Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Project Considerations

5.1 Introduction

There’s really no such thing as a “typical” project. No two are alike. Each will have its own uncertainties, issues, and challenges that need to be addressed in a slightly different way. Taking a one-size-fits-all approach to projects would be an exercise in futility, while there aren’t enough pages in this book to describe all the variations in project types that can and do occur. But projects do tend to have some commonalities that allow most of them to be divided up into several bins, if the bins are made large enough. Projects can be divvied up between these bins based on both the initial strategy (screening) and the kind of drug they’re expected to result in (market strategy). Needless to say, these particular alphas and omegas will influence many of the letters in between.

As a scientist, project team member, or project proposer, a basic knowledge of these project types can help you fit the pieces of the puzzle together to answer questions like, “Why are we emphasizing these factors and not those?,” “Why the short timelines?,” “Why are we hung up at this stage?,” “Why are we trying to make a drug when there are similar ones already on the market?,” or even, “Why aren’t we developing this perfectly good compound?” Knowledge is power, and being able to answer these questions is not only intellectually satisfying but also enables the researcher to do a better, more focused job as well.

Keeping this in mind we find that, much like Gaul, the landscape of drug discovery targets can be divided into three parts, as shown in Box 5.1. Each of these has its own strengths and weaknesses, advantages and disadvantages, and risks and benefits. Projects of these different types will have a different flavors and tend to run somewhat different courses. Figure 5.1 shows these schematically.

Box 5.1 Categories of New Drug Discovery Projects

- Established targets for which clinically effective drugs exist
- Novel targets which so far lack clinical validation
- Undetermined targets uncovered through phenotype screening
5.2 Established Targets

Doing drug discovery research on an established target implies a number of things of great importance to the resulting project, as described in Box 5.2. In this and the other boxes, what many would consider as the most important advantages and disadvantages have been italicized.

**Target validation** (TV, discussed in more detail later), that is, systematically building up evidence that the interaction of a proposed target with an agent like a small molecule or a

---

### Box 5.2 Advantages and Disadvantages of Working on Established Targets

**Advantages**

- Target ID and target (but not compound) validation can be skipped
- Known compounds and assays are available as tools
- *Higher than average probability of finding an efficacious agent*

**Disadvantages**

- Crowded market space
- Work needs to be fast: competition is far ahead
- Drug needs to display a clinical advantage to be viable
- Market may disappear due to newer targets or mechanistic toxicity
- Formularies or price regulation may limit profitability
- *Crowded IP space*
monoclonal antibody (mAB) can have a therapeutic effect in a human disease state, is often far from trivial as we’ll see. It represents an ongoing, labor-intensive process that can be rate-limiting, but is necessary to avoid disastrous losses in time, money, and manpower that occur when a clinical drug fails to work because its target is not at all crucial to disease progression or symptoms. Target identification (TID), in this case the process whereby a new target is first proposed based on some minimal evidence, is less cumbersome but still requires devoted resources. The beauty of working on a project based upon an already established target is that both of these steps can be skipped.

Examples of established targets would be HMG-CoA reductase inhibitors for cholesterol lowering (e.g. atorvastatin, Figure 5.2), and PPARγ agonists for type II diabetes (e.g. rosiglitazone, Figure 5.3). A company going after either of these targets as of today would have the advantage of not having to do the expensive, messy work of first identifying and then validating it. This has already been done for them on somebody else’s dime (or, more likely, hundreds of millions of dollars). Their new program could therefore begin with compound screening and as it progressed they wouldn’t have to deal with the annoying question that people involved in financing such projects often ask, “How can we be sure this is a valid target?”

But the companies that originally worked out the preclinical and clinical TVs didn’t do it out of humanitarian benevolence or to confer good assays and tools to future competitors. Instead, by definition they already have an approved drug or advanced clinical candidate well on its way to approval from which they hope to recoup the costs of this project as well as their less successful research and make a shareholder-pleasing profit to boot. And with the “blockbuster mentality” currently no less ensconced in big pharma management than it is in Hollywood, one can be sure that such a target is likely to draw whole flocks of companies to that particular watering hole. Trying to find enough room to squeeze in and enjoy may not be easy. After the first-in-class drug, a plethora of competing drugs will probably follow. And, as a recent study has shown, that profitable period of market exclusivity, when a single first-in-class approved drug exists for a given target, is becoming shorter and shorter, down to an average of 1.8 years in the late 1990s\(^1\) and probably to only a matter of months today.

![Mevacor (lovastatin)](image1.png) ![Lipitor (atorvastatin)](image2.png)

**Figure 5.2** Structures of two HMG-CoA reductase inhibitors. Lipitor is actually the calcium salt of the compound depicted.
That same study uses a statistical analysis of drug approvals by date to demonstrate the eye-opening fact that in recent years second- and third-in-class drugs were more likely to have been competitors vying for initial approval with the eventual first-in-class winner, than copycat compounds trying to shoehorn into its established market later. These competitors just didn’t happen to win first place in the race, but they were very much in it, having started before the target was an “established” one. All of which begs the question of what exactly a “me-too” drug might be these days. Dr Albert Wertheimer of Temple University uses the following scenario: “Three different labs all start out within a month of each other and one is more successful with FDA or they had fewer questions or their CRO was more efficient or was luckier in getting more patients faster. So one of them got FDA approval first and we say, ‘Ah, that’s the originator!’ But in France the one that came in second here could have been first there, and the American ‘innovator’ here is ‘me-too’ there. So what’s a ‘me-too’?”

Most likely, in the time it takes for a brand new drug project based on an established target to go from lead ID to approved drug, probably at least 8 years under the most favorable circumstances, any target for a major market which had Phase II clinical validation when it began will already be treated by several approved drugs. So starting that late raises a new annoying but incisive question, “Do you really want to bring the fourth- or fifth-in-class drug to market?”

Now, most researchers whom I’ve ever met treasure innovation and are highly competitive. To them this prospect seems about as exciting as entering a race where the first three winners were announced before the gun was even fired. They won’t be the only ones wondering how such a drug could get any significant market share when facing such competition. The research chemist, particularly if he ends up working with a close structural analog of the marketed drug, which is frequently the case in this kind of project, will inevitably think of it as a “me-too” project, while management will likely use a more elegant term like innovative improvement. Fast follower is yet another tag sometimes used for this type of project, although, as mentioned above, “fast” here only applies on a geological time scale. There’s a certain element of truth to all of these terms, but neither they nor any other two-word phrases can explain all that’s really going on with such a project.

In pursuing a clinically established target, a company has made the corporate decision to minimize risk, at least the risk of target failure, which is responsible for a significant fraction of clinical drug failures. “Risk” here is relative. It’s not quite the same as making the decision to buy a bank CD versus buying Internet stocks as an investment. In the drug discovery business there is no truly low risk strategy. Everyone involved is something of a gambler. But in this approach, everyone is confident that at least the target is a good one and it has been shown to lead to a salable therapeutic agent. Many useful, profitable drugs have come out of this approach. And this approach per se doesn’t imply that chemists will be bored to tears generating compounds nearly identical to a known drug but falling into an IP loophole: patent poaching (another two-word epitaph). At least in theory, an established target can be approached with molecules resembling established drugs or it can be targeted with compounds looking nothing like any of them. The target doesn’t necessarily define the chemistry, but, as we’ll see, since using a known drug or pharmacophore as a starting point shortens the time to having an identifiable lead, it sometimes does for this type of project.

