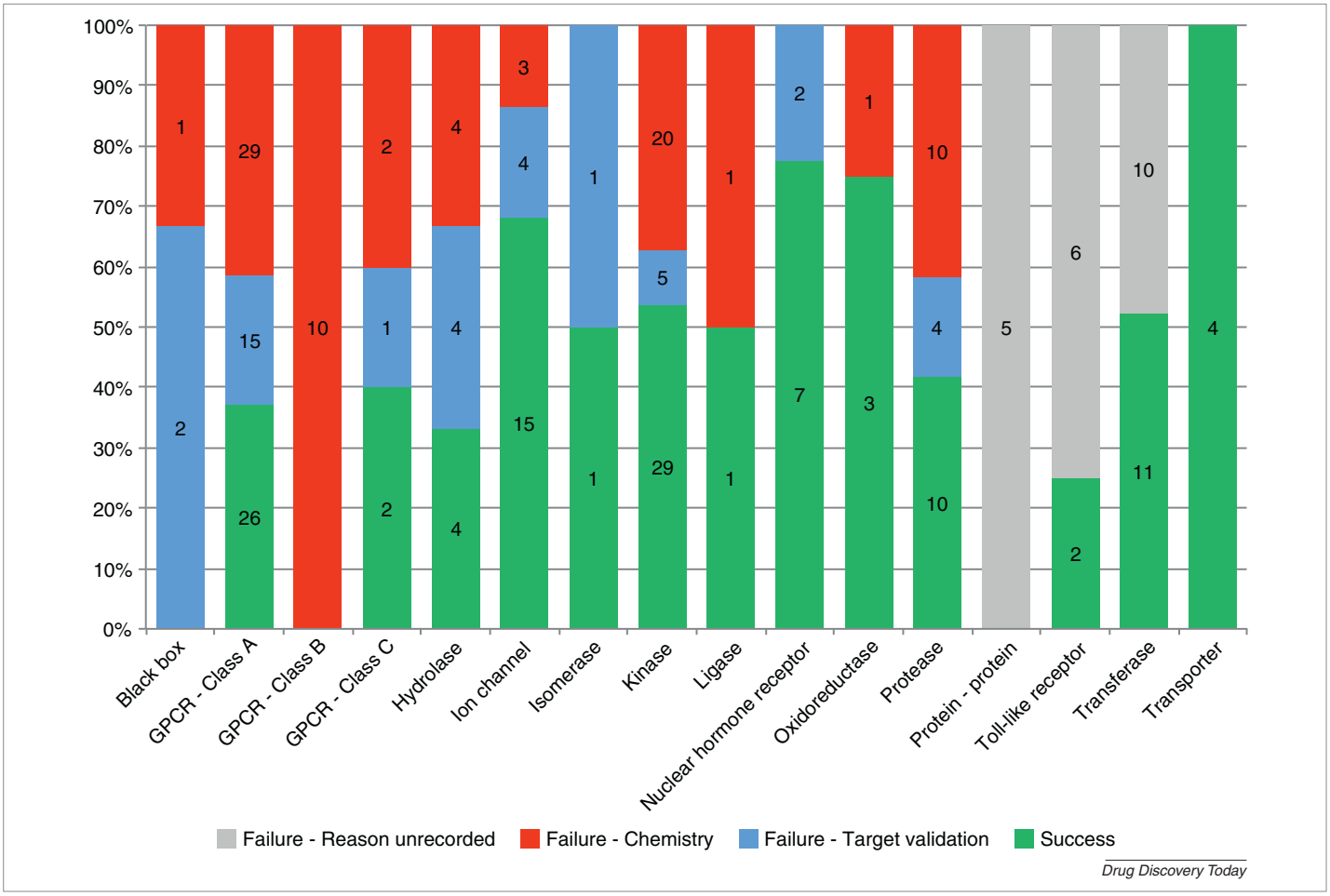

FIGURE 1

HTS screening success rates by target class in AstraZeneca 2004–2008. Success rate is defined as the percentage of screening campaigns that led to a project transitioning into the lead identification phase – typically demonstrating on-target effects in a cellular assay, with evidence of exposure in a rodent species. Unsuccessful screens are broadly categorised according to reason for failure – chemistry, where no tractable hits were found, or target validation. Numbers indicate the total number of screens run in each class and/or category.

targets is largely empirical, based on the inability to find small molecule leads against novel target classes [5,6]. At present, all currently approved small molecule drugs interact with just over 200 protein targets and approximately 50% of these fall into just four protein classes: GPCRs, nuclear receptors, and voltage-gated and ligand-gated ion channels [7]. The molecules that interact with these targets rely on a relatively small number of molecular scaffolds, which unsurprisingly form the basis of most major compound collections. Although such collections have been augmented by efforts from combinatorial chemistry, targeted libraries, privileged scaffolds, co-factor and secondary structure mimetics, and so on, it is clear that historical screening sets are not particularly diverse and have not provided the increase in hits that was anticipated. It is also clear that many of the new, highly attractive targets being identified fit in relatively unexplored bioactive space which traditional lead generation approaches and existing compound libraries are not well placed to exploit [8].

