Review

Discovery of small molecule cancer drugs: Successes, challenges and opportunities

Swen Hoelder*, Paul A. Clarke, Paul Workman*

Cancer Research UK Cancer Therapeutics Unit, Division of Cancer Therapeutics, The Institute of Cancer Research, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK

ABSTRACT

The discovery and development of small molecule cancer drugs has been revolutionised over the last decade. Most notably, we have moved from a one-size-fits-all approach that emphasized cytotoxic chemotherapy to a personalised medicine strategy that focuses on the discovery and development of molecularly targeted drugs that exploit the particular genetic addictions, dependencies and vulnerabilities of cancer cells. These exploitable characteristics are increasingly being revealed by our expanding understanding of the abnormal biology and genetics of cancer cells, accelerated by cancer genome sequencing and other high-throughput genome-wide campaigns, including functional screens using RNA interference. In this review we provide an overview of contemporary approaches to the discovery of small molecule cancer drugs, highlighting successes, current challenges and future opportunities. We focus in particular on four key steps: Target validation and selection; chemical hit and lead generation; lead optimization to identify a clinical drug candidate; and finally hypothesis-driven, biomarker-led clinical trials. Although all of these steps are critical, we view target validation and selection and the conduct of biology-directed clinical trials as especially important areas upon which to focus to speed progress from gene to drug and to reduce the unacceptably high attrition rate during clinical development. Other challenges include expanding the envelope of druggability for less tractable targets, understanding and overcoming drug resistance, and designing intelligent and effective drug combinations. We discuss not only scientific and technical challenges, but also the assessment and mitigation of risks as well as organizational, cultural and funding problems for cancer drug discovery and development, together with solutions to overcome the ‘Valley of Death’ between basic research and approved medicines. We envisage a future in which addressing these challenges will enhance our rapid progress towards truly personalised medicine for cancer patients.

* Corresponding authors. Tel.: +44 20 8722 4301; fax: +44 20 8722 4324.
E-mail addresses: swen.hoelder@icr.ac.uk (S. Hoelder), paul.workman@icr.ac.uk (P. Workman).

© 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

doi:10.1016/j.molonc.2012.02.004
1. Introduction: progress and frustration

The transition from cytotoxic chemotherapy to molecularly targeted cancer drug discovery and development has resulted in an increasing number of successful therapies that have impacted the lives of a large number of cancer patients. The administration of anti-oestrogens and anti-androgens to treat breast and prostate cancers that are driven by the respective hormones is well established. The curative activity of all-trans retinoic acid in the treatment of most patients with acute promyelocytic leukaemia harbouring translocations in the RARα-retinoic acid receptor gene established the validity of the concept of targeting pathogenetic driver abnormalities with a small molecule in the clinic (Huang et al., 1988). Following on, the ABL inhibitor imatinib is generally regarded as a trail blazer drug that most impressively validated the concept of designing a small molecule therapeutic to treat a defined patient population – in this case chronic myeloid leukaemia in which the malignancy is driven by the BCR-ABL translocation and for which the improvement in survival has been dramatic (O’Brien et al., 2003; Druker et al., 2006).

These successes were followed by a number of other small molecule drugs inhibiting critical cancer targets, e.g. the epidermal growth factor receptor (EGFR) kinase inhibitors gefitinib and erlotinib that potently inhibit EGFR in patients with non small cell lung cancer (NSCLC); the EGFR/erbB2 inhibitor lapatinib for erbB2-positive breast cancer; and the vascular epidermal growth factor receptor (VEGFR) kinase inhibitor sorafenib in renal cancer (Yap and Workman, 2012). Most recently the CYP17A1 inhibitor abiraterone, which blocks androgen synthesis, has been approved for late stage, castration-resistant prostate cancer and is likely to change the standard of care for these patients (de Bono et al., 2011). In addition, inhibitors of the protein kinase ALK by crizotinib (Kwak et al., 2010) and of another kinase BRAF by vemurafenib (Chapman et al., 2011) have recently been approved for the treatment of NSCLC patients with a pathogenic rearrangement of the ALK gene and metastatic melanoma with the BRAF V600E mutation, respectively. The progress with small molecule drugs is mirrored by the successful introduction of protein-based therapeutics, particularly antibodies, as exemplified by the anti-erbB2 monoclonal antibody trastuzumab in erbB2-positive breast cancer (Slamon et al., 2001). These examples provide ample evidence of the success in targeting the pathogenic drivers to which cancer cells are ‘addicted’ (Weinstein, 2002; Weinstein and Joe, 2008).

However, despite the considerable progress made with the new molecularly targeted therapies, including advances in diseases like NSCLC and melanoma for which few treatment options are available (Yap and Workman, 2012) for many patients the therapeutic options are still limited and the process of bringing a new drug to patients is still frustratingly slow with high failure rates (DiMasi and Grabowski, 2007; Kola and Landis, 2004), a problem often referred to as the ‘Valley of Death’ between basic research and new drug approval. There are several reasons why progress is not as fast as we would like it to be.

Firstly, it has only relatively recently been fully appreciated that within a particular anatomically and histologically defined solid tumour type patients need to be treated with a particular class of kinase inhibitor that matches the predominant pathogenetic driver mutation. Thus NSCLC patients with EGFR mutations respond to EGFR inhibitors while those with ALK translocations respond to ALK inhibitors (Collins and Workman, 2006; Yap and Workman, 2012). Recognition of the value of targeting specific oncogene addictions and the need for companion biomarkers for patient selection is now having a major impact on cancer drug discovery and development.

Secondly, although good progress has been made in identifying and then drugging pathogenic driver gene targets, the sequencing of increasingly larger numbers of cancer genomes has revealed extraordinary complexity, including the presence of thousands of genetic alterations and considerable genetic heterogeneity, not only between different tumours but also within an individual cancer (McDermott et al., 2011; Sellers, 2011; see also De Palma and Hanahan, 2012). This heterogeneity undoubtedly contributes to drug resistance and the need for drug combinations (see later).

Further on, once a potential novel therapeutic target has been identified, there can be significant scientific and technical hurdles to discovering a novel and effective drug. Importantly, it is increasingly recognised that selecting and validating the best targets can be very challenging (Benson et al., 2006). Not only must causal linkage of the proposed target to the clinical disease be firmly established and a clear clinical hypothesis developed, but the quantitative consequences of target modulation must be shown to be sufficient to deliver a therapeutically meaningful biological effect in relevant experimental models.

Yet another concern for drug discovery is the ‘druggability gap’. Many targets with very promising disease linkage, such as mutated RAS proteins or transcription factors like c-MYC or hypoxia-inducible factor (HIF), are currently regarded as technically undruggable – or at the very least as extremely challenging to target by medicinal chemists using small molecules (Verdine and Walensky, 2007). Even in the case of more druggable proteins it takes several years to identify a drug candidate that satisfies the stringent requirement for clinical development (Paul et al., 2010).

Importantly, we have learned that disease hypotheses do not always translate from cellular and in vivo models into the clinic, as particularly illustrated by the plethora of agents, such as histone deacetylase inhibitors and antimitotic and antiangiogenic drugs, which have entered clinical trials but so far shown only limited efficacy and/or high failure rates, often in large Phase III randomised clinical trials (Bates et al., 2012). Drugs acting on other interesting targets, such as the cancer-supportive molecular chaperone HSP90 and the frequently mutated oncogenic lipid kinase PI3 kinase, remain works in progress in terms of moving towards regulatory
approval (Neckers and Workman, 2012; Travers et al., 2012; Clarke and Workman, 2012) – an achievement that represents the true and ultimate measure of target validation.

Another frustrating problem is that for drugs which have been brought successfully into clinical trials and shown therapeutic efficacy, resistance frequently emerges, as illustrated recently by crizotinib (Sasaki et al., 2011) and vemurafenib (Poulikakos et al., 2011). Resistance to molecularly targeted agents can be due to mutation of the target itself, as in the case of kinase gatekeeper mutations (Gibbons et al., 2012), the activation of adaptive feedback loops (Rodrik-Outmezguine et al., 2011) or alternative oncogenic pathways (Johannessen et al., 2010; Nazarian et al., 2010). It is interesting that in most cases the resistance mechanism preserves the original overall pathway addiction, e.g. to the RAS-RAF-MEK-ERK or PI3 kinase-AKT kinase signal transduction cascades. However, the rational design of combinations to overcome such problems, as well as the issue of clonal heterogeneity, still proves very challenging (Whitehurst et al., 2007).

Yet there are many reasons to be optimistic for cancer drug discovery and development, not least due to recent scientific and technological breakthroughs that help us to tackle the challenges we face. Novel concepts such as ‘non-oncogene addiction’ (Luo et al., 2009) and ‘synthetic lethality’ (Ashworth et al., 2011) have widened the scope beyond the exploration oncogenic pathway addictions and have helped guide the identification of novel targets either through hypothesis-driven research or large-scale screening campaigns. Massive genome sequencing and molecular pathology efforts in conjunction with bioinformatics and systems biology approaches are allowing us continuously to refine our understanding of how cancer cells are wired and how they can be targeted through single agents or on several fronts through drug combinations (Stratton, 2011; Macconal1 and Garraway, 2010; Kitano, 2003).