Having eliminated the TID and TV steps in a project based on an established target, keep in mind that compound validation is not dispensed with. The target may be a good one, but
compounds can fail for plenty of other reasons (poor PK, toxicity, etc.). And although it might seem that, since the two stages of the drug discovery process have been amputated from the scheme, fewer resources should be required for this type of project, the need to work fast when starting from behind means that this isn’t always so. All the full time employees (FTEs) and the budget saved in eliminating the first two steps and maybe even more could be required to get the resulting drug to market while there’s still time. But, if the competition is so far ahead—assuming one isn’t competing with oneself in a second generation project—that it won’t even be remotely possible to beat them to the market, how can the company hope to sell the drug? Why should any physician prescribe it or any patient take it?

Because of some clinical advantage. It needs to have an advantage over existing drugs that will give people a reason to use it. This can take the form of greater efficacy (clinical studies prove that it just works better), better dosing regimen (say once daily as opposed to several times a day), or reduced toxicity (causing fewer or less severe side effects). What do these things mean to the medicinal chemist? They mean that while synthesizing and selecting the clinical candidate the emphasis will be on superiority to the existing drug in terms of some important and hopefully measurable property such as in vivo half-life in PK studies, where superiority could mean once (q.d.) versus thrice (t.i.d.) daily dosing, or selectivity against known antitargets implicated in observed toxicity for the existing drug. If no such advantage is found and the best that can be obtained is something equivalent to what’s already out there or soon will be, the company may decide not to spend many millions of dollars required to try and bring it to the market. But the catch is that often advantages or disadvantages may not become evident until Phase II or Phase III clinical trials or even later, so who can really say short of that? A candidate with no obvious advantage might still be moved forward, but the project and the drug become a much harder sell, both internally and externally.

With a new drug that’s only “just as good as” an old one, the best pharmaceutical detailers (sales reps) in the business will not be able to significantly “move the product” regardless of how many free samples they give away. But, if the existing drug causes some particular, significant toxicity in patients and yours doesn’t, you stand a good chance of taking away a large amount of their market share. Or if the existing drug, like the oral anticoagulant warfarin, is difficult to dose correctly due to a narrow therapeutic window and lots of individual variability while your new vitamin K reductase inhibitor for some reason isn’t, all other things being equal, your drug will likely take over that market.

A couple of myths that exist in the minds of many researchers need to be exploded here. The first is that “first-in-class” means “best-in-class” or at least “as-good-as-any-other-in-class.” In some ways coming up with a new drug isn’t terribly different from coming up with any other new product, say a new car. Does Henry Ford’s Model T outperform the car in your driveway? Does Ford still sell a lot of them? The new drug business is an evolutionary process both in terms of targets for a given disease and drugs to treat any one of those targets, which means that “first-in-class” can eventually even become “worst-in-class.” Prescription drugs, like anything else, are imperfect and there’s plenty of room for improvement.

An example that can provide hope to those who labor to come up with a new drug for an established target is that of Lipitor (atorvastatin, Figure 5.2). Currently the #1 prescription drug sold anywhere, with over $13 billion in annual worldwide sales, it was the fifth statin to be approved by the FDA, making it to market nine years after the first-in-class compound, Mevacor (lovastatin, Figure 5.2). At least one major study has demonstrated improved
effects on LDL-lowering for atorvastatin relative to other statins, and it has a low incidence of a nasty but fortunately rare side effect of compounds in this class, rhabdomyolysis. Other factors, including the marketing might of Pfizer are, no doubt, a big part of Lipitor’s success story, but the point is that working on yet another drug for an established target is sometimes a legitimate formula for success, both from the point of view of the corporate executives and of the patients who take it.

It can even be exciting for the project’s chemists. Consider the structures of lovastatin and atorvastatin, which are shown in Figure 5.2. Other than the lactone-containing side chain of the former, which upon hydrolysis generates the important pharmacophore contained in the side chain of the latter as well as all other drugs in this class (see Appendix, “The Periodic Table of Drugs”), there’s really not much similarity. Lovastatin looks like, and is, a fungal metabolite, while atorvastatin doesn’t look like a natural product because it isn’t. Instead, it was the result of a determined synthetic effort by a group of chemists at Warner-Lambert who worked out an interesting, commercial scale enantioselective synthesis.6 Here, indeed we have “innovative improvement” and one doubts that any of the chemists involved found this project trivial or boring.

Differences in the molecular structures of the drugs are key to the second myth that needs exploding. This is that finding a new drug for a given target that’s structurally quite different from known drugs for this target provides some sort of advantage for the drug. As a synthetic chemist myself, I’ve never been fond of true “me-too” or “patent poaching” projects aimed at finding a “hole” in an existing patent via something like a one atom replacement. Only desperate corporate executives transfixed by Gantt charts could love such a project. Chemists would rather make novel—meaning highly dissimilar—structures. But dissimilar structures per se convey no advantage in the clinic. No physician is going to ask a patient, “Wouldn’t you rather take an oxazole than a pyrimidine?” Once a drug has reached clinical testing, structure is largely irrelevant. Only two things really matter: safety and efficacy. That’s where your compound needs to demonstrate its uniqueness. If the oxazole is less likely than the pyrimidine to be biotransformed into a toxic metabolite, allowing the question to become, “Wouldn’t you rather take a safer compound than a more dangerous one?” then the clinical and marketing advantages are obvious. Structural dissimilarity probably will provide a basis for new chemistry, new papers, and patents, all of which really are benefits. But unless the resulting drug has greater efficacy and/or less toxicity than known drugs for the same target, it will fall flat.

Another potential disadvantage of using an established target strategy is that the market for drugs acting by a given target can suddenly disappear when another, better target comes along, as it did for older cholesterol lowering drugs with the introduction of the statins beginning with Mevacor. After all, how can one look at the market conditions as assessed a few years ago, and then project that the market for a given disease is going to be the same for another 10 or 15 years down the road? Few drug markets today are what they were in the early 1990s and crystal balls are hard to come by in this business. You might end up with a new and improved Edsel, a significant improvement that nobody will buy.

Dangerous, rare toxicity, the bane of the pharmaceutical industry these days, might also be encountered. Recall that we’ve mentioned several times that a real clinical advantage is required for this strategy to work. Just as a clinical advantage can’t always be predicted in advance, neither can idiosyncratic toxicity. Should such toxicity be seen with a known drug
acting on a given target, it isn’t going to immediately be clear whether said toxicity was **compound toxicity** or **mechanistic toxicity**. The latter would doom all drugs acting on the target while the former might instead provide a better opportunity for competing drugs in the same class to grab market share.