Accessing biologically relevant chemical space

There have been several analyses of the chemical space covered by drug molecules, chemical libraries and natural products. Two basic methods have been used: the first visualises chemical property space based on physicochemical properties [9,10]; the second is based on chemical structure [11,12]. These two approaches are complementary and independently conclude these three broad classes of molecules have different properties and cover different areas of chemical space. Notably, the space covered by drug molecules is relatively small, whereas natural products access areas not covered by the other two classes and contain structural motifs that also make them distinct from other drugs. In general, natural products are more rigid and have more fused, bridged and spirocarbocyclic rings. They also possess more oxygen and fewer nitrogen atoms than traditional synthetic molecules and a greater number of stereocentres and increased scaffold diversity. Several authors have argued that these properties make natural products, their derivatives and chemical libraries based on their scaffolds and other features more likely to provide hits against newly emerging target classes [13,14].

In addition, it has been argued that the limited number of chemical reactions that are presently used to synthesise drug molecules restricts the space accessible [15]. Synthesising more-complex molecules and using different synthetic methodologies should be explored to provide access to new chemical space and new scaffolds and functionality. Aligning these efforts with

analyses that identify gaps in chemical space should increase the chances of providing a diverse set of compounds that cover more space and increase the chances of finding hits against novel biological targets. An analysis of attrition in the development phase has indicated that compounds with a higher proportion of sp^3 centres than sp^2 centres have a better chance of surviving to become drugs [16]. Intuitively, the fraction of sp^3 centres in a molecule increases when going from commercially available compound libraries to diversity-oriented synthetic libraries and to natural products, again suggesting that more focus on natural-product-inspired compounds could prove a successful route in accessing hits to newer targets [17].

Diversity-oriented synthesis and biology-oriented synthesis

It is broadly accepted that the parallel explosion in the volume of high throughput combinatorial chemistry techniques combined with HTS techniques of the early 1990s has failed to yield tractable starting points for many drug discovery programmes. The reasons for this are probably vast, but certainly include a lack of chemical diversity (with emphasis placed on varying single scaffolds with large numbers of reagents) and biological relevance (where targets were defined based on synthetic accessibility rather than more-formal considerations of bioactive space). The area of combinatorial chemistry known as diversity-oriented synthesis (DOS), a term first coined by Schreiber [18], has emerged over the past decade as a response to this failure. DOS libraries aim to cover bioactive chemical space (known and speculative) through the synthesis of compounds with a high degree of structural diversity, in terms of functional groups and stereochemistry, but with a particular emphasis on scaffold diversity. It has been defined as the deliberate, simultaneous and efficient synthesis of more than one target compound in a diversity-driven approach [19].

Many DOS libraries are claimed to be natural-product-like, although this definition appears to be a relatively abstract one based only on a tendency toward molecules rich in stereogenic centres and with 3D skeletally complex architectures [20]. Key to DOS library chemistry is the elaboration of simple starting materials, using robust chemical reactions over a short sequence, and includes branching points where common substrates can be diversified into different scaffolds by the application of diverse reagents and conditions. It is currently unclear whether DOS libraries really do represent a more efficient way to target biologically relevant space or whether this is simply the next generation of combinatorial chemistry which is still largely agnostic to the drug target. Nevertheless, successes have been claimed for the screening of DOS libraries against a range of targets, some of which could certainly be considered traditionally intractable, including kinesin motor protein Eg5 [21], transcription factor HOXA13-DNA interaction [22], Sonic Hedgehog pathway signalling [23] and protein–protein interaction target Bcl-2 (Fig. 2) [24].