Medicinal chemists, working closely with their biological collaborators, have become very efficient at discovering effective and safe clinical candidates acting on new targets and are pushing the boundary of what is regarded as technically druggable. It is increasingly common to see the use of genetically engineered mouse models (GEMMs) to test a drug under conditions that may closely mimic the real-life clinical situation (Politi and Pao, 2011).

Finally, there can be no doubt that we have moved into the era of personalized or stratified medicine which ensures that the right drug is given to the right patient at the right time, thereby ensuring fast progression through clinical trials and maximum therapeutic benefit to patients (de Bono and Ashworth, 2010; Yap and Workman, 2012). All of these advances are illustrated by the recent approval of the BRAF inhibitor vemurafenib only 9 years after the BRAF mutations had first been published (Chapman et al., 2011).

So the current position is one in which we have seen some extraordinary successes in our exploitation of the cancer genome and tumour biology, while at the same time encountering a number of frustrating challenges that nevertheless represent opportunities to drug discovers and developers.

In this article, we will review the current state of cancer drug discovery and development, focusing on small molecules that act on new molecular targets that represent therapeutic dependencies and vulnerabilities. This translational activity can at the highest level be broken down into four steps that reflect the entire value chain (Figures 1 and 2). These stages are:

- Target selection and validation
- Chemical hit and lead generation
- Lead optimisation to select a clinical candidate
- Biology-led clinical trials

We will discuss each step individually in the following sections and we will conclude with a discussion of solutions to overcome the ‘Valley of Death’ between basic cancer research and approved cancer drugs. We have written this review from our perspective as members of a drug discovery and development group that is actively involved in bringing forward personalised medicine.
2. Target validation and selection

Selecting the target to work on is the most important decision a drug discovery organisation faces (Bunnage, 2011). This is because if the wrong target is selected then the considerable resources that are applied downstream of target selection — including, potentially, the enormous costs of clinical trials that fail to show therapeutic activity — will be entirely wasted. In addition, there is the opportunity cost associated with not working on a better target.

The selection of a new target for cancer drug discovery is increasingly based on the strength of the evidence that it represents a dependence or vulnerability for a given stratified set of cancer patients, commonly defined by their molecular genetic status (Benson et al., 2006). Such evidence is important not only to provide confidence that an anti-tumour effect will be achieved if the target is modulated pharmacologically, but also to help ensure as far as possible that selectivity for tumour versus normal healthy cells will be achieved by the eventual drug that emerges from the drug discovery project.

The selection of a right target is almost inevitably a question of balancing opportunities with risks. Opportunity is generally reflected by the unmet medical need that a novel drug will address. The risk can be broadly assessed under two categories:

1) Biological risk — i.e. will an inhibitor of the target, if discovered, show the desired biological and therapeutic effect in patients?

2) Feasibility risk — i.e. is it possible that an inhibitor that acts on the target and exhibits drug-like properties (see later) can actually be discovered?

These individual risks should be viewed collectively to assess the overall risk associated with a novel target and it should be stressed that any one particular risk on its own should rarely be regarded as a show-stopper. For example, it might well be appropriate to progress a target with high biological risk provided the feasibility or technical risk is low and it is likely that a chemical tool compound can be identified relatively quickly that will help to de-risk the target. Likewise, it may make sense to progress a target with high technical risk if the biological validation is very strong and the drug would meet a significant unmet medical need.

2.1. Assessing biological risk in target validation

Candidate targets come from a variety of sources that are either based on biological high-throughput screening efforts or on hypothesis-driven research. Whilst these approaches provide us with a range of proteins that show a link to tumour biology, these potential targets have to be regarded as candidates that will essentially always have to undergo further in-depth validation experiments that are designed to build sufficient confidence that modulation of the particular target would lead to the desired therapeutic effect.

The assessment of the validity of a given cancer target can conceptually be focused onto two key questions:

Figure 2 — Drug discovery and development: From gene to drug. The four key steps of cancer drug discovery that are covered in this review are highlighted in the centre. ‘Reverse translation’ from the clinic back to the laboratory is covered under the target validation and selection section. Modified with permission from (Collins and Workman, 2006).
1. Is there sufficient evidence that the target is overactive or part of an overactivated pathway that drives a defined set of cancers and contributes to cancer growth? — or that it supports the oncogenic process or represents a vulnerability that can be exploited through synthetic lethality?
2. Is there evidence that inhibition of the target will lead to the desired effect in patients, for example tumour regression and increased survival at well tolerated doses?

Over the last decade or so, molecular therapeutic targets for cancer treatment have been identified as individual gene products following rigorous and focused assessment of their contribution to tumourigenesis and their potential as for drug discovery. Increasingly, more systematic approaches are being employed that involve genome wide unbiased cancer genome sequencing. Examples include the UK-based Cancer Genome project (http://www.sanger.ac.uk/genetics/CGP/) and the NCI Cancer Genome Atlas project (http://cancergenome.nih.gov/) in the USA. An exhaustive worldwide global cancer genome sequencing campaign across a wide range of cancer types is now underway under the auspices of The International Cancer Genome Consortium or ICGC (http://www.icgc.org). Indeed the goal of the ICGC is: ‘To obtain a comprehensive description of the genomic, transcriptomic and epigenomic changes in 50 different tumour types and/or subtypes which are of clinical and societal importance across the globe’. These large-scale efforts aim to characterise many tumours of different types comparing global genome abnormalities, gene mutations, gene expression and epigenetic changes. The scale and speed of these efforts is becoming ever more feasible with the recent advent of higher throughput and more cost-effective Next-Generation Sequencing technologies for DNA and RNA (Boehm and Hahn, 2011).

Cancer genome sequencing has identified some important and frequently mutated oncogenes for which addiction has been demonstrated and against which drugs have been discovered, developed and approved — the most notable example being BRAF. However, it is clear that oncogenes that are now being identified will be much less frequently mutated, and therefore represent smaller, though still potentially important patient populations.

As well as defining the full repertoire of cancer genes and other molecular abnormalities, a further objective of the comprehensive profiling campaigns will be to understand how cancer genes interact together in dynamic networks. These effects are likely to identify further new targets and certainly help to prioritize certain target loci that may be especially important in particular genetic systems and disease contexts. Also critical is to identify genes that are not necessarily themselves genetically altered but that are essential for viability in cancer cells. Alongside sequencing and expression profiling efforts, the complimentary approach of functional screening based on genome-wide RNA interference (RNAi) is currently being used to identify such targets (Brough et al., 2011; Mohr et al., 2010; Quon and Kassner, 2009).

It can be argued that the strongest evidence that a target is important for the biology of a certain tumour is of genetic nature, for example detection of activating mutations as with BRAF (Davies et al., 2002); gene amplification such as for ERBB2 (Slamon et al., 1989); or fusion genes as in the case of ALK (Soda et al., 2007). Importantly, it is essential to distinguish between passenger mutation/genetic alterations that do not contribute to tumour biology and the significantly less frequent driver mutations/genetic alterations to which a tumour is addicted (Sellers, 2011). Although statistical evidence is important in this respect, ‘wet biology lab’ experimentation is important in identifying and validating driver oncogene targets. This can be done firstly by showing that introduction of the genetic defect or increased expression of a particular gene is able to transform normal cells into those having a more cancerous phenotype. And secondly by showing that the reversal of the molecular abnormalities, for example by RNAi or a dominant negative construct, is able to block malignant transformation or kills the cancer cell. Such experiments are generally performed in cancer cell lines in tissue culture, but genetically-engineered mouse models (GEMMS) are increasingly being used to validate that introduction of a genetic aberration leads to the formation of tumours and that correction of it yields an antitumour effect (Politi and Pao, 2011).

Showing that a target is genetically activated leading to addiction certainly builds confidence in the target. But not all targets fall into this category. Recently a concept sometimes referred to as non-oncogene addiction has been proposed. This model suggests that a variety of non-oncogenic gene products/pathways that lack oncogenic activity in their own right may still be essential for supporting the transformed cancer phenotype and as such may also represent potential therapeutic targets (Luo et al., 2009).