An apparent example of compound-dependent toxicity can be seen in the metformin/phenformin story. Glucophage (metformin, Figure 5.3) is a biguanidine thought to act via multiple and not yet fully elucidated mechanisms to exert its blood glucose lowering effects. Another structurally similar drug, phenformin (Figure 5.3), is thought to utilize the same sort of mechanisms, but years of use by diabetic patients have shown an association with an unacceptably high-risk of life-threatening lactic acidosis, which fortunately metformin doesn’t share. So in this case, although a theoretical basis for mechanistic toxicity might be postulated (buildup of lactic acid due to hepatic gluconeogenesis inhibition, one of the effects of both drugs), it appears that the toxicity was largely compound dependent. Phenformin was withdrawn from the US market while metformin still does brisk sales. Two other diabetes drugs, the PPARγ agonists Rezulin (troglitazone), which was withdrawn from the market due to liver toxicity, and Avandia (rosiglitazone) still an important drug today, although one recently highlighting another toxicity concern, share a similar story with a different type of toxicity involved.

Because such situations can occur one might even wonder if fast follower or “me-too” approaches could count this type of possible scenario as a potential advantage. But that would be like setting up shop next door to your competitor in the fervent hope that lightning will strike and eliminate his business, only worse. It would be a ridiculous strategy.

To the risks that the market for a drug acting on an established target might disappear either because better targets have been found or unexpected mechanistic toxicity has been discovered, one can add the risk that in the future fewer drugs will compete in a given indication. This could be due to the increasingly restrictive formularies which are only looking to cover prescriptions for the cheapest drug in each class or it could be due to the
increasing pressures on the US government to regulate drug prices. If that happens, and, as we’ve seen in Chapter 2, there are those who believe it will, it would be very hard for pharma companies to justify high-initial pricing for a fourth-in-class or fifth-in-class drug regardless of its advantages. For these and other reasons many people believe that backing too many drugs based on established targets is just not a good idea and that the risk level of this approach isn’t so low after all.

All of these objections are built on future scenarios that may or may not come to pass. But there’s another difficulty with this approach that’s seen every day in the new drug discovery business. It’s just plain hard to find “IP space” for drugs acting on established targets. There are often already hundreds of existing composition-of-matter patents issued for compounds acting on such a target, each with a Markush structure implying more compounds than there are stars in the heavens. Gazing at page after page of descriptions of such structures brings the chemist closer to a true understanding of infinity, not to mention despair of ever finding something not covered. Add to that the fear that even if you do happen upon something not claimed in existing patents it doesn’t mean that work isn’t already underway somewhere to cover it in future ones. As we saw in Chapter 3, it takes 18 months from the date of filing for a patent application to be published. So whatever holes you find in their published ones may be in the process of being patched before you ever find them. Even leads which may have come out of screening and appeared to be proprietary may not be. All of this means that it might just be impossible to find a patentable compound of your own to develop when it comes to an established therapeutic target. And no patent generally means no project.

Studies of the industry consistently show that projects based on established targets are more likely to reach the clinical stage and are more likely to succeed in the clinic than projects based on non-established targets. In short, they’re seen as being winners. So keep in mind that although disadvantages may outnumber advantages as shown in Box 5.2, the single advantage of “higher than average probability of finding an efficacious agent” is often thought to outweigh all the negatives. As a result, a large percentage of new drug projects will fall into this category and most research chemists can expect to get plenty of experience working on them.

5.3 Established “Tough Targets”

A second type of established target project exists where a number of these disadvantages are obviated. Dr Hans Maag, Vice President of Chemistry at Roche Palo Alto, refers to these as “tough targets where we don’t know how to actually start chemistry, but where the biology is pretty clear cut.”

“There are all these biologics—antibodies and so on—out there which are commercially successful. So if you want to make a small molecule, basically the target is already validated. There’s no question. So it’s not a matter of validation of the target. You just want to find some small molecule that interferes with that target either allosterically or otherwise,” he explains.

An example of this kind of target is tumor necrosis factor (TNFα), an important cytokine shown to be a key player in many inflammatory disorders including rheumatoid arthritis, a painful, debilitating, and difficult-to-treat disease. The anti-TNFα antibodies Remicade (infliximab), Humira (adalimumab), and Enbrel (etanercept), a soluble form of the ligand
binding region of the TNFα receptor, have proven beyond any doubt the validity of TNFα as a target. All of these, as biologics, require injection and carry the risk of immunogenicity, not to mention high-price tags.

A small molecule interfering with the binding of TNFα to its receptor wouldn’t carry these liabilities and would likely become a billion dollar drug. Although other small molecules are known to interact somewhere along the TNFα pathway, like Thalomid (thalidomide) which somehow reduces TNFα levels, and thus have at least some sort of pathway validation, a molecule that could directly interfere with the binding of TNFα to its receptor, thereby mimicking the effects of Remicade and Humira, has proven hard to come by. Projects directed at this have encountered the usual difficulties of trying to mimic protein–protein interactions with small molecules, although recently a publication from Sunesis describes a small molecule inhibitor of this interaction which acts by essentially displacing one of the subunits of the normally trimeric ligand.

For any such project directed at a small molecule interacting with a target that’s already been validated in the marketplace by biologics, the drug’s advantages are obvious. IP space is wide open (since probably no one else has such an agent either, although they may be looking for one), and the market for the resulting drug is almost guaranteed. The only stumbling block, which must not be underestimated, is finding such an agent—it’s just not clear that it can be done. But the saving grace is that the failure to identify a good lead would at least happen early on in the project, once again as in a synthetic scheme where the trickiest step is one of the first ones and not the last step where you’ve invested much more time, energy, and hope. An oft-repeated aphorism in drug discovery is that if you’re going to fail, fail early.

We’ve now looked at two “flavors” of projects based upon established targets. But what about the ones nearer and dearer to a chemist’s heart and at the vanguard of biotech and pharmaceutical research, new targets?

5.4 Novel Targets

5.4.1 Identifying New Targets

Although the human genome contains slightly more than 20,000 genes which, due to post-translational modifications, alternative splicing, and protein complex association, might make for in excess of one million different proteins (not to mention other, non-protein targets) in the human “targetome,” only a very small percentage of these are expected to provide good therapeutic targets. As we’ve seen in Chapter 2, genomics has till date proven to be far more capable of elucidating rare inherited disorders stemming from mutations in a single protein than in identifying new targets for common diseases, and the expected flood of new targets hasn’t materialized. These days new TID has largely shifted to the science of proteomics.

The differential expression of a given protein in diseased versus normal cells or tissues is often the starting point for investigating whether or not it constitutes a good target for a drug. Potential targets can stem from other sources too, such as progress in understanding biochemical pathways and their rate-limiting steps, or new targets identified after phenotype screening, described later, but differential protein expression is probably the most common source of proposed new targets. Examples would include the overexpression of a given
protein in a number of different tumor cell lines or its correlation with more aggressive or more highly metastatic tumors. Probably thousands of such proteins have been examined in industry and academia already, since such data raise the possibility that they might be crucial to tumorigenesis, tumor growth, or tumor spreading.