Biology-oriented synthesis (BIOS) is a term coined by Waldmann *et al.* [25] to describe the generation of compound libraries based upon iterations close to scaffolds of known biological relevance – often natural products. BIOS in this form can be seen as the combination of two concepts also from this group, specifically the distillation of the Dictionary of Natural Products into a structural classification of natural products or hierarchical 'scaffold tree' [26] and the clustering of target proteins based around similarities in the ligand-binding sites [27]. The aim then is to modify known bioactive ligands in a way that yields activity against other, often

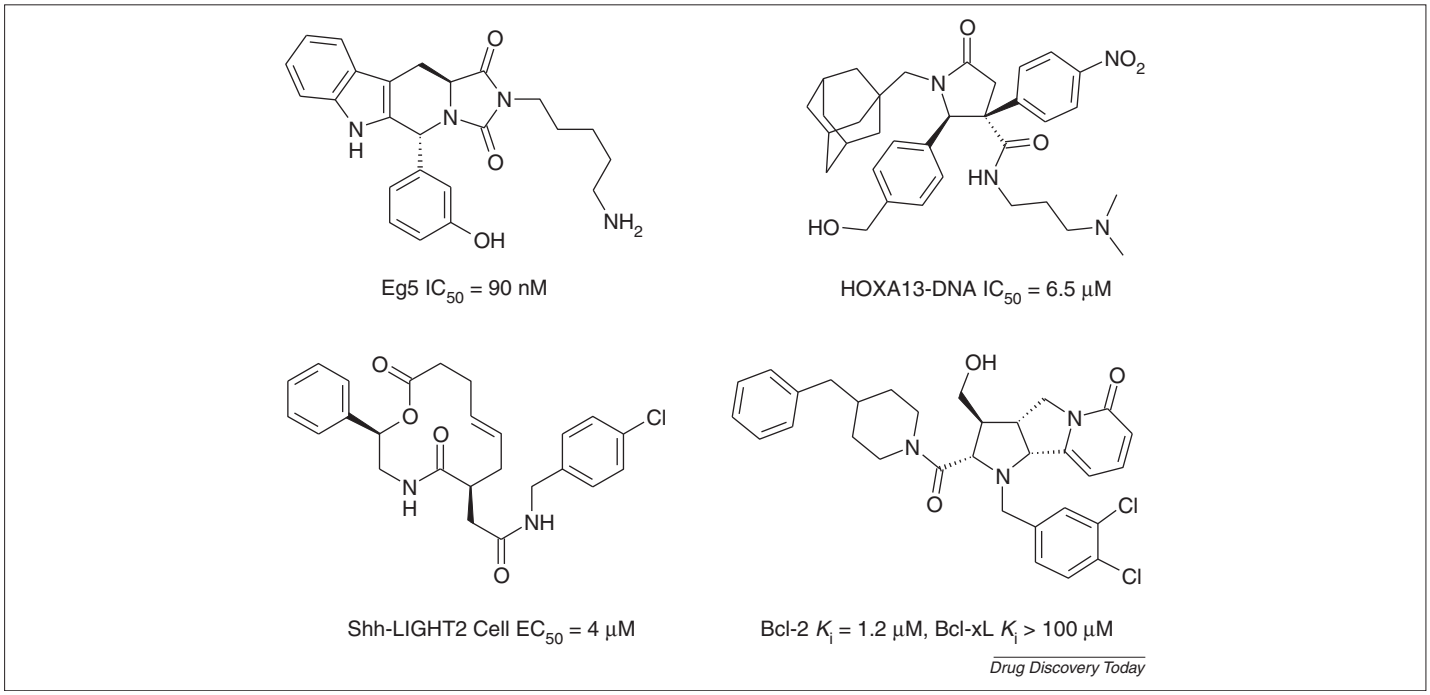

FIGURE 2

Successful examples of DOS libraries producing inhibitors of difficult or intractable targets. Clockwise from top left, an inhibitor of motor protein Eg5, an inhibitor of the interaction between DNA and the transcription factor HOXA13, an inhibitor of Bcl-2 with selectivity over Bcl-xL and 'robotnikinin', a cellular inhibitor of Sonic Hedgehog pathway signalling.

unrelated, areas of target space, an activity that has a long and productive history in medicinal chemistry research.

Natural products

The rise of combinatorial chemistry in the early 1990s was similarly associated with the cessation of natural product research at nearly all of the major pharmaceutical companies. The reasons for this effectively irreversible loss of such a historically important drug discovery capability can be debated but no doubt include incompatibility of natural product mixtures with HTS techniques, the often aggressive timelines of a modern drug discovery programme precluding rapid structure elucidation and analogue synthesis, and the current vogue for concepts such as 'lead likeness' [28] and 'drug likeness' leading to the modern medicinal chemist assessing such structures as ugly [29]. What is evident however is that natural products and natural product derivatives continue to play a significant part in the discovery of new approved therapies, as they have done since the earliest days of medicinal research.