This model of non-oncogenic addiction also encompasses the concept of synthetic lethality (Luo et al., 2009; Ashworth et al., 2011). According to this concept, cancer-associated genes can be identified that represent a vulnerability that results in cell death only when another gene is inactivated. In the cancer therapy context, this can mean the identification of a drug target that when inhibited pharmacologically results in the death of cells harbouring a particular tumour suppressor gene. The initial worked example is the dramatic observation that tumours with mutations in BRCA1 or BRCA2 become much more dependent on the activity of poly (ADP-ribose polymerase) or PARP and are much more highly sensitive to inhibitors of PARP compared to normal cells, even though PARP expression or activity is generally not increased in these tumours (Lord and Ashworth, 2008). Selective killing is due to the fact that BRCA1/2 mutant cells are deficient in DNA double-strand repair by homologous recombination and thus become extremely sensitive to cell death when a second DNA repair mechanism, mediated by PARP, is inhibited.

The identification of such synthetic lethal targets is attractive because it provides a means to exploit pharmacologically loss-of-function tumour suppressor gene defects, which are otherwise essentially undruggable. Use of PARP inhibitors in BRCA mutant patients has been validated clinically, initially with the drug olaparib (Fong et al., 2009). Another, albeit somewhat different example of indirectly attacking a tumour suppressor is inhibiting the protein kinase AKT/PKB that is activated upon mutation/deletion or epigenetic suppression of tumour suppressor gene PTEN in many cancers (Hollander et al., 2011). In this case, inhibition of the target antagonises the effect of tumour suppressor loss rather than normalising...
the activity of a target that is overactive due to a genetic activation. These examples, particularly the synthetic lethality approach, are generating increased interest in identifying targets that can be inhibited in cancers with defined loss of function mutations.

Having established that the target of interest is important in the biology of cancer, the second important step when validating a target is to increase confidence in the hypothesis that inhibition of the target will lead to the desired therapeutic effect in patients. These validation experiments are now frequently performed, as discussed above, by suppressing the expression of a protein through RNAi, since a suitable chemical inhibitor is often not available at this stage of the project. However, where such chemical tools are available, these can be very valuable, subject to a number ‘fitness factors’ and controls that we have set out (Workman and Collins, 2010).

The limitations of using RNAi to mimic a chemical inhibitor have recently received widespread attention (Palchaudhuri and Hergenrother, 2010; Quon and Kassner, 2009; Sigoillot and King, 2010). RNAi will rarely completely suppress target expression and it is now widely accepted that most if not all RNAi reagents can have off-target effects additional to inhibition of target gene expression. Multiple sequences and appropriate non-targeting controls should be used to mitigate the risk of off-target effects.

An additional complication is that inhibition of target gene expression by RNAi leads to the depletion of the entire protein which may influence the composition of protein complexes or protein scaffolding functions. In contrast, a small molecule inhibitor will generally only inhibit one particular function, e.g. the enzymatic activity. If the biological effects observed through RNAi knockdown is due to inhibition of a scaffolding function they may not be recapitulated by a small molecule inhibitor. These concerns can be addressed to some extent by suitable control experiments. For example, exogenous expression of an RNAi knockdown-resistant version of the target gene should rescue the effect of the RNAi unless the RNAi activity is a result of off-target effects (Falschlehner et al., 2010; Palchaudhuri and Hergenrother, 2010; Sigoillot and King, 2010). Conversely, the effect should not be rescued by exogenous expression of mutant that reflects inhibition through a small molecule, e.g. expression of a catalytically dead mutant in the cases of enzymes.

Achieving the same biological effect with RNAi and with chemical tool compounds (especially using inhibitors from more than one chemical class or chemotype and incorporating inactive analogues) is especially persuasive (Workman and Collins, 2010).

Once confidence has been built that the effect of RNAi is a genuine on-target mechanism, the observed phenotype has to be closely scrutinised. In many cases, it is expected that the effect of target depletion or inhibition should be selective to cancer cell lines that are addicted to the particular target or pathway. Determination of whether this leads to cytostasis or cell death, e.g. through apoptosis or autophagy, is important — with cell death being a more desirable therapeutic outcome. It is important to demonstrate that target modulation leads to the expected effect in multiple cell lines which show that particular genetic lesion (Brough et al., 2011). It is equally important that no or minimal effect is observed in cancer cells with a different genetic background, and ideally also in normal, non-transformed cell lines from the same tissue, although the latter are challenging to work with in culture and to transduce.

The robustness and reproducibility of effects is important. A recent publication has shown that a high proportion of potential targets identified in the literature could not be validated in-house within the laboratories of a pharmaceutical company (Prinz et al., 2011). This emphasizes the importance of obtaining robust and reproducible data — ideally, as mentioned, in multiple cell lines and with multiple reagents and orthogonal technical approaches. Also important is our own experience of obtaining quantitative data, not only on the degree of target inhibition or knockdown but also on the extent of the resulting phenotype observed. Thus the amount of cell cycle arrest or cell death induced must be sufficient to result in a meaningful therapeutic effect.

In addition to target validation studies in vitro, there is an increasing need to look at the effects of target modulation in vivo animal models, enabled by the increasing availability of validated short hairpin RNAs for target knockdown, ideally using an inducible system. Furthermore, it should be remembered that although inhibition of cancer cell proliferation and cell death predominate as desirable therapeutic endpoints, other malign cancer hallmarks that may need to be evaluated are properties such as motility, invasion, angiogenesis and metastasis (Hanahan et al., 2000; see De Palma and Hanahan, 2012).

2.2. Assessing technical risk

Even if a target shows strong biological validation, the discovery of a safe and effective small molecule drug candidate is still a scientifically challenging and risky endeavour that can fail at several stages. For this reason it is essential to think through the whole project to identify key roadblocks, challenges and risks that will threaten the identification and clinical development of the drug that will emerge. Key risks or challenges include that compound screening will not deliver chemical hits that act on the target or that a biological assay that is critical for the progression of the project cannot be made available. Many technical risks or gaps can be tackled and managed if they are recognised sufficiently early. Some typical questions and issues are discussed below.

As mentioned, a key question is whether the proposed target is druggable with a small molecule agent. By druggability we refer here to the question of whether a target can be modulated in the desired fashion by a small molecule (Hopkins and Groom, 2002; Russ and Lampel, 2005). To bind with strong potency, i.e. Kd < 20 nM, a small molecule has to engage in a number of hydrophobic and polar interactions with its target. For drug-like molecules, this is generally only possible if the drug binds to an enclosed and hydrophobic pocket or cavity on the protein. Or to use Paul Ehrlich’s classic lock and key analogy (Kaufmann, 2008), a druggable target must have a keyhole into which a small molecule key is able to fit. Evaluating druggability in the context of small molecules thus comes down to analysing if the target contains such a cavity or pocket. Different methods exist to evaluate druggability, a few important and topical examples of which will be reviewed briefly here.
chance that it will be druggable as well, especially if the sequence homology is high (Hopkins and Groom, 2002; Russ and Lampel, 2005). The publicly available database canSAR (https://cansar.icr.ac.uk/), which was developed in our Unit, is a powerful tool with which to search for homologous targets based on protein sequence and structure as well as to identify and download published chemical hit matter (Halling-Brown et al., 2012). Such a search, which we use ourselves extensively and which we recommend should always be performed for a novel target, has the added benefit that any inhibitors of a related target can serve as the launch pad with which to kick-start a new drug discovery project.

Structure-based assessment can be very valuable if 3-dimensional structural information is available on the target. Thus if an X-ray crystal or NMR structure is in hand this can be inspected for enclosed pockets suitable for the binding of a small molecule. The analysis can be done qualitatively by visually inspecting the structure or more quantitatively using computational tools that rank targets according to predicted druggability (Fuller et al., 2009).

It is increasingly recognised that pockets that are able to bind drug-like molecules are not necessarily fully present in static picture of an apo crystal structure (Lee and Craik, 2009). Instead these pockets may represent a particular conformation of the protein that does not crystallise but can be trapped by binding to a drug. These transient pockets represent an exciting opportunity for targets that are difficult to drug but are still very challenging to identify and explore (Surade and Blundell, 2012).

The druggability of a protein can never be predicted with absolute accuracy, especially if little information on the target is available. It is often possible to assess fairly accurately the extreme cases, i.e. the very druggable (for example protein kinase-like) or the essentially undruggable (such as the flat surfaces involved in many protein–protein interactions). But there is a considerable grey zone in between these two extremes. In these cases attempting to identify chemical hits is frequently the only way to find out if the target is likely to be druggable.

Whilst a high-throughput screening campaign using a large library of drug-like molecules is expensive to run, the screening of around 1000 very low molecular weight compounds known as ‘fragments’ has been proposed as a cost-effective test of druggability (Hajduk et al., 2005). Fragment screening will be discussed further below. But it is important to mention in this present context that fragment binding can only be taken as an indication of druggability if the fragments have been shown to bind to a site that is relevant for drug discovery.

There are a number of technical enablers that must be considered ahead of committing to a drug discovery campaign on a new target. Particularly important is the so-called ‘screening cascade’, a term that is often used to describe the series of biological assays that are needed and the relative order in which compounds are progressed through them. Figure 3 shows a typical example that we used for our HSP90 drug discovery programme (in collaboration with Vernalis) that ultimately led to the drug candidate NVP-AUY922 that is now in Phase II clinical trials (Brough et al., 2007). The test cascade assays are run on compounds generated in successive iterative make-and-test cycles, so that the desired properties of the drug, which are usually decided by the team at the outset, are progressively achieved.