In such cases, a number of caveats need to be kept in mind. First of all, expression data obtained using microarrays reflect only mRNA levels and don’t necessarily correlate with protein concentrations which in turn don’t necessarily correlate with protein activity levels. Concentrations of two proteins with the same relative mRNA levels might differ by more than 20-fold, showing that “transcript levels provide little predictive value with respect to the extent of protein expression.” And not all proteins are produced in active form. Some require further processing or association with other proteins or are modulated by, say, endogenous inhibitors whose levels might vary. So in interpreting such data, the scientist should keep in mind that “there’s many a slip ‘twixt the cup and the lip.”

Second, the dynamic nature of protein expression means that it’s always cell- and time-dependent. Expression levels in early stage tumor cells will probably vary considerably from those in later stage tumors, and both may have little in common with protein expression levels in, say, endothelial cells that line the tumor vasculature. A tumor cell from an aerobic environment (near blood vessels) will have a different proteomic profile from the one buried deep inside a tumor in an anaerobic environment. These kinds of issues can complicate the interpretation of protein expression in any type of tissue.

But most importantly of all, correlation can’t establish causation. “Omics technologies are giving us correlative data while what we really need is causative data,” says Dr David Szymkowski, Director of Biotherapeutics at Xencor of Monrovia, California. He points out that the reasoning behind the assumption that a protein overexpressed in a given disease state is a valid therapeutic target is analogous to the following bit of (il)logic:

- Firemen are found at burning houses
- Firemen are not found at normal houses
- Therefore, firemen cause house fire
- Therefore, eliminate firemen to prevent fires

Overexpression of protein in a disease state might actually be a defensive response and not something one wants to eliminate. A second possibility is especially worth considering:

- More weddings occur in June than in any other month
- More suicides occur in June than in any other month
- Therefore, weddings cause suicides
- Therefore, eliminate weddings to prevent suicides
That is, the increase or decrease of protein expression in a given disease state can be entirely coincidental, the proteins in question being more like innocent bystanders than causative agents. The vast majority of the proteins differentially expressed in a given disease will fall into these two categories.\textsuperscript{18}

Clearly, launching a new project based on such a target, however exciting, with the subsequent commitment of money and manpower, can only be justified by an equal commitment to find firmer evidence as soon as possible, and to drop the project quickly if no such evidence is forthcoming. Everything, then, depends on TV.

5.4.2 Target Validation

5.4.2.1 Levels of Validation

The term “target validation” can mean different things to different people. To a chemist proposing a new project it might refer to affirmative evidence from a mouse knockout or even just an RNA interference (RNAi) experiment in cells. To a clinician it probably means no less than proven statistically significant efficacy by a drug in large scale, controlled human trials. In the strictest sense, the clinician is right, but in the real world it’s obviously impossible to put agents acting on every new proposed targets into the clinic to find out whether they’re valid or not. Since, however, the need for new drugs acting on new targets is very real, a series of milestones which demonstrate a target’s importance in cells and animals need to be met between the time that a target is identified and the time a drug acting on it can enter the clinic. The closer a compound gets to the clinic and the farther along a project advances, the higher the bar becomes. As mentioned earlier, this means that TV doesn’t consist of a single experiment but instead is an ongoing process. It’s best thought of as “target confidence building.”\textsuperscript{19} The closer a project gets to the clinic the more expensive it becomes and the more confidence that indeed the target is a valid one is required to satisfy scientists, managers, and investors.

A number of authorities have commented on the need for a staged TV and suggested some of the milestones that might be used to define each stage.\textsuperscript{20} Table 5.1 shows one possible system that might be used, and is worth a good look.

“Level None” is included as a reminder that initially identified targets should in no way be assumed to be valid ones. Using our cancer analogy, Level I validation might consist of

| Level | Milestone |
|-------|-----------|
| None  | Correlative evidence only |
| I     | In vitro POC (e.g. works in a cellular disease model) |
| II    | First in vivo POC (works in animal disease models) |
| III   | Efficacy in one or more Phase II clinical studies |
| IV    | Efficacy in Phase III trials sufficient for NDA approval |
evidence showing that blocking the function of Protein X induces apoptosis in various cancer cell lines but has less of an effect on normal cells. Tools for thus perturbing protein function (RNAi, knockout, antibodies, etc.) can vary, as discussed in the next section, and different results can sometimes be obtained with different tools. Early stage projects usually lack small molecules capable of selectively interfering with a single desired target, so at that point TV experiments often use other methods, but ultimately small molecules will provide the best validation when a small molecule drug is the goal.

Many proposed targets will drop out even at this lowest validation level. For those that don’t, the push will be onto develop a tool (small molecule, antibody, etc.) capable of selectively perturbing the target in an animal model of disease. Achieving this first in vivo proof-of-concept (POC), thus reaching Level II validation in our scheme, is a major milestone and cause for celebration in any project, but it still doesn’t prove that an efficacious drug is just around the corner. Level II validation in cancer might mean tumor growth inhibition, or, better still, regression in a xenograft (using human cancer cells in an immune compromised animal) or syngeneic (using a cancer native to the species) mouse model. For cancer especially, efficacy in a single model doesn’t really predict human clinical success. But it’s generally true that a compound that works in many different models has a better shot at clinical efficacy than the one that’s only effective in a few. The upshot is that for many indications Level II validation needs to be established in multiple models for it to mean much.

Once a program has produced a clinical candidate for a new target, the true test of its validation approaches. Failure to achieve a viable safety profile, pharmacokinetic properties, or withdrawal due to “business reasons” sometimes means that the new target ends up being neither validated nor invalidated. With proof of efficacy in a Phase II study, though, the first studies really designed to look for it, the first level of validation in man (Level III) has been achieved. Even here validation is not unequivocal. In some disappointing cases clinical efficacy can’t be reproduced in a second Phase II study or a later, Phase III trial. Hence the highest level of target confidence, Level IV validation, is reserved for drugs and targets that have run the full gamut.

Successfully reaching each of these milestones to validate a new target is probably the biggest challenge to be met in all of drug discovery. Its difficulty is a major factor in the relative dearth of drugs aimed at new targets that get approved each year. So drug discovery projects directed at novel targets need all the “target confidence building” they can get along the way. Fortunately, a wide variety of validation tools are nowadays available, as discussed in the next section.