In a comprehensive analysis of the period 1981–2006, Newman and Cragg suggest that up to 50% of drug approvals are for molecules that can trace their lineage to a natural product, a semi-synthetic derivative of a natural product or a natural-product-inspired total synthesis effort [30]. In a concurrent analysis, Butler highlights that 21 natural-product-derived drugs were approved for use in Western markets during the period 1998–2004 [31]. A more recent update to this indicates an additional 19 natural-product-derived approvals during the period January 2005 to April 2010 [32].

Through their significant structural and stereochemical diversity, natural products are uniquely placed to modulate multiple biological processes and, as industry focus moves into difficult, traditionally intractable but nevertheless high value targets, historical compound collections built up through a productive focus on GPCR, nuclear receptor and ion-channel research might become increasingly less relevant for hit finding. In a comparison of natural products and approved drugs, principal component analysis was used to highlight that both groups overlap, but that the former exemplifies a much broader range of chemical space [6]. Distinguishing features such as higher molecular weight, lower hydrophobicity, greater stereochemical complexity and fewer aromatic rings underscore the divergence with modern drug discovery efforts. Natural products have proven successful modulators of difficult targets such as a range of antibacterial targets and protein–protein interactions. Historically they have been invaluable as tools to elucidate molecular targets that elicit a particular biological response [33]. The DOS and BIOS concepts outlined in the preceding section are reasonable, if synthetically driven, responses to this challenge and, although these might help in the quest for leads against new biological space, they might be no substitute for a return to active natural product screening efforts if the inherent challenges of doing so can be overcome [34].

Phenotypic screening

Drug discovery efforts over the past few decades have been target-centric approaches, directed at a specific biological hypothesis with links to disease pathogenesis and, thus, clinical relevance. Coupling this approach with modern, highly automated target-based screening strategies (e.g. HTS) resulted in the discovery of many clinical candidates. However, despite these efforts, successful registration of new drugs has not significantly increased [35,36]. One possible contributing factor to high attrition is that HTS generally relies on equilibrium binding assays, whereas it is observed that a disproportionate number of approved drugs have non-equilibrium kinetics [37]. A broader historical perspective of drug discovery efforts also provides a potential alternative, in that these typically focused on extensive phenotypic or even *in vivo* screening and optimisation approaches, delivering many successful products, often with limited or no information on the molecular target or mechanism of action. The potential benefits and impact on attrition in clinical development have been recently reviewed [38], concluding that the majority of small molecule first-in-class NMEs discovered over the past decade resulted from phenotypic screening approaches.

The potential beneficial impact of such approaches on modern drug discovery was recently reinforced by Eli Lilly where the PD^2 screening programme was established [39]. This is a partnership aimed at identifying new chemical matter from diverse sources that has a clear impact on disease-relevant phenotypic readouts. An important but neglected area receiving renewed attention is that of drug repositioning or repurposing [40]. As modern drug discovery organisations align research into discrete target- or disease-oriented silos to improve focus and productivity, it is easy to forget that many targeted drugs end up being successful in indications for which they were not originally intended. Observations of unexpected effects, be they in cell lines, preclinical disease models or patients, can lead to significant medical advances. Notable examples include sildenafil, originally trialled in hypertension but marketed in the treatment of erectile disfunction, and tamoxifen, a product of fertility research shown to be effective in the treatment and prevention of hormone-sensitive breast cancer. Ultimately, high throughput *in vivo* screening in relevant and ethical animal models would be the greatest challenge to overcome – offering potentially high impact on modern drug discovery and closing a circle started many decades ago. *In vivo* phenotypic screening assays have been reported in a variety of model systems, including nematode, fruit fly and zebrafish, although relevance to human target biology and disease states is unclear [41].

Advances in screening technologies

Systems cell biology is the study of the living cell and of how the complex interaction of genes, proteins and signalling networks link through to function. This has opened up the opportunity to study individual molecular targets in the context of their signalling and functional networks, potentially generating a much clearer link between molecular targets and disease [42]. Advances in imaging technologies as well as automated data processing and analysis of vast datasets have recently enabled the coupling of high content biology approaches with HTS platforms [43]. This could provide potential access to studying diseases in novel ways [44] but could also give rise to new, primary, high throughput, high content cell-based screening paradigms. These could probe multiple endpoints in parallel thus offering much higher quality and creating confidence in primary screening data and could significantly extend early drug discovery strategies and approaches [45].

Such approaches are important because many small molecules are rarely entirely specific for one target and, indeed, interaction at one specific protein could lead to multiple divergent cellular consequences.