The biological test cascade will usually be specific to a particular target and thus tailored to the drug discovery project concerned. Most projects will require a biochemical assay with recombinant protein(s). This assay needs to have

![Figure 3](image-url)
sufficient throughput and fidelity, particularly if the screening of large compound libraries is planned.

Next, both a cancer cell proliferation assay or similar together with a mechanism-based cellular endpoint are pivotal to show that anticancer effects are mediated through inhibition of the target in the cellular setting (Collins and Workman, 2006). These assays are particularly relevant if the desired cellular outcome is cell death since cell killing can be induced by many (off-target) mechanisms. The observation that a compound induces apoptosis does not necessarily mean that it does so through inhibition of the desired target. Examples of mechanism-based cellular read outs are phosphorylation of downstream targets in the case of a target kinase or depletion of client proteins as in the case of HSP90 (Brough et al., 2009; Sharp et al., 2007). Several technologies now exist to perform these cellular mechanistic assays with significant throughput and with quantitation, e.g. high-content cellular imaging or ELISA-based assays.

Quantitation is important for ranking compounds and understanding their structure-activity relationships and activity is commonly expressed as the concentration of compound required to achieved a 50% inhibitory effect.

In the case where the target concerned belongs to a protein family, selectivity assays are commonly necessary to assess preferential activity for the desired target against other family members and to minimise off-target effects.

In addition to improving potency and selectivity on the desired target, medicinal chemists must also improve the physicochemical properties of a prototype drug lead, particularly solubility and lipophilicity, in order to achieve cell permeability and appropriate in vivo pharmacokinetic properties. Thus it is important to run assays that predict for appropriate characteristics with respect to ADME (Absorption, Distribution, Metabolism and Excretion).

Prior to testing for in vivo therapeutic activity in an animal model, most modern cancer drug discovery project teams will optimize pharmacokinetic (drug exposure) and pharmacodynamic (target engagement and downstream pathway inhibition) properties, the latter measured by appropriate target-specific biomarkers (Collins and Workman, 2006). Defining appropriate in vivo models with sufficient throughput to support the make-and-test cycles involved in medicinal chemistry optimisation — often requiring data turn around on a weekly or bi-weekly basis — is frequently a challenge. We will cover this topic further below.

We believe that it is essential for the project team to consider the likely clinical trial strategy early on in the project. In particular, key questions of how patients will be stratified using predictive biomarkers and which pharmacodynamic biomarkers will be used to demonstrate target modulation should be addressed. In our own Unit we have adopted a ‘no biomarker no project’ policy which can present a substantial hurdle, particularly for newer targets that lack a substantial body of background data. Nevertheless, we believe that it is critical to plan for these downstream activities at an early stage.

It is also important consider, both early on and then continuously after, the competitive landscape around the target. A new drug will only be successful if there are not already drugs available that adequately address the medical need. When evaluating a potential target it is thus important to obtain an overview of the competition. Assessing the competitive landscape is often not straightforward for several reasons. It involves the analysis of patent applications which are notoriously difficult to penetrate. In addition, patents are only published 18 months after submission and there is a significant time gap between patents that have been filed and patents that can be analysed. Moreover, the situation at the start of a drug discovery project has to be forecasted into the future when the project reaches clinical trials or even the market. These predictions can only be done with limited accuracy, yet assessing the competitive situation at an early stage will give important clues. For example, if there many competitors of which several are ahead, the chances are that the project will face strong competition when entering the market and clear advantages will need to be demonstrated.

On the other hand, it should be noted that embarking on a drug discovery project on a target for which competitors are somewhat ahead can be advantageous because a project team can learn from their progress as evident from patent applications and publications to generate drugs with a competitive edge. In addition, it has been shown that for novel targets, the first drug to reach clinical trials — although potentially important for a proof-of-concept — is often not the first to reach the market. Moreover, even if a drug is already approved, a follow-on drug can still receive priority review from the regulatory authority such as the US FDA (DiMasi and Faden, 2011). These considerations are more likely to be relevant for industry teams, as academic drug discovery groups are more inclined to focus on innovative high risk targets, although the competition may still be strong.

However, at the latest when any competitor compounds are reaching the clinic, these compounds will need to be prepared or sourced and then profiled in depth to define their properties (DiMasi and Faden, 2011; Giordanetto et al., 2011), hence allowing an assessment of whether the in-house project already has or can produce superior compounds.

In summary, choosing a promising target is a matter of assessing opportunities and risks, many of which we have discussed here. It goes without saying that eliminating all risks early on is impossible. Moreover, cancer is an area with high unmet medical need and thus, rather than shying away from risks, a more productive approach is to assess and manage them effectively from the start of the project. Also worth noting is that a non-profit drug discovery group like our own may frequently be prepared to take on higher levels of risk than a commercial company, since our raison d’etre is to be innovative.

### 3. Hit and lead generation

Once a target is chosen, the project team decides on a strategy to generate chemical hit matter. This is the term used to describe chemical compounds that appear to be early prototypes that act on the drug target.

Several different hit generation approaches have successfully been employed. These approaches can either be used individually or in combination, depending on the nature and need of the drug discovery project.
The decision on the type of the hit identification strategy to be employed is an important one for a project. Hit discovery is time-consuming and particularly in the case of high-throughput screening is very resource-intensive. It is also frequently a point of no return: once a high-throughput screening campaign has been conducted, it is often impossible to perform another one, not least because of the costs involved, or too late to embark upon another strategy while still remaining competitive on a hot target. The chosen strategy will also directly impact the properties of the hit matter that will subsequently be obtained and thus influence how much time will be required to optimise the hits, first to obtain lead compounds and then to produce a preclinical candidate (Nadin et al., 2012).

Conceptually, two overall types of approaches for hit finding can be distinguished: knowledge-based design and random screening. Designing hits requires some sort of prior knowledge. This can be a crystal structure of the target, or the chemical structure of known inhibitors or a natural ligand. This information can be used to design or select relatively few (10^1 to 10^3) compounds which will then be screened.

In contrast, random screening does not require prior knowledge of target structure or of inhibitors/ligand, and involves screening of large compound collections. In reality, most hit discovery campaigns commonly involve a mixture of these two approaches and represent a trade-off between required target knowledge, on the one hand, and the number of compounds screened, on the other (Figure 4).

Nowadays, even large screening collections are carefully designed to contain so-called drug-like compounds – i.e., those with properties resembling known drugs (Gribbon and Sewing, 2005) – while even the most confident scientist following a knowledge-based approach will design a number of compounds to be screened.

3.1. High-throughput screening

Since the advent of targeted therapy, high-throughput screening (HTS) has been a cornerstone for hit generation. The first cancer drugs originating from HTS have reached the market, e.g. soranfinib (Macarron et al., 2011). Others are undergoing clinical trials, as in the case of inhibitors of HSP90 (Cheung et al., 2005; Brough et al., 2007), PI3 kinase (Folkes et al., 2008; Raynaud et al., 2009) and the Hedgehog pathway (Robarge et al., 2009).

One advantage of HTS is that it is a relatively unbiased approach and can identify compounds with novel binding modes. Recent examples include allosteric inhibitors of the kinase AKT/PKB (Figure 5). Initially compound 1 was identified by HTS and shown not to be competitive with ATP, in strong contrast to most other kinase inhibitors (Lindsley et al., 2005). Optimisation of this series of compounds eventually led to the drug candidate MK-2206 (2) (Hirai et al., 2010) which is now in Phase II clinical trials.

The mechanism of action was revealed by solving the crystal structure of AKT1 bound to 3, another potent inhibitor from this class of compounds (Wu et al., 2010). 3 binds to a pocket at the interface of the regulatory PH domain and the kinase domain, approximately 10 Å away from adenine pocket where ATP competitive inhibitors bind. Binding to this allosteric pocket stabilises the kinase in an inactive (PH-in) conformation. This PH-in conformation is significantly different from the PH-out conformation observed with ATP competitive inhibitors. Importantly Wu et al. hypothesise that the PH-in conformation explains why the allosteric inhibitors potently inhibit the phosphorylation of AKT itself and do not trigger the well known feedback activation observed with ATP competitive inhibitors (Okuzumi et al., 2009). A key step of the activation of AKT is recruitment of the protein to the

Figure 4 — Choice of screening strategies for finding hit matter acting on the desired target. Most drug discovery project teams tend to use hit identification approaches that represent a mixture of design and screening elements. Increasing knowledge of the target facilitates the screening of fewer compounds, whereas if many compounds are screened little knowledge on the target is required.