5.4.2.2 Target Validation Tools

5.4.2.2.1 Knockouts and Knock-Ins

The ability to produce an animal lacking a particular gene product of interest allows researchers to look for disease-relevant phenotypes and has become a powerful tool for validating targets (Figure 5.4). Because of their genetic similarities to humans, mice are generally used, although much information has also been gained from knockout experiments in lower organisms like Caenorhabditis elegans and drosophila. Three types of animals can be compared: wild type (+/+), heterozygotes (+/−), and homozygotes (−/−) so that a crude sort of dose-dependence can sometimes be noted in phenotype.
Knockout mice have proven to be useful in the validations of many drug targets, one good example of which is the cysteine protease Cathepsin K. This enzyme is mostly found in osteoclasts, a specialized type of cell involved in bone remodeling. Cathepsin K is a proteolytic enzyme whose purpose was thought to be the degradation of the collagen matrix that’s left over after decalcification during the process of bone resorption. Osteoporosis, a disease characterized by low bone density and one that’s particularly prevalent and problematic in aging populations, is believed to represent an imbalance between bone buildup and bone breakdown, so if Cathepsin K plays an important role in the latter, inhibiting its function might provide an effective way to treat the disease. When Cathepsin K \(-/\) mice were generated, they were found to have abnormally dense bone, with severely impaired bone resorption.\(^{23}\) This provided a measure of validation for this target in osteoporosis,\(^ {24}\) and further validations along with several clinical candidates aimed at it have since appeared.

A study by researchers at Lexicon Genomics (now Lexicon Pharmaceuticals) looked at the targets for the top 100 prescription drugs and found that for the targets where knockout models could be generated (most of them) in 85% of the cases the knockouts produced phenotypes that were “informative” for the therapeutic indication involved.\(^ {25}\) Not only Lexicon but also National Institutes of Health (NIH) has now embarked on projects aimed at knocking out most if not all of the genes found in mouse.\(^ {26}\)

Like any tool, though, knockouts aren’t universally applicable and free of all problems. Although many knockout mice are now commercially available (but expensive), most require quite a bit of time and money to generate. In many cases (some say 20% of the time, others say more frequently) knockouts result in embryonic or neonatal lethality, the absent protein apparently being required for proper development whether or not they’re needed later. When this happens, of course, information about the target protein’s role in mature animals can’t be obtained, although sometimes the lethality itself means something: finding embryonic lethality due to defective blood vessel formation in heterozygous vascular endothelial growth factor (VEGF) knockouts implied that nature didn’t have any good ways of compensating for its loss during angiogenesis. This provided a key piece of information in developing anti-angiogenic anticancer compounds like Avastin, a VEGF antibody.\(^ {27}\)

In general, nature is adept at compensating for the loss of an incredible variety of proteins, and this presents a problem when phenotypes one expects to observe just aren’t there because compensation by redundant mechanisms happened during the development. And, surprisingly, different phenotypes are sometimes observed for the same knockout when different strains of mice are used.\(^ {28}\) Furthermore, the phenotypes that are observed in knockout mice

**Figure 5.4** Points of intervention utilized by different TV tools.
aren’t always straightforward in predicting what will happen with small molecule inhibitors. For the important PPARγ receptor, heterozygous knockout mice demonstrated improved sensitivity to insulin when fed with a high-fat diet, which suggests the strategy of using PPARγ antagonists to treat diabetes. However antidiabetic thiazolidinediones like rosiglitazone, which improve insulin sensitivity are PPARγ agonists. Needless to say, this took a bit of explaining.

One big reason that knockout results may not correspond with results obtained using small molecules is that prospective target proteins may well be involved in more than the one particular activity that a small molecule might target. A small molecule might just interfere with its binding to a ligand at one site, but proteins are often multidomain and can form necessary complexes with other proteins via interactions not affected by the small molecule. This effect, illustrated graphically in Figure 5.5, means that small molecules are capable of tweaking proteins while knockouts, as the name implies, aren’t exactly a light tap on the cheek.

Figure 5.5 A fundamental difference between pretranslational and posttranslational tools for TV, as illustrated for a receptor tyrosine kinase. (a) Posttranslational methods like inhibition by small molecules can affect target protein function without affecting its association with other proteins. (b) Pretranslational methods like knockout or RNAi knockdown result in the loss of target protein and all of its interactions with other proteins. (Reproduced with permission from Fitzgerald, K. RNAi versus small molecules: Different mechanisms and specificities can lead to different outcomes. Curr. Opin. Drug Discov. Dev. 2005, 8, 557–566, copyright 2005, The Thomson Corporation.)
“Knockout of a gene results in complete loss of the target protein, which can disrupt protein complexes or impair functional domains that would be unaffected by a drug. These secondary effects can dominate the phenotype of a knockout, such that a knockout and a small molecule inhibitor targeting the same protein produce very different outcomes.” The authors of this quote recommend a less blunt-force transgenic method, replacing target proteins with mutants lacking catalytic activity (knock-ins) rather than deleting them altogether, as a more predictive alternative to knockouts. They provide an example (that of the kinase p110γ) where the two techniques produced opposite phenotypes. Along with other techniques like conditional knockouts, where the absence of a protein can be induced specifically in a desired tissue, such knock-ins provide additional useful tools for TV.

### 5.4.2.2.2 Antisense Oligonucleotides

In 1978, researchers at Harvard reported that a 13-base deoxynucleotide segment which was complementary to a section of Rous sarcoma virus mRNA could block virus production in cell culture. Later work demonstrated that such antisense oligonucleotides (AOs), varying in length from 7-mers to about 30-mers, bind to their target mRNA sequences through Watson–Crick base pairing and either lead to cleavage of the mRNA by the enzyme RNase H or prevent translation by other mechanisms. The end result is the same, however: production of the target protein is shut down (see Figure 5.6).

![Figure 5.6](image)

**Figure 5.6** How AOs block protein expression. An AO binds to the target sequence in the sense strand of the target mRNA. In some cases, RNase H then cleaves the mRNA part of the resulting duplex, freeing up the AO for further cycles of binding and cleavage, while in other cases different mechanisms resulting in a blockade of the ribosome are induced. (Reprinted with permission from Taylor, M.F et al. Antisense oligonucleotides: A systematic high-throughput approach to target validation and gene function determination. *Drug Discov. Today* 1999, 4, 562–567, copyright 1999, Elsevier.)
“Naked” oligos like the 13-mer used for Rous sarcoma virus are rarely used nowadays. They normally aren’t very stable because of the presence of nucleases, endogenous enzymes which can cleave them in minutes. This means that chemical modifications are normally required for them to stick around long enough to have a good effect. One of the earliest and the most consistently used modifications involves replacing a non-bridging oxygen atom in the phosphate linkage with a sulfur atom, resulting in phosphorothioate deoxyoligonucleotides. Others involve alkylation of the 2’ hydroxy groups of ribonucleotide oligomers with methyl or methoxyethyl groups (which also seems to increase their affinity for target mRNA), generation of methylphosphonates in place of phosphate linkages, and various modifications in the bases or in the ribose moiety, some even going so far as to replace the ribose rings with morpholines or polyamides. Sometimes chimeric oligos incorporating several of these modifications at different positions are used.34

Not every stretch of a given protein’s mRNA sequence is a good site for AO binding, so some experimentation is required, but since only gene sequence information, not secondary or tertiary protein structural information is required to design such tool compounds, they can be generated much more quickly than small molecule inhibitors. Getting them into cells, however, isn’t always as straightforward, and works better in some types of cells than in others. Methods for transfection include the use of polycationic reagents, liposomes, and special techniques like electroporation and microinjection. Antisense oligonucleotides have proven to be useful for TV both in cells and in animals, where the embryonic lethality often seen in knockouts is not an issue.35

Limitations, though, include possible non-specific effects due to interactions with various proteins and toxicity due to the transfection reagents that need to be used. As with the use of knockouts for TV, the absence (not just presence in inactive form) of target protein might invoke phenotypes that differ from those seen with small molecule inhibitors. Unlike knockout, however, the target protein will probably not be 100% absent when using antisense techniques, although >70% knockdown is often achieved. Importantly, since antisense reagents interact with cells or animals that already have the target proteins present and merely prevent future production of them, a lot depends on the half-life of the existing protein. If it’s on the order of days (which can happen) it may take at least that long for antisense knockdown to have any effect on phenotype, yet another reason why AOs need to be very stable.