Ion-channel modulators are an important marketed therapeutic drug class, and remain a significant area of further research in the pharmaceutical industry. Advances in miniaturisation of electrophysiology have opened up new ways of pursuing this important target class. Physiologically relevant screening approaches with medium-to-high throughput capacity could give access to a much wider spectrum of chemical diversity which has thus far not been accessible or associated with this traditional area of pharmacology [46]. Protein–protein interactions, an area of drug discovery that has traditionally been regarded as 'undruggable', might benefit from newly developed protein–fragment complementation assay (PCA) or bimolecular fluorescence complementation assay (BiFC) technologies that have seen a range of applications in isolated enzyme and cell-based formats. These techniques enable dissection of cellular networks in real time and are independent of imaging technologies. PCA and/or BiFC approaches have also been applied to GPCR-focused research to further the understanding of receptor oligomerisation and its impact on signalling [47]. New label-free endpoint detection methods such as mass spectrometry [48] or the resonant waveguide grating (RWG) optical biosensor methodology [49] offer the potential for plate-based, high throughput approaches distinct from classical, fluorescence-based assays which can be prone to artefacts and interference. Currently, however, throughput and cost are factors that might limit label-free screening techniques to applications in screening hit characterisation and lead optimization, rather than broader collection screening. To date, it is mainly small scale, high value primary screening campaigns such as those used in fragment-based lead generation approaches that exploit label-free techniques best (e.g. NMR, surface plasmon resonance). In parallel to the development of new HTS approaches considerable attention has been given to post-HTS hit evaluation, more specifically the characterisation of non-specific inhibitors, the most likely source of false positives [50]. Potential causes of non-specific inhibition in biochemical assays arise from the interplay between redox properties of putative inhibitors and assay buffers and the influence of physicochemical properties that could lead to aggregation under the assay conditions. High throughput methods used for the characterisation of inhibitors as potential redox actives [51] or aggregators [52] have been described in the literature and demonstrated to be effective in triaging HTS screening output. A useful general approach used to identify potential false positives, independent of all non-specific inhibition mechanisms, is the 'ratio test' in which compounds are assayed using different enzyme concentrations. This can often be highly diagnostic in combination with the detailed analysis of concentration response data [53].

Fragment-assisted lead generation

Fragment-based approaches [54,55] are fully integrated in the early drug discovery processes across the pharmaceutical industry and have resulted in the delivery of a number of clinical candidates [56]. The principal attraction of fragment-based lead generation is the highly efficient sampling of chemical space with small molecules of low complexity. In addition to a direct impact on early drug discovery, fragment-based approaches also offer exciting opportunities in a number of additional fields relevant for lead generation activities such as druggability assessments, HTS evaluation or targeted, strategic corporate screening collection enhancements. Owing to the well documented high attrition in the drug discovery process, it is desirable to assess the likely success of a project even before a new target is formally established in a lead generation portfolio. Historically, *in silico* methods have been used for the assessment of druggability or chemical tractability of potential drug targets [57,58]. More recently, fragment screening approaches have been proposed that offer additional qualitative experimental evidence for the assessment of druggability that could be used to assess the risk of failure in early projects [59,60]. Fragment approaches against intractable targets are often still the hit finding technique of last resort, notably when HTS campaigns have failed to yield any suitable hits. Nevertheless, there have been some notable successes; the most advanced of which is Abbott's B-cell lymphoma (Bcl)-2 inhibitor navitoclax currently in Phase II clinical trials.

Recent independent publications have disclosed fragment ligands targeting oncogenic GTPase K-Ras, perhaps the most significant of all high value intractable targets. Small ligands bind in the switch II region and inhibit GTPase activation by blocking binding of guanine nucleotide exchange factor son of sevenless (SOS) [61,62].

Delivery mechanisms

Following Lipinksi's seminal publication [63], the pharmaceutical industry has rightly focused on accessing small molecules with drug-like properties such as appropriate lipophilicity, aqueous solubility and limits on the number of H-donors and acceptors to try and reduce attrition caused by issues such as poor bioavailability, high clearance, poor permeability as well as other physicochemical causes. Although clearly worthwhile, a secondary effect has been the restricting of chemical space being accessed by synthesis. It is highly probable that a significant portion of useful bioactive space lies outside this 'Lipinski space' [64] where analysis focused on marketed oral drugs and excluded certain pharmaceutical classes.