Figure 5 — Allosteric inhibitors of the kinase AKT/PKB.
membrane. However, the phospholipid binding which is responsible for attachment of AKT to the membrane is blocked in the inactive PH-in conformation, thus keeping the kinase in the cytosol and preventing its phosphorylation. AKT competitive inhibitors on the other hand are hypothesised to bind to the active PH-out conformation thus triggering recruitment to the membrane, phosphorylation and feedback activation (Wu et al., 2010).

Binding to the allosteric pocket also resulted in inhibitors with unprecedented selectivity over closely related proteins from the AGC family of protein kinases (Lindsay et al., 2005). It is very unlikely that this series of allosteric inhibitors would have been identified by design or focused screening, as opposed to HTS.

Whilst most HTS campaigns are conducted using biochemical assays, large-scale screening can also be performed in a cellular context. Two different scenarios for cellular HTS can be distinguished. In the first scenario, a defined target is screened in a cellular context. A recent example is an assay developed for the screening of ninety thousand compounds to identify inhibitors of the epigenetic regulator histone demethylase JMJD3. A high-content cellular mechanistic imaging assay was used to determine both cellular viability and, as an on-target biomarker, the level of enzyme-dependent demethylation of the H3K27(Me)3 mark produced by an exogenously expressed JMJD3 (Mulji et al., 2012). Such screening of a defined target readout in a cellular context can be particularly advantageous when a biochemical assay is challenging to generate, e.g. because the protein or substrate is not stable or multiple reagents are needed.

In the second scenario (sometimes referred to as chemical genetic screening) a cellular HTS is performed to identify compounds that induce a particular phenotype or inhibit a pathway that is relevant to say, the proliferation or survival of cancer cells. The aim of the chemical screening is thus twofold in this case: Firstly to discover chemical hits and secondly to identify a novel target that is modulated by these hits. However, the second step has proven very challenging even though different target identification technologies can be explored, as reviewed elsewhere (Sato et al., 2010). A recent example is the identification of inhibitors that antagonise WNT signalling (Huang et al., 2009). The authors performed screening using a WNT-responsive luciferase reporter assay. Through a series of experiments, including a quantitative chemical proteomics approach, they then showed that the obtained hits target the WNT pathway through inhibition of the poly-ADP-riboseylating enzymes tankyrase 1 and tankyrase 2, previously not known to be key regulators of the pathway. This is thus an elegant example of a chemical genetic screen that identified chemical hits and at the same time revealed a novel target for the WNT pathway.

The examples described here illustrate the potential of HTS to utilise biochemical and cellular assays to discover chemical compounds with novel and innovative modes of action. However, it is also important to point out that maintaining and screening a large screening collection is resource-intensive (Macarron et al., 2011). In addition, HTS assays have to be carefully optimised to be fit for the screening of hundreds of thousands of compounds, otherwise meaningless data may be produced. The process of screening and following up on the results can be time-consuming, not least because frequent false positives and undesired hits have to be eliminated. False positives often arise from compounds that aggregate (Shoichet, 2006) or interfere with the assay read out (Thorne et al., 2010). Undesired hits often have chemically reactive groups that react in an unspecific fashion with the target. Unfortunately, compounds that are prone to give false positives or represent chemically reactive compounds are still fairly prominent in commercial libraries but ways to identify them have recently been described (Shoichet, 2006; Baell and Holloway, 2010; Thorne et al., 2010). Lastly, the ideal size and composition of a compound screening file is still a matter of discussion that goes beyond the scope of this review (Akella and DeCaprio, 2010; Macarron et al., 2011).

The costs associated with HTS are obviated to some extent if smaller, more focused screening is performed. Focused screening is particularly relevant if the target is a member of a well-studied family of proteins, e.g. kinases or G protein-coupled receptors (GPCRs). The idea behind focused screening is that inhibitors of different members of a protein family, such as protein kinases, often have common molecular features. Focused libraries thus consist of compounds having these features and accordingly frequently show much higher hit rates compared to divers screening collections.

Similarly, cross-profiling of inhibitors generated for one particular kinase, has traditionally been a rich source for hits of other kinases. In the extreme case, one clinical candidate can be explored as an inhibitor of more than one kinase, as in the case of imatinib acting on ABL, PDGR and KIT (Kindler et al., 2004; Matei et al., 2004) and crizotinib inhibiting MET and ALK (Cui et al., 2011).

3.2. Fragment screening

A special application of the screening concept is fragment screening or fragment-based drug discovery (FBDD) (Hajduk and Greer, 2007). In sharp contrast to HTS, in FBDD only a limited number of compounds (typically 1000) of relatively low molecular weight is (~300) are screened with the aim of identifying weakly potent fragment hits (~100 μM).

This approach addresses one of the key challenges of HTS: That is that an HTS hit must engage in several interactions with the target in order to have sufficient potency to be identified as a hit. To be able to engage in these interactions, the hit must feature the required chemical groups, e.g. hydrogen bond donors or acceptors and/or hydrophobic groups, positioned in the correct geometrical arrangement. Particularly for molecular targets that are challenging to drug, the chances are that such a compound is not present even in large screening sets, thus rendering the HTS campaign unsuccessful.

Fragment screening, on the other hand seeks deliberately to identify small fragment hits that are of lower potency. As a result only a few key interactions with the receptor are required, greatly reducing the number of compounds that have to be screened (just as it is easier to find a parking spot for an Austin Mini than for a 1970’s Cadillac). However, the reduced screening effort comes at a price: The initial, weakly potent fragment hit has to be optimized through chemical modification to improve the potency into the range of a hit compound. The optimisation of a fragment will in most cases
require the detailed knowledge of the binding mode to the receptor, e.g. through crystal or NMR structures, thus limiting FBDD to targets where this information can be derived.

A recent example of successful FBDD is the BRAF inhibitor vemurafenib (5), recently approved for the treatment of late-stage melanoma by the FDA (Bollag et al., 2010; Tsai et al., 2008). Screening yielded compound 4 as a low molecular weight fragment hit with weak potency (Figure 6). The co-crystal structure of this compound was then explored through several iterative rounds of design, synthesis and testing ultimately to identify vemurafenib. As can be seen in Figure 6, vemurafenib shows higher molecular weight and potency than 4 but retains its chemical core (depicted in red). This approach, where one fragment serves as starting point and is optimised through the appending of additional chemical groups is often referred to as fragment growing.

An alternative approach to fragment growing is known as fragment linking and has been explored in pioneering work to identify inhibitors of the BCL-2 family of proteins (Oltersdorf et al., 2005; Petros et al., 2005). These proteins are overexpressed in many cancers types and contribute to tumour initiation, progression, survival and in particular to resistance to therapy. Obtaining potent small molecule inhibitors proved difficult using alternative approaches such as HTS owing to the necessity of targeting what is a very challenging protein–protein interaction that does not score highly for conventional druggability. The activity of the BCL-2 protein family is mediated via the binding of the BH3 α-helix of one protein with a large hydrophobic pocket present on binding partners. A high-throughput NMR-based method called ‘SAR by NMR’ was used to screen a fragment library to identify small molecules that bind to the hydrophobic BH3-binding groove of the BCL-2 family member Bcl-XL. In this case, two different fragments (Figure 7, compounds 6 and 7) were identified that bound to distinct but proximal pockets of the protein. These weakly binding fragments were then linked through several rounds of synthesis and design, ultimately resulting in ABT-737 (8). Although of relatively high molecular weight for a conventional small molecule drug, this agent proved to be an extremely potent inhibitor of BCL-XL, BCL-2 and BCL-w that had oral activity and was progressed into clinical trials.

The power of the fragment linking approach lies in the fact that due to the entropic contributions to binding, the linking of two weak binders can lead to a much more potent inhibitor. However, this example also illustrates that linking two fragments in a productive manner and eventually discovering a drug candidate is a challenging endeavour that may require the synthesis and testing of hundreds of compounds and the commitment to the project over a longer than usual period.

### 3.3. Virtual screening

The approach of selecting compounds from large databases by using computational tools rather than physically screening them is generally referred to as virtual screening. Conceptually two different approaches can be followed. Ligand-based approaches select compounds from databases that are in one way or another similar to an already existing inhibitor of the target in question (Schneider, 2010). Structure-based approaches seek to evaluate computationally the fit of compounds to a binding pocket, e.g. by exploring crystal
Virtual screening has obvious advantages over physical screening. It is significantly less resource-intensive and faster. In addition, even compounds that are not available can be evaluated by virtual screening and if found promising, can be bought or synthesised. Millions of compounds can thus be analysed by virtual screening.

It is important to keep in mind, however, that virtual screening is still a relatively coarse filter. That applies particularly to structure-based screening because prediction of binding affinities still remains one of the holy grails of computational chemistry (Schneider, 2010). Nevertheless, several successful examples have been published and recently reviewed (Ripphausen et al., 2010).