As is the case for every other TV tool mentioned in this section except for knockouts, AOs have human therapeutic potential as well. Vitravene (fomiviren) is a phosphorothioate 21-mer from the pioneering antisense company Isis Pharmaceuticals that’s been approved for use in CMV retinitis in patients with HIV. The drug is injected intravitreally (into the eye). Another Isis antisense drug, ISIS 30102, which targets apoB-100, has shown promise in Phase II clinical trials with a cholesterol lowering effect comparable to that obtained with atorvastatin.36 This drug was given by subcutaneous (SC) injection. Other antisense drugs are in clinical trials as well.37 Antisense oligonucleotides remain important for both validation and therapeutics, but in many ways they’ve recently been eclipsed in popularity by another technique directed at mRNA, called RNAi.
5.4.2.2.3 RNAi

Twenty years after the discovery that AOs could block the translation of mRNA into target proteins, another method of posttranscriptional target knockdown was found. In 1998, Dr Andrew Fire, then at the Carnegie Institute of Washington, Dr Craig Mello of the University of Massachusetts, and their co-workers published a paper demonstrating that injecting low levels of *double stranded RNA (dsRNA)* into *C. elegans* could inhibit the expression of the target protein in a sequence specific manner by a mechanism different from that observed for AOs. This and subsequent work which uncovered the details of this new type of *RNA interference (RNAi)* eventually led to them being awarded the Nobel Prize for Medicine in 2006. Figure 5.7 outlines our current understanding of the processes involved.

Long dsRNAs like the one used in the above report are cleaved by the nuclease Dicer to produce double stranded 21–25mers with several unpaired nucleotides (overhangs) at each 3'-end. These also have 5'-phosphate groups and are called *small interfering RNAs (siRNAs)*. Unwinding of the strands and association with the *RNA-induced silencing complex (RISC)* allows the antisense strand to guide to the complex to the targeted mRNA by hybridization with the right mRNA sense sequence. Once bound, the mRNA is cleaved in the center of the cognate sequence by the

![Figure 5.7](image)

**Figure 5.7** The mechanism of RNAi. (a) Long dsRNA is cleaved by the ribonuclease Dicer to form double stranded 21–25 nucleotide fragments called siRNAs. (b) siRNAs can also be introduced directly into the cells. In either case, association of the antisense strand with the RISC guides the resulting complex to its cognate mRNA site, where the mRNA is cleaved, selectively blocking protein translation. (Reprinted with permission from Manoharan, M. RNA interference and chemically modified small interfering RNAs. *Curr. Opin. Chem. Biol.* 2004, 8, 570–579, copyright 2004, Elsevier.)
Argonaute component of RISC, thus selectively preventing protein translation, and the antisense strand of the siRNA is free to do it all again.

This catalytic mechanism is believed to have evolved as a defense against viruses in simple organisms lacking more complex immune systems, but it isn’t used that way in mammals. Putting ds RNA segments longer than 30 nucleotides into mammalian cells is problematic: an interferon response is induced and the cells tend to die. Because of this RNAi is normally achieved using presynthesized 21–23 nucleotide siRNA reagents that are transfected into cells using methods (carriers, electroporation) similar to those used for AOs (Figure 5.7(b)), or by introducing a DNA vector for short hairpin RNA (shRNA), which is converted to the relevant siRNA reagent within the cells.39

“Naked” siRNAs, being double stranded, are more stable than “naked” AOs, and are, unlike the latter, normally fine for cellular experiments. But for in vivo use where they’re added exogenously, their stabilities,40 PK parameters, and perhaps affinities might need to be improved by the same sorts of chemical modification (phosphorothioates, etc.) used for antisense reagents.41 Reagents for siRNA experiments in cells are probably the easiest of all TV tools to come by. They can be designed solely on the basis of gene sequence and ordered over the Internet. Vast siRNA libraries can now be purchased as well. As with AOs, getting the compound into the cells can be the trickiest part of the whole process, multiple controls containing some mismatches need to be used, and actual target protein levels need to be determined before any conclusions are arrived at. But compared to the efforts involved in finding a potent, selective small molecule inhibitor of the target, RNAi can constitute a much easier route to early TV, particularly for classes of targets where good small molecule tools are hard to come by. And most, but not all, studies comparing the use of siRNA with that of AOs have found the former to be effective at lower concentrations.42

These advantages go a long way toward explaining the increasing popularity of RNAi in TV. In a typical recent example, a commercially available siRNA targeting matrix metalloprotease-11 (MMP-11) impeded cell proliferation, resulted in apoptosis in the mouse hepatocarcinoma cell line Hca-F, and attenuated the metastatic potential of these cells in a mouse model.43 Researchers at Genentech took advantage of conditional RNAi, a newer technique44 to show that knocking down BRAF, a serine–threonine kinase, slowed tumor growth in one xenograft mouse model of metastatic melanoma and caused regression in another, thus providing in vivo validation for BRAF as a cancer target.45 Small molecule inhibitors of this novel target exist but are not selective.

Originally RNAi was thought to be exquisitely specific for the desired target, but recently this has been called into question, and off-target effects have been reported.46,47 Some mismatches between the reagents and the target mRNA may thus be tolerated, and one possibility is that siRNAs might in some cases be acting like their cousins, microRNAs, important, newly discovered regulators of mammalian protein translation which seem to act non-specifically by recognizing as few as seven nucleotides.48 Another possibility is that the sense strand of the siRNA duplex, normally assumed to be degraded, can still play a role. The sometimes non-specific effects of siRNA are an area of ongoing research, but things need to be kept in perspective: “While off-target activity of siRNA has received a large press coverage lately, the difficulty in obtaining small-molecule specificity has been an ongoing struggle within the pharmaceutical industry for many years.”49 This would make arguments against the use of siRNA reagents by chemists because of an alleged lack of specificity a little like the
pot calling the kettle black. Specificity is an issue for any TV tool, or for that matter any therapeutic modality. Methods for demonstrating specificity, especially the use of rescue experiments (adding back the gene product via a transcript sequence that the siRNA doesn’t target) and using multiple siRNA reagents targeting different parts of the transcript, have been suggested.50

Another assumption that dsRNA with less than 30 nucleotides (read, siRNAs) will never cause an interferon response has also been called into question.51 These kinds of problems are to be expected for such a young technology, but in many ways the most remarkable thing is how far it’s come in such a short time. RNAi therapeutics are already undergoing clinical trials and one of the companies most involved in the technology, Sirna Therapeutics, was recently acquired by Merck for more than $1 billion, which indicates just how promising the industry feels this technology ultimately is.