Novel delivery technologies for drugs such as liposomal formulations, depots, antibody conjugates, pro-drugs and nanotechnology approaches could make delivery of agents that have limited oral exposure more feasible. In addition, exploiting natural delivery methods such as active transport systems or methods that result in the targeting of drugs to specific compartments could enable the exploitation of newer chemical space. An example of an approach currently receiving attention is the area of stapled peptides where chemical cycling of a small peptide sequence results in α-helical peptides that show higher stability than the natural linear peptide and have an ability to cross cell membranes and interact with intracellular targets. The approach has enabled the design of novel Myeloid cell leukaemia sequence 1 (MCL-1) inhibitors with high selectivity [65]. Success in these areas will require close collaboration between chemists and formulation scientists.

New approaches to biological modulation

Traditional approaches to drug discovery usually aim to produce highly potent molecules that interact with a single target. This is

done for good reasons – to limit potential toxicity and to control and understand pharmacological responses. However, many established drugs are now known to interact with multiple targets and it is increasingly being understood that biological systems are extremely complex with multiple proteins effectively communicating with each other through various feedback loops and other interactions. One possible way to exploit this is to interact intentionally with multiple parts of a biological system and Hopkins [66] has developed network pharmacology to help identify molecules that interact with multiple components of a biological system to deliver activity. Another facet of this analysis is that structural relationships between what were previously thought to be unconnected proteins have become apparent and this could lead to small molecule leads being identified against new proteins. Other promising new approaches for modifying biological activity include modulation of mRNA stability [67], delivery of DNA and RNA fragments [68] and a whole range of approaches that are being used to interfere with protein–protein interactions and that are beginning to provide novel inhibitors of important biological function [69,70].

Concluding remarks

The proportion of screening methods directed against the main traditional target classes is in decline and in their place is a range of new, diverse and increasingly complex targets that the pharmaceutical industry must respond to. Because compound collections are a reflection of the targets we have targeted in the past, and chemical libraries are a reflection of the chemical reactions we can do, there is a real risk that currently accessible chemical space might not address the areas of biological space the industry needs to focus on. To date, collections have occupied a proportion of chemical space that overlaps with biological space, but undoubtedly there is biological space that does not conform to the Rule of Five [63].

Natural products broaden and diversify the current chemical space, and history shows that they can be drugs. DOS libraries are an attempt to 'marry' natural product diversity with the chemical reactions we can do, and is as yet unproven. However, just as chemists need to diversify structures to enable new hits to be identified, biologists need to consider how to find those hits most effectively. The field of kinase research has demonstrated that there are multiple mechanisms of action by which targets can be inhibited, such as targeting catalysis, activation and allosteric mechanisms. We must be similarly creative in the field of less-tractable targets, where multiple modes of inhibition can be considered – for example binding to and stabilising large protein complexes as well as inhibiting their formation. HTS has been the main hit generation approach of the past few decades and, for certain chemical classes, has proved valuable; but as targets change so the range of hit finding methods we call upon needs to be expanded. What is found through broad screening is as much a product of the assays used as it is the equity screened. Because some hits against high value targets now break our conventional thinking on what attractive start points look like, new technologies used to deliver the molecules to the target will need to change and a renewed focus on drug delivery mechanisms will be crucial. The next wave of truly transformational drugs will come through collaboration across the traditional scientific disciplines of chemistry and biology and between industry and academia. New institutes are being set up precisely to tackle this chemical biology challenge, and it is imperative that industry and academia can unite to solve the challenge of drugging the undruggable.

Acknowledgements

The authors gratefully acknowledge Dr Peter Greasley for the data used in Fig. 1 and Dr Mike Waring for helpful discussions during the preparation of this manuscript.

References

1 Dandapani, S. and Marcaurelle, L.A. (2010) Accessing new chemical space for 'undruggable' targets. *Nat. Chem. Biol.* 6, 861–863

2 Bohacek, R.S. *et al.* (1996) The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* 16, 3–50

3 Overington, J.P. *et al.* (2006) How many drug targets are there? *Nat. Rev. Drug Discov.* 5, 993–996

4 Drewry, D.H. and Macarron, R. (2010) Enhancements of screening collections to address areas of unmet medical need: an industry perspective. *Curr. Opin. Chem. Biol.* 14, 289–298

5 Payne, D.J. *et al.* (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* 6, 29–40

6 Schmidtke, P. and Barril, X. (2010) Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem.* 53, 5858–5867

7 Bauer, R.A. *et al.* (2010) Expanding the range of druggable targets with natural product-based libraries: an academic perspective. *Curr. Opin. Chem. Biol.* 14, 308–314

8 Bansal, A.T. and Barnes, M.R. (2008) Genomics in drug discovery: the best things come to those who wait. *Curr. Opin. Drug Discov. Dev.* 11, 303–310