### 3.4. Knowledge-based design

Despite the advent of HTS and random screening, intelligent design remains an important component of the medicinal chemist’s tool box. This is illustrated by the recent publication of ABL inhibitors that are effective against the imatinib-resistant mutant in which threonine is mutated to isoleucine at the gatekeeper position known as T315I that is also resistant to second generation inhibitors dasatinib and nilotinib (Chan et al., 2011; Packer and Marais, 2011). Currently available ABL inhibitors have shown impressive efficacy against CML. However, the T315I mutation frequently emerges and confers resistance to all of them, partially through steric interference of the mutant isoleucine with the inhibitor. The researchers (Chan et al., 2011) recognised that potent inhibition of the kinase can be achieved without binding to the space that is altered through the T315I mutant, particularly if an inactive (DFG out) conformation of the kinase is explored (Packer and Marais, 2011). The design started with substituted aminopyrazoles that were already known to bind to a region distant to the mutant gatekeeper (Figure 8, red part of compound 10). The same region of the ATP pocket is addressed by other ABL inhibitors, e.g. imatinib (Figure 8, red part of compound 9). Chan et al. took this concept further by recognising that the aminopyrazole moiety can also be explored to engage in additional interactions distant from the gatekeeper in particular with Arg386 and Glu282 of ABL1 (Figure 8, blue part of compound 9). These interactions also further stabilise the inactive conformation of the kinase. As a first proof of concept, compound 10 showed reasonable potency against the T315I mutant and wildtype kinase. To further improve the activity, the researchers then extended the compound into the adenine binding pocket but in a manner that is not obstructed by the gatekeeper mutation, thus leading to very potent inhibitors of both wildtype and T315I mutant ABL, in particular the clinical candidate DCC-2036 (Figure 8). Interestingly, this appended moiety resembles the corresponding group in the marketed multi-kinase inhibitor sorafenib (Macarron et al., 2011). The work thus represents an elegant example how different and previously known components of a kinase inhibitor can be explored and extended to engage in an additional interaction, leading to an inhibitor with innovative properties.

### 3.5. Hit discovery strategy considerations

As described above, several approaches are available for hit finding and can be used either individually or in combination. Particularly when exploring challenging targets with low predicted druggability, it is often advisable to follow more than just one approach so as to increase the chances of finding hits that can be progressed by medicinal chemistry.

Whilst following different hit discovery approaches increases the likelihood of finding hit matter, it is also time- and resource-intensive. For druggable targets, e.g. the ATP pocket of a kinase, the extra effort might not be warranted and the screening can be limited to a focused collection to save time and resources. Designing a hit strategy is thus a question of carefully reviewing what is known about the target to assess which approaches will be followed so as to be as confident as possible that hits can be obtained.

### 3.6. Assessing the quality of the hit matter

The quality of the chemical hit matter will to a large extent determine how fast and efficiently it can be optimised to a preclinical candidate. A hit of high quality already shows

---

**Figure 8** — Design of ABL inhibitors active against the T315I gatekeeper mutant that is resistant to imatinib, dasatinib and nilotinib. In the first step, it was shown that potent inhibition of ABL can be achieved by addressing a region of the ATP pocket that is not altered by the T315I mutant. The red part of 10 explores a region occupied in a similar manner by imatinib (9) through the moiety outlined in red. Further potency is gained through adding another substituent (outlined in blue) to 9 that interacts with Arg386 and Glu282. The clinical candidate DCC-2036 (11) was then obtained by extending the compound further towards the 315 residue in a manner that is not obstructed by the T315I mutation (Chan et al., 2011).
many of the properties desired for a lead or even the future drug candidate.

However, it is equally important to consider properties generally referred to as lead-likeness (Hann and Oprea, 2004), in particular molecular weight and lipophilicity (Hann, 2011; Nadin et al., 2012). When comparing two hits with equal potency, it is generally the case that the one with lower molecular weight has an advantage. This due to the observation that both molecular weight and lipophilicity often increase during the optimisation of a hit towards a preclinical candidate. In addition, it has been shown – as a rule of thumb based on past experience – that oral drug candidates with high molecular weight (> 550 Da) and excessive lipophilicity have a higher chance of failing during development (Lipinski et al., 1997). When relationships between physicochemical drug properties and toxicity were inferred from a data set comprising animal in vivo tolerability studies on 245 preclinical Pfizer compounds, an increased likelihood of toxic events was found for less polar, more lipophilic compounds (Hughes et al., 2008). A smaller hit thus gives the medicinal chemist more room to add substituents in order to dial in properties like potency and selectivity. Increasingly, ligand efficiency indices are used to evaluate and compare screening hits which represent the activity against the target normalised to the molecular weight (Abad-Zapatero and Metz, 2005; Hopkins et al., 2004; Leeson and Springthorpe, 2007).

4. Lead optimisation to select a clinical candidate

The optimisation of a lead structure through iterative rounds of medicinal chemistry design, synthesis and testing is often referred to as multi-parameter or multi-objective optimisation (Ekins et al., 2010). It requires the simultaneous optimisation of several compound properties such as potency, selectivity, tolerability, bioavailability and metabolic stability to generate a safe and efficacious drug. The particular multidimensional challenge can be illustrated by comparing it to solving the Rubik’s cube. It is easy to solve one or two sides of the cube but so much more difficult to get them all right. But even this analogy only illustrates the medicinal chemist’s challenge (and sometimes dilemma) to some degree because in the case of Rubik’s cube the outcome of every move can be precisely predicted, whereas this is not the case in drug design, with many properties being inter-related and somewhat unpredictable. Thus it is not uncommon that despite a strong hypothesis to introduce certain changes into the chemical structure of a hit compound, the biological results for the new compounds are dramatically different from what was expected.

Having said that, the art of multidimensional optimisation has evolved considerably in recent years. In particular, our advanced understanding of factors influencing pharmacokinetics, together with the widespread use of structural biology and computational tools, now help us greatly to define design hypotheses concerning how properties can be optimised in a rational fashion (Ekins et al., 2010; Lusher et al., 2011; Plowright et al., 2012). Figure 9 describes the example of multi-objective optimization of PI3K inhibitors (Folkes et al., 2008; Raynaud et al., 2009). Starting from a lead structure with promising biochemical and cellular potency (PI-103, an inhibitor which is still used as a chemical tool) the selectivity, solubility and bioavailability as well as other properties were optimised to finally yield GDC-0941 that is currently in Phase II clinical trials.

A detailed discussion of the strategies and tactics employed by medicinal chemists goes far beyond the scope of this article. Instead we will comment in the following sections on selected technologies and strategies that are relevant to the chemical optimisation of cancer drugs, specifically slow off-rate or covalently binding agents, the choice of preclinical cancer models for drug discovery, and approaches to preclinical safety testing.

![Figure 9](https://example.com/figure9.png)

**Figure 9** – An illustration of multidimensional optimisation leading to GDC-0941, a pan-Class I PI3K inhibitor that is currently in Phase II clinical trials (Folkes et al., 2008; Raynaud et al., 2009). The lead compound PI-103 already showed promising biochemical and cellular activity (and is used as a chemical tool) but also exhibited poor physicochemical and pharmacokinetic properties. These were optimised through several rounds of synthesis and testing. Selectivity for PI3K versus DNA-PK was also improved. The properties which initially did not fulfil our criteria are marked in red.
4.1. Drugs with prolonged residence time at the target and covalently binding agents

Most drugs have to modulate their target in a sustained fashion to be efficacious. This can be achieved by optimising the pharmacokinetic properties to achieve prolonged exposure to the drug. Alternatively, sustained target modulation can arise from a slow off-rate such that a drug will take several hours to dissociate from its target even when the compound is cleared from the organism. Whilst optimisation of pharmacokinetic properties has traditionally been a key focus of medicinal chemists, the beneficial effect of extended residence time of the drug at the target has only recently received wider attention (Copeland et al., 2006). This might partially be due to the availability and wider use of surface plasmon resonance (SPR) that enables measurements of off-rates with reasonable throughput (Rich et al., 2008).

A particularly good example is the dual EGFR/ERBB2 inhibitor lapatinib (Figure 10). It has been shown to have a very slow dissociative half-life of 300 min compared to less than 10 min for gefitinib and erlotinib, two other approved EGFR inhibitors (Wood et al., 2004). Interestingly, the available co-crystal structure of lapatinib shows that the drug binds to an unusual inactive conformation of EGFR. Based on this structure it has been hypothesised that dissociation from the protein requires a slow conformational change of EGFR, thus significantly prolonging the off-rate (Wood et al., 2004). The long half-life may potentiate drug efficacy by prolonging drug-induced down-regulation of receptor-mediated tyrosine kinase activity (Gilmer et al., 2008).