RNAi certainly deserves a seat at the TV table, its main advantage over small molecules being the ease of reagent design and construction.52 No method is perfect, of course, and it’s been suggested that much of the utility of RNAi or any other TV tool lies in combination screening with other tools. An example of this is the recent use of a genome-wide siRNA screen to identify genes that sensitize non-small cell lung cancer (NSCLC) cells to the effects of paclitaxel, an experiment that could help point the way to new therapeutic targets as well as more rational combination therapies for cancer.53

5.4.2.2.4 Antibodies

Antibodies represent an endogenous method for targeting extracellular proteins. Dozens of monoclonal antibodies (mABs), those derived from one clone and recognizing one antigen, have been approved for therapeutic use, mostly in cancer and inflammatory diseases, and literally hundreds more are in the clinic, which accounts for the massive amounts of research going on in this area. But they can also be used as exquisitely selective TV tools, provided that the target is an extracellular one. Figure 5.8(a) shows the structure of a full length antibody.

Their high molecular weights (~150 kDa) partly explain why they normally can’t be used for intracellular targets: they’re highly cell impermeable. And even introducing a full length

![Figure 5.8](image_url)
mAB into cells through microinjection or having them expressed by the cells themselves is problematic. The reducing environment inside the cell doesn’t allow for the disulfide bonds necessary for antibodies to maintain the right conformation and antibodies produced intra-cellularly don’t fold properly.

For extracellular targets, although they aren’t quite as simple to make as siRNAs, antibodies have proven to be useful in the early stage drug discovery via libraries generated through phage display or as tools to study the physiological role of extracellular proteins like selectins. One of the nice things about using them for TV is that if the target proves valid and useful, mABs used to establish this might themselves be developed into therapeutic agents. But generating mABs that bind selectively and have the desired effect on function (not a given) can be challenging.

Recent years have seen a growing interest in intrabodies (intracellular antibodies). As shown in Figure 5.8, these usually consist of the light chain (VL) and the heavy chain (VH) antibody domains containing binding recognition elements, held together by a peptide linker. These are referred to as single chain Fv fragments (scFvs). Genetic engineering is normally used to express them in cells, and their effectiveness sometimes depends on their ability to redirect target proteins to different locations within the cell via peptide sequences attached to them that function as “cellular Zip codes.” It’s been observed that the biggest barriers to the effectiveness of scFvs are often stability and solubility, not binding affinity per se. Despite these limitations, intrabodies have proven to be useful in validation experiments for the cysteine protease Cathepsin L in cancer, and the hepatitis C serine protease NS3 in liver cancer, among others.

Because of the size of these molecules, both intrabodies and antibodies have the exciting ability to affect protein–protein interactions, which small molecules are usually not good at, and because of their mode of action they’re not limited by the time lag involved in slow protein turnover, but can take effect as soon as the reagent is there.

5.4.2.5 Aptamers

Besides their use as AOs and siRNA reagents, oligonucleotides find yet another use as perturbogens for TV studies as aptamers. The term refers to short (<100 nucleotide) stretches of DNA or RNA that fold in a way conferring shape complementarity with a target protein. Peptide aptamers exist as well, these having more in common with antibody fragments like those discussed above. Oligonucleotide aptamers are selected from pools of up to 10^15 candidates by an automated in vitro process called systematic evolution of ligands by exponential enrichment (SELEX) that can now be completed in just a few days. Not only proteins but also small molecules and metal ions can be targeted: an RNA aptamer was found to bind to theophylline with 10,000-fold greater affinity than for its methyl analog, caffeine. Aptamers don’t seem to induce immunogenicity when given in vivo, which is often a major limitation for antibodies.

One recent example of the use of aptamers to study potential targets makes use of an RNA Spiegelmer (“spiegel” being German for mirror), a mirror image oligonucleotide made with L-ribose sugars in place of the normal D-ribose rings so that the resulting compound can resist degradation by nucleases for days. This aptamer, called L-NOX-B11, binds to octanoylghrelin with nanomolar potency. Ghrelin is a 28-amino acid peptide that’s an endogenous ligand for the growth hormone secretagogue receptor 1a (GHS-R1a) and a potential obesity
target that gets a lot of press coverage. Pegylated ϵ-NOX-B11 was found to inhibit ghrelin’s effects in central nervous system (CNS) in an in vivo experiment, and a successor spiegelmer produced weight loss in a diet-induced mouse obesity model.

Most aptamers find use against extracellular targets, but, as with antibodies, an intracellular version of this technology also exists, in this case called intramers. A good review of the use of aptamers for TV was recently published, emphasizing the utility of aptamers in HTS assays as well. An aptamer drug, the VEGF antagonist Macugen was recently approved by the FDA for the treatment of age-related macular degeneration (AMD). Like the antiviral AO drug Vitravene, Macugen is injected into the eye. Several other aptamers are in the clinic as well.

### 5.4.2.2.6 Small Molecules

After reading the preceding sections, most readers probably won’t need too much convincing that small molecules usually make the best tools to use in validating a target. Perhaps the best way to consider the advantages that small molecules bring is to look at the disadvantages they lack: embryonic lethality, compensation during development, structural effects due to the absence of normal protein–protein interactions, long delays between administration and effect, difficulties in determining dose–response, immunogenicity, interferon response, extreme challenges with intracellular delivery, and general unsuitability for oral administration.

They’re not without problems of their own. Finding a selective small molecule inhibitor from scratch will probably be more difficult than finding a functional siRNA tool. A chemist can’t look at the sequence of a gene, design a small molecule inhibitor, and order it online. Problems with selectivity, cell permeability, and oral bioavailability are no strangers to the medicinal chemist. But keep in mind that most drugs on the market are and will continue to be small molecules. Ultimately any project directed at such a compound will need to use them in validation experiments up to and including clinical studies. Although other validation tools can be helpful along the way, particularly when used in combination, for such projects only small molecules are “the real thing.”

### 5.4.3 Working on Novel Target-Directed Projects

For purposes of this book, the line between an “established” target and a “novel” one has been drawn at clinical efficacy in Phase II, which represents Level III TV according to the scheme described earlier. This definition of target novelty therefore encompasses everything from targets completely lacking any validation through targets which have been proven to work in multiple animal models of a given disease. You can see that these represent widely varying degrees of “novelty,” including targets only just now existing in the mind of a single scientist to targets for which positive animal model data have already been published in *Science* or *Nature*. How will such a project differ from those based on established targets?