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

10 Grabowski, K. and Schneider, G. (2007) Properties and architecture of drugs and natural products revisited. *Curr. Chem. Biol.* 1, 115–127

11 Wetzel, S. *et al.* (2007) Cheminformatic analysis of natural products and their chemical space. *Chimia* 61, 355–360

12 Lee, M. and Schneider, G. (2001) Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *J. Comb. Chem.* 3, 284–289

13 Kumar, K. and Waldmann, H. (2009) Synthesis of natural product inspired compound collections. *Angew. Chem. Int. Ed.* 48, 3224–3242

14 Harvey, A.L. *et al.* (2010) Current strategies for drug discovery through natural products. *Expert Opin. Drug Discov.* 5, 559–568

15 Roughley, S.D. and Jordan, A.M. (2011) The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J. Med. Chem.* 54, 3451–3479

16 Lovering, F. *et al.* (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756

17 López-Vallejo, F. *et al.* (2012) Expanding the medicinally relevant chemical space with compound libraries. *Drug Discov. Today* 17, 718–726

18 Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287, 1964–1969

19 Spring, D.R. (2003) Diversity-oriented synthesis: a challenge for synthetic chemists. *Org. Biomol. Chem.* 1, 3867–3870

20 Arya, P. *et al.* (2005) Exploring new chemical space by stereocontrolled diversity-oriented synthesis. *Chem. Biol.* 12, 163–180

21 Hotha, S. *et al.* (2003) HR22C16: a potent small-molecule probe for the dynamics of cell division. *Angew. Chem. Int. Ed. Engl.* 42, 2379–2382

22 Ng, P.Y. *et al.* (2007) Synthesis of diverse lactam carboxamides leading to the discovery of a new transcription-factor inhibitor. *Angew. Chem. Int. Ed. Engl.* 46, 5352–5355

23 Peng, L.F. *et al.* (2009) Syntheses of aminoalcohol-derived macrocycles leading to a small-molecule binder to and inhibitor of Sonic Hedgehog. *Bioorg. Med. Chem. Lett.* 19, 6319–6325

24 Marcaurelle, L.A. *et al.* (2009) Diversity-oriented synthesis of a cytisine-inspired pyridone library leading to the discovery of novel inhibitors of Bcl-2. *Bioorg. Med. Chem. Lett.* 19, 2500–2503

25 Kaiser, M. *et al.* (2008) Biology-inspired synthesis of compound libraries. *Cell. Mol. Life Sci.* 65, 1186–1201

26 Koch, M.A. *et al.* (2005) Charting biologically relevant chemical space: a structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. U. S. A.* 102, 17272–17277

27 Koch, M.A. *et al.* (2004) Compound library development guided by protein structure similarity clustering and natural product structure. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16721–16726

28 Teague, S.J. and Davis, A.M. *et al.* (1999) The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed. Engl.* 38, 3743–3748

29 Rishton, G.M. (2008) Natural products as a robust source of new drugs and drug leads: past successes and present day issues. *Am. J. Cardiol.* 101, 43D–49D

30 Newman, D.J. and Cragg, G.M. (2007) Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 70, 461–477

31 Butler, M.S. (2005) Natural products to drugs: natural product derived compounds in clinical trials. *Nat. Prod. Rep.* 22, 162–195

32 Mishra, B.B. and Tiwari, V.K. (2011) Opportunity, challenge and scope of natural products in medicinal chemistry. 1–62 (ISBN: 978-81-308-0448-4), avalable online at: http://www.trnres.com/ebook/uploads/tiwari/T_1302158431Tiwari%201.pdf

33 Hong, J. (2011) Role of natural product diversity in chemical biology. *Curr. Opin. Chem. Biol.* 15, 350–354

34 Kingston, D.G.I. (2011) Modern natural products drug discovery and its relevance to biodiversity conservation. *J. Nat. Prod.* 74, 496–511

35 Paul, S.M. *et al.* (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* 9, 203–214

36 Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* 3, 711–715

37 Swinney, D.C. (2006) Biochemical mechanisms of new molecular entities (NMEs) approved by United States FDA during 2001–2004: mechanisms leading to optimal efficacy and safety. *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)* 6, 461–478

38 Swinney, D.C. and Anthony, J. (2011) How were new medicines discovered? *Nat. Rev. Drug Discov.* 10, 507–519

39 Pineiro-Nunez, M. (2010) Lilly's phenotypic drug discovery initiative. *Abstracts of Papers, 240th ACS National Meeting*, Boston, MA, United States, August 22–26