An obvious way to prolong the off-rate is to design a drug that engages in a covalent bond with the target leading to irreversible or slowly reversible binding (Singh et al., 2011). Examples are the pan-ERBB inhibitors neratinib and afatinib currently undergoing clinical trials. These drugs carry a reactive group (a so-called Michael acceptor) and engage in a covalent bond with Cys797 of the EGFR protein (Figure 10). Crucially, neratinib and afatinib have been shown to be active against the T790M/L858R double mutant that renders most reversible inhibitors ineffective. Covalent inhibitors in general have been proposed to be more resilient towards target mutation (Singh et al., 2011). Several examples of covalent drugs as well as the challenges of tuning the reactivity of the chemically reactive group to avoid toxicity have been reviewed elsewhere (Potashman and Duggan, 2009; Singh et al., 2011; Smith et al., 2008). The recent co-crystal structure of abiraterone shows that a ‘covalent’ interaction is formed between the pyridine moiety and the haem iron present in the active site of the CYP171A1 target (DeVore and Scott, 2012).

4.2. Preclinical cancer models

In therapeutic development a major hurdle is translating efficacy demonstrated during the drug discovery phase into efficacy in the clinic, as there are often discrepancies between drug efficacy demonstrated in the current preclinical experimental models and final efficacy, or lack thereof, in patients. Preclinical models need to take account both of the molecular nature of the target and also of how the chemical compound will behave. Different models will be required for compounds targeting genetic dependency, as compared to those exploiting host-tumour interaction through stromal or hormonal signalling or compounds targeting non-oncogenic addiction, perhaps through synthetic lethal interactions (Caponigro and Sellers, 2011).

During lead optimisation, tumour cell lines grown in vitro culture are frequently used as the first line of study as cell proliferation, apoptosis and other cellular endpoints can easily be measured using different technologies in high-throughput formats as required. Cancer cell line-based platforms have recently been reviewed (Sharma et al., 2010). As already mentioned above, it is valuable to use proliferation/apoptosis assays in combination with mechanism-based cell assays for target engagement and pathway inhibition. Using both types of assays together ensures that the target is modulated at concentrations similar to or lower than the levels required for cell growth inhibition and thus that the growth inhibitory effect is achieved by the anticipated mechanism. Confidence is increased by seeing a correlation between cancer cell growth inhibition and target/pathway modulation in the structure-activity relationships across a range of analogues.

There are, however, some caveats with regard to the use of tissue culture cell lines. In vitro culture of cancer cell lines can result in substantial phenotypic, genetic, and epigenetic alterations induced by artificial environment of cell culture,
which may not reflect the biology of original tumours in situ. Moreover, prolonged culture of tumour cells will inevitably result in genetic drift with time. In addition, cancer cells from some tumour types or genetic backgrounds fail to grow under in vitro culture conditions and will be under-represented in cell line panels. As adherent cell growth on a plastic dish has limitations and may not really reflect tumour biology, 3-dimensional growth conditions are now being more frequently used as in vitro models that may be more representative of tumour biology in the host organism. Cells are either grown as spheroids in suspension, or in matrices that may represent the extra-cellular environment of the tumour. However, it is currently quite challenging to use these 3-dimensional models in sufficiently high throughput, although methodologies to increase throughput are in development.

Compounds that show promising activity in cell-based assays will progress to in vivo animal studies, provided that their ADME properties are suitable for such studies. Human tumour xenografts have been used extensively and when selected to exhibit the relevant molecular characteristics and pathogenic drivers can mimic many of the features and response to targeted drugs of the corresponding human cancers (Caponigro and Sellers, 2011; Ocana et al., 2011; Workman et al., 2010). However, the above reviews also highlight limitations of human tumour xenografts: They do not replicate human stromal–tumour cell interactions, they grow much faster than human tumours in patients and hence can be more sensitive to targeted drugs. The human tumour xenograft approach also generally relies on the use of established cancer cell lines with all their limitations that are described above. Finally, the tumours are usually not implanted at their native site. Recently, the latter two shortcomings have been mitigated by using orthotopic models in which the tumour is transplanted into the correct organ site and by relying on early passages of fresh biopsies from patients to maximize retention of the original patient characteristics (Caponigro and Sellers, 2011; Ocana et al., 2011; Workman et al., 2010). We believe that despite the limitations listed above, molecularly relevant human tumour xenografts will retain their role as workhorses in drug discovery since they allow testing of agents in a variety of genetically-defined human tumours with reasonable throughput. Sufficient throughput is critical once an advanced stage of lead optimisation has been reached and often 50–100 efficacy experiments per year may have to be performed. Our view is that the key factors in using human tumour xenografts successfully are to ensure molecular relevance and in addition to avoid overly optimistic interpretation of the outcome. Thus it seems likely that a higher hurdle for efficacy should be used than is often currently the case in the selection of clinical development candidates.

Referred to briefly earlier, GEMMs have now emerged as in vivo models for cancer drugs and are complimentary to the use of human tumour xenografts (Caponigro and Sellers, 2011; Ocana et al., 2011; Workman et al., 2010). In GEMMs, the tumours are initiated by inducing a specific genetic lesion, e.g. through activation or overexpression of oncogenes such as KRAS or MYC or via depletion of tumour suppressors, ideally in a tissue- and time-dependant manner. GEMMs thus have the advantage that the tumours arise in a more natural fashion and exist in their natural, albeit murine environment. In addition, they allow the investigation of drug effects on tumour development, as well as studies on established cancers. Crucially some GEMMs have shown patterns of sensitivity to chemotherapeutic agents and development of resistance that are similar to their human tumour counterparts.

On the other hand, GEMMs also have some limitations. From the drug discovery viewpoint, as distinct from more basic research, the logistics of maintaining sufficient stocks of tumour-bearing mice can be challenging, particularly for models with long tumour latencies which in most cases renders it impossible to use them routinely to support make-and-test cycles in the context of lead optimisation (Politi and Pao, 2011). They may also suffer from significant protein sequence differences between the human and the murine target and will likely also show limited genetic aberrations whereas human tumours generally exhibit very large numbers of mutations, although more sophisticated cancer gene-combination models are now emerging. Perhaps the most valuable role for GEMMs in molecularly targeted drug discovery is to show proof of concept for in vivo animal model activity in the context of clearly defined and clinically relevant genetic abnormalities. Of note in the present context, a promising and practical compromise for drug discovery is the use of transplantable lines derived from GEMMs.

The question whether human tumour xenografts or GEMMs are more predictive for the response of human cancers continues to be intensely debated. We believe that the predictive value of neither model is fully understood and that they are of complimentary value for drug discovery. Human tumour xenografts (and transplantable GEMM tumours) are the better suited for routine in vivo testing and GEMMs can be used to explore proof of concept once an advanced candidate has been identified. There are now many examples where both models give similar results to a targeted agent when the pathogenic driver is the same. Where this is seen it can build additional confidence in the target hypothesis.

4.3. Preclinical safety testing

Once a preclinical candidate has been identified, sufficient preclinical data have to be generated to support a clinical trial. A new guidance from ICH (International conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use) topic (S9) provides recommendations for non-clinical evaluations to support clinical trials of cancer drugs. This new ICH topic has recently been reviewed (Joness and Jones, 2011). For safety testing of small molecule drug candidates, generally the use of one rodent and one non-rodent species is recommended but some exceptions exist where a single species may be possible, e.g. with cytotoxics. However, many organisations perform the more extensive studies required for a non-cancer agent (under the ICH topic M3).

5. Biology-led clinical trials of targeted drugs

It has become increasingly apparent that the traditional one-size-fits all clinical trial paradigm that has been successful
for many cytotoxic drugs is no longer appropriate for the generation of molecular cancer therapeutics and must be drastically changed in order to achieve full and rapid benefit from targeted and personalised cancer therapies, as reviewed in detail elsewhere (de Bono and Ashworth, 2010; Yap et al., 2010; Yap and Workman, 2012).

Clinical trials for targeted drugs should be led by the biology and the clinical hypothesis. These trials should be hypothesis-testing and biomarker-led. They should be designed to test a strong scientific hypothesis, for example that a particular drug acting on a specific molecular target is efficacious in patients with a particular type of genetic aberration or certain molecular feature (de Bono and Ashworth, 2010; Yap et al., 2010; Yap and Workman, 2012). Two key scientific advances are fundamental to enable hypothesis-driven trials, ideally starting as early as the first-in-human Phase I study. The first of these is the continuing shift away from patient selection based on the anatomical site and histological classification of the cancer to patient stratification based on genomic aberrations and other relevant molecular characteristics. The second is the extensive use of biomarkers (de Bono and Ashworth, 2010; Tan et al., 2009; Yap et al., 2010; Yap and Workman, 2012).