40 Loging, W. *et al.* (2011) Cheminformatic/bioinformatic analysis of large corporate databases: application to drug repurposing. *Drug Discov. Today: Ther. Strateg.* 8, 109–116

41 Giacomotto, J. and Ségalat, L. (2010) High-throughput screening and small animal models, where are we? *Br. J. Pharmacol.* 160, 204–216

42 Giuliano, K.A. *et al.* (2006) Systems cell biology based on high-content screening. *Methods Enzymol.* 414, 601–619

43 Zanella, F. *et al.* (2010) High content screening: seeing is believing. *Trends Biotechnol.* 28, 237–245

44 Jain, S. and Heutink, P. (2010) From single genes to gene networks: high-throughput-high-content screening for neurological disease. *Neuron* 68, 207–217

45 Mayr, L.M. and Bojanic, D. (2009) Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* 9, 580–588

46 Dunlop, J. *et al.* (2008) High-throughput electrophysiology: an emerging paradigm for ion-channel screening and physiology. *Nat. Rev. Drug Discov.* 7, 358–368

47 Vidi, P.A. and Watts, V.J. (2009) Fluorescent and bioluminescent protein–fragment complementation assays in the study of G protein-coupled receptor oligomerization and signaling. *Mol. Pharmacol.* 75, 733–739

48 Roddy, T.P. *et al.* (2007) Mass spectrometric techniques for label-free high-throughput screening in drug discovery. *Anal. Chem.* 79, 8207–8213

49 Dodgson, K. *et al.* (2009) 100K well screen for a muscarinic receptor using the Epic label-free system – a reflection on the benefits of the label-free approach to screening seven-transmembrane receptors. *J. Recept. Signal Transduct. Res.* 29, 163–172

50 Shoichet, B.K. (2006) Screening in a spirit haunted world. *Drug Discov. Today* 11, 607–615

51 Lor, L.A. *et al.* (2007) A simple assay for detection of small-molecule redox activity. *J. Biomol. Screen.* 12, 881–890

52 Feng, B.Y. *et al.* (2007) A high-throughput screen for aggregation-based inhibition in a large compound library. *J. Med. Chem.* 50, 2385–2390

53 McGovern, S.L. *et al.* (2002) A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* 45, 1712–1722

54 Hajduk, P.J. and Greer, J. (2007) A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 6, 211–219

55 Congreve, M. *et al.* (2008) Recent developments in fragment-based drug discovery. *J. Med. Chem.* 51, 3661–3680

56 Chessari, G. and Woodhead, A.J. (2009) From fragment to clinical candidate – a historical perspective. *Drug Discov. Today* 14, 668–675

57 Perot, S. *et al.* (2010) Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discov. Today* 15, 656–667

58 Ward, R.A. (2010) Using protein-ligand docking to assess the chemical tractability of inhibiting a protein target. *J. Mol. Model.* 16, 1833–1843 (Online)

59 Hajduk, P.J. *et al.* (2005) Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* 48, 2518–2525

60 Edfeldt, F.N.B. *et al.* (2011) Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discov. Today* 16, 284–287

61 Sun, Q. *et al.* (2012) Discovery of small molecules that bind to K-Ras and inhibit SOS-mediated activation. *Angew. Chem. Int.* 51, 6140–6143

62 Maurer, T. *et al.* (2012) Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5299–5304

63 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3–25

64 Khanna, V. and Ranganathan, S. (2011) Structural diversity of biologically interesting datasets: a scaffold analysis approach. *J. Cheminf.* 3, 30

65 Kritzer, J.A. (2010) Stapled peptides: magic bullets in nature's arsenal. *Nat. Chem. Biol.* 6, 566–567

66 Hopkins, A.L. (2008) Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* 4, 682–690

67 Eberhardt, W. *et al.* (2007) Modulation of mRNA stability as a novel therapeutic approach. *Pharmacol. Ther.* 114, 56–73

68 Hirata, K. *et al.* (2007) Design of PCR-amplified DNA fragments for in vivo gene delivery: size-dependency on stability and transgene expression. *J. Pharm. Sci.* 96, 2251–2261

69 Dekker, F.J. *et al.* (2005) Protein structure similarity clustering (PSSC) and natural product structure as inspiration sources for drug development and chemical genomics. *Curr. Opin. Chem. Biol.* 9, 232–239

70 Betzi, S. *et al.* (2009) Protein–protein interaction inhibition (2P2I): fewer and fewer undruggable targets. *Comb. Chem. High Throughput Screening* 12, 968–983