We introduced the Pharmacological Audit Trail (PhAT; Figure 11) to provide both a conceptual framework and practical guidance for defining a biomarker-driven early clinical trial strategy that enables the rational evaluation of drugs and the testing of the biological hypothesis behind new molecular cancer therapeutics (Workman, 2003a, 2003b; Tan et al., 2009; Yap et al., 2010). Moreover, the PhAT provides the basis for making key decisions as the drug advances from preclinical studies through the different clinical stages. Examples of these key decisions are: 1) Is the right patient with the right molecular pathology being given the drug? and 2) Is the optimal dose range and schedule appropriate to deliver appropriate target engagement, pathway modulation, biological effect and therapeutic benefit? Use of the PhAT also empowers making logical decisions concerning whether to continue with the clinical development of the drug or to terminate the drug development programme. And if the latter, whether seeking an alternative, improved drug would be worthwhile or whether this would be inappropriate because the biological hypothesis is flawed.

Importantly, by validating, at least in part, the clinical hypothesis early on, the PhAT addresses key clinical development risks at an early stage and thus minimises the chances of failing in late clinical trials due to lack of efficacy — a major element in the ‘Valley of Death’ problem. We have reviewed the PhAT in detail elsewhere (Yap et al., 2010) and here we will highlight some key elements of it.
An important feature of the PhAT is the seamless transition from the drug discovery phase into clinical trials. Pharmacokinetic/pharmacodynamic and related efficacy data from suitable preclinical models are for example used to generate a dose regimen hypothesis that is then explored for the initial Phase I trials with targets for drug exposure and target modulation.

Another key feature of the PhAT is the molecular stratification of patients even in the initial Phase I setting. This allows, subject to the achievement of appropriate target and pathway modulation, the potential observation of indications of response at the earliest possible time, as was the case for many of the molecularly targeted agents described earlier in this review, including inhibitors of BRAF. Such audit trail data provide a more rational basis to refine the dose regimen rather than simply basing it on Maximum Tolerated Dose. It should be noted, however, there is often concern about selecting a pharmacodynamically effective dose that is below the Maximum Tolerated Dose, especially where very profound (e.g. >90%) target inhibition is required. Nevertheless, demonstration of target modulation is essential. Moreover, there is now accumulating evidence that non-randomised Phase I trials are of limited use for targeted cancer drugs and we propose that detection of early responses in extended and stratified Phase I trials enable the direct transition into randomised Phase II/III studies for those drugs that show obvious (but not marginal) clinical benefit. Moreover, early regulatory approval of drugs showing strong activity in Phase II may also be appropriate, subject to stringent follow-up.

Examples of drugs tested in stratified Phase I studies with use of elements of the PhAT are the ALK inhibitor crizotinib in patients with NSCLC who have tumours with the pathogenic EML4–ALK translocation (Kwak et al., 2010), vemurafenib with driver V600E BRAF mutations in their melanomas (Chapman et al., 2011) and the PARP inhibitor olaparib in patients with cancer who have BRCA1 or BRCA2 mutations which result in synthetic lethality (Fong et al., 2010).

The upfront use and testing of putative predictive biomarkers in early clinical trials can potentially minimize the need for retrospective analysis of patient subgroups in later stage trials that would otherwise be carried out in unselected populations. Notable examples of subgroup dredging are clinical studies of EGFR targeted drugs where analysis of tumour tissue led to the discovery that patients who had wild-type KRAS and mutant EGFR had significantly increased benefit (Allegra et al., 2009). Such effects could however be missed with such a retrospective approach and prospective stratification is to be preferred. Indeed it is often quoted that trastuzumab would not have shown activity if tested in non-selected breast cancer patients, rather than the sensitive ERBB2-positive group. These examples emphasise the importance of a priori drug evaluation in the appropriate molecular contexts early on in the drug development process, starting from Phase I/II clinical trials. Early implementation of biomarkers also facilitates their subsequent qualification to meet required regulatory standards.

An important final advantage developed in the updated PhAT is that the detection of molecular alterations that are induced by the treatment enables the generation of hypotheses to explain the acquisition of resistance and to suggest how resistant tumours can be drugged.

Of course critical to the PhAT is the availability of validated, robust biomarkers, especially for patient selection, demonstration of target engagement and assessment of response (de Bono and Ashworth, 2010; Tan et al., 2009; Yap et al., 2010; Yap and Workman, 2012). These biomarkers need to be discovered through basic research and clinical molecular pathology studies in patient tumour samples, and may also be developed and used during the preclinical drug discovery phase. Such biomarker studies can encompass analysis of a single specific biomarker in a tumour biopsy specimen all the way through to whole genome sequencing. Genomic, transcriptomic, epigenomic, proteomic and metabolomic studies are all becoming much more common. Use of circulating tumour cells and circulating DNA have advantages and non-invasive functional imaging approaches are also important, particularly for assessing multiple tumour deposits around the body. Validation and qualification of biomarkers is important and cost-effectiveness and quality control become key factors in the routine use of biomarkers for patient selection and decision-making for personalised medicine.

In summary, the PhAT provides a conceptual framework and practical guidance with which to test new targeted drugs under the new molecularly targeted clinical paradigm. The PhAT in particular aims to make more rapid, efficient and logical the currently slow and costly clinical development phase and in particular to address key questions and risks as early as possible so as to minimise the chance of expensive failure in late stage clinical trials.

6. Summary and concluding remarks

Many new molecularly targeted cancer drugs have gained regulatory approval over the last few years and have improved and extended the lives of a large number of patients. However, the discovery and development of new targeted drugs is still frustratingly slow with high failure rates, particularly in late stage clinical trials, with the eventual development of drug resistance remaining as a constant challenge.

On the other hand, our understanding of the genetic and molecular basis of cancer initiation and malignant progression has improved enormously, especially as a result of genome sequencing and other large-scale profiling methods. This has opened up incredible opportunities for selective therapeutic targeting to exploit addictions, dependencies and vulnerabilities in cancer cells.

In addition, evolving scientific and technological breakthroughs have enabled faster and more efficient drug discovery, together with more sophisticated and biology-led clinical trials. We believe that the progress that has been made in the last decade and ongoing large scale genetic and molecular characterisation of cancer will lead to a further significant acceleration in the process of bringing new personalized drugs to cancer patients worldwide.

Many of the advances and recommendations described in this review should shorten the timescale from the initial biological discovery to drug approval and decrease attrition. They therefore directly address the pernicious problem of the
Valley of Death that still separates basic research discoveries from therapeutic innovation.

In addition to the scientific and technical advances discussed here, there are organizational and cultural improvements that need to be made to help us bridge the Valley of Death (Bornstein and Licinio, 2011). In particular there is a need for closer working relationships and better collaboration between the academic, healthcare, regulatory and industrial sectors in order to help drive the enormous advances emerging from genetic and biological research into medical and commercial benefits for society. There is a clear requirement for more productive partnerships that will require ongoing changes in behaviour and reward systems. Academia needs to encourage and reward entrepreneurial innovation and team science. Industry needs to work more closely and effectively with academia. One of the areas in which progress is being made is the establishment of academic drug discovery centres, of which our own is one example. When appropriately scaled, staffed and supported, these groups are able to take on what are initially seen by pharma as unacceptably high risk drug targets and to de-risk them through some degree of proof-of-concept before subsequently partnering the project with industry to progress rapidly to the clinic and the delivery of personalised cancer treatment. By working together as a collaborative international community, harnessing the scientific power and technical capabilities of both public and private sectors, it will be possible to enhance significantly our ability to exploit our increasingly sophisticated understanding of the complex and dynamically evolving genomes and network biology configurations of human cancers so as to achieve further significant gains in tumour responsiveness and overall survival – bringing major benefits to all cancer patients worldwide and to society as a whole.

Acknowledgements

We thank our colleagues in the Cancer Research UK Cancer Therapeutics Unit for valuable discussions, especially Professor Julian Blagg. We thank Ann Ford and Val Cornwall for administrative assistance. The authors gratefully acknowledge core programmatic support from Cancer Research UK (grant number C309/AB274/) and also acknowledge NHS funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and Royal Marsden NHS Foundation Trust. The ICR is a Cancer Research UK Life Fellow. We apologize to the authors of many excellent publications that could not be cited due to space considerations.

Conflict of interest

The authors declare a conflict of interest. Swen Hoelder, Paul Clarke and Paul Workman are employees of The Institute of Cancer Research, which has a commercial interest in the development of various anticancer agents, including inhibitors of CYP171A1, PI3 kinase, HSP90, AKT/PKB, CDKs, Aurora/FLT3, histone deacetylases and WNT, and operates a rewards-to-inventors scheme. The authors are involved in a range of commercial collaborations (see www.icr.ac.uk). Paul Workman is or has been a founder of, consultant to, and Scientific Advisory Board (SAB) member of Piramed Pharma (acquired by Roche) and Chroma Therapeutics. He is also on the SAB of Astex Pharmaceuticals, Wilex and Nextech Invest.

R E F E R E N C E S


Fong, P.C., Yap, T.A., Boss, D.S., Carden, C.P., Mergui-Roelvink, M., Gourley, C., Del Greve, J., Lubinski, J., Shanley, S.,Messiou, C.,