Box 5.3 summarizes the advantages and disadvantages of projects directed at novel targets. First of all, when working on a project directed at a novel target, research scientists are generally going to be more excited and therefore more motivated, a factor that should not be underestimated. The prospect of designing and making with one’s own hands a first-in-class
drug, especially a structurally novel one that fills an unmet need, will probably translate into
greater efforts, more creativity, and more focus for even the most jaded veterans. Few are the
medicinal chemists who haven’t dreamt about coming up with the next Gleevec (imatinib), the
breakthrough Bcr-Abl kinase inhibitor which has dramatically improved and extended
the lives of patients with chronic myeloid leukemia (CML). Contributing to a first-in-class
Hsp90 inhibitor for cancer or Cathepsin K inhibitor for osteoporosis is just more appealing to
a scientist than coming up with a fourth-in-class drug. And as everyone knows, the difference
between a motivated employee and a lackadaisical one is night-and-day.

One might think at first that new targets would bring with them wide open IP space. But
novel targets do not, per se, imply novel compounds. A lot depends on just how different that
target is from known ones. As we’ll see later, leads for a new target that’s a member of a
known family, such as kinases or aspartyl proteases, can often be found in libraries of
compounds active against other protein family members, as in the situation where screening
an aminoquinazoline library turns up a hit for new kinase target. This class of compounds is
extensively claimed as kinase inhibitors in the patent literature, so although the target may be
new the inhibitor probably is old and there may not be much IP room. But if the novel target
itself is significantly different from popular, well-explored targets, there’s a better chance that
a structurally novel lead that binds to it will be identified, if one can be found at all. So once
again, although the target doesn’t dictate the compound, it has a certain amount of
“influence” over what it will be.

For novel targets, although patent lebensraum doesn’t necessarily come with the territory,
market space does. Even in a crowded therapeutic area where many effective drugs exist, a
drug intervening at a different point is always welcome. The first-in-class renin inhibitor
aliskiren (Figure 2.8), recently approved for the treatment of hypertension, is a good example
of this. Renin, an aspartyl protease, is involved early on in the angiotensin pathway, which
the well known angiotensin converting enzyme (ACE) is involved in later, and which largely
regulates blood pressure. Renin inhibitors have been around for decades. But these inhibitors
tended to be rather large, peptide-like compounds and problems with oral bioavailability

---

**Box 5.3 Advantages and Disadvantages of Working on Novel Targets**

**Advantages**
- More motivating to project scientists
- *Market space and potential market exclusivity*
- Potential for advantage over or synergy with other drugs

**Disadvantages**
- Requires validation
- May not have benchmark compounds or relevant assays
- *High attrition rate*
kept them from becoming drugs. Before the appearance of aliskiren, renin inhibitor projects were most notable for having eventually provided leads for another aspartyl protease that came along much later, HIV protease. Not surprisingly, oral bioavailability always seemed to be the major obstacle to developing this class of drugs, but with great effort and after decades of work that obstacle has been overcome.

Three points are apparent from this story. First, targets in a given family do tend to prefer structurally related compounds. Second, even the most difficult problems can eventually be overcome given sufficient incentive. And third, our original point that market space does exist for drugs with novel mechanisms of action (MOAs). Now that aliskiren has made it to market, both single agent and combination use with other antihypertensives can be expected to provide a win–win situation for its maker, Novartis.

Why should there be a need for such a new drug, albeit one with a different target, when so many are available already? For the same two reasons we’ve encountered before: better efficacy and/or better safety, either with the new agent alone or in combination therapy with existing drugs. One could just as easily ask, “Why should an organic chemist need more than one way of putting on a silyl protecting group?” Because sometimes the “favorite” way either doesn’t work or works too well and silylates other things. The answer breaks down to reactivity and selectivity, which are curiously similar to a drug’s crucial attributes, efficacy and safety. As chemists we intuitively realize that transformations for which dozens of reagents and conditions exist are the most popular and/or the most difficult ones. An analogous situation applies to drugs.

The use of a drug acting against a new target in combination with another existing drug is a particularly attractive proposition, especially when it comes to difficult, life-threatening diseases, where therapeutics need all the help they can get, like hepatitis C virus (HCV), HIV, and cancer. While the hope remains that someday a single agent efficacious enough to at least keep such a disease in check or even cure it may be available, the reality is that no such drug exists today. Existing therapeutics can work but have limitations. So when different classes of them are available they’re frequently used together in the kind of combination therapy as discussed in Chapter 1. Here, the beneficial effects can be additive or, better yet, synergistic, although studies will need to be carried out to make sure that their toxicities aren’t synergistic too.

This approach is best exemplified by HIV drugs. The first target to be broached in the war on this fearsome virus was an enzyme called reverse transcriptase (RT), which the virus requires to convert its genomic RNA into DNA, and the first approved drug to target it was the nucleoside analog Retrovir (zidovudine, AZT, Figure 5.9), which had originally been looked at for cancer. Although this compound alone proved a lifesaver for many AIDS patients, efficacy was not ideal and toxicity in the form of neutropenia, severe anemia, and liver disorders was encountered in some patients. Furthermore, the fast and sloppy replication process the virus uses soon gave rise to mutations, some of which were resistant to the effects of the drug.

Other RT inhibitors soon reached the market including other nucleosides and later some non-nucleosides (NNRTIs), but all acting on the same target—incremental improvements but still far from ideal. A major advance in AIDS therapy came with the identification of another target, HIV protease where, as mentioned above, samples of renin inhibitors in the compound collections of pharma companies provided leads. Even before the first-in-class
compound, Roche’s Invirase (saquinavir), was approved for sale, it was general knowledge that if it worked, such a drug’s real future would lie in combination with RT inhibitors, and so it has proven to be. Of course, additional targets for HIV (fusion inhibitors, integrase inhibitors, etc.) continue to be pursued, with each new first-in-class drug expected to find a market for combination therapy of AIDS and each improving the lives of many patients.

If a drug acting on a new target can still find market space in a therapeutic area where other drugs already exist, imagine what it can do where there is no effective current therapy! This might include new diseases like severe acute respiratory syndrome (SARS) (assuming a company wanted to gamble on that eventually becoming a lucrative market), or old ones for which existing drugs only treat the symptoms. Finding something there would be a major accomplishment and would be valued as such.

Counterbalancing all of these advantages, of course, are the several major disadvantages. We’ve already seen that despite all the new, improved technologies that can be used, the process of preclinical TV is often slow, resource intensive, and, worse still, sometimes not predictive of clinical results. This last factor is the most discouraging of all, as late stage efficacy failure is the most expensive as well as discouraging sort of project failure.

A less severe, but still real limitation involved in pursuing novel targets is the fact that assays may need to be developed and reagents (things like the protein of interest or target substrates for proteolytic enzymes) may need to be made before validation can even begin. It might not be simply a matter of browsing through catalogs for reagents or kits or tracking down a procedure in Methods in Enzymology. Real work might need to go into it, which means FTEs will need to be approved by the management for a speculative project that hasn’t even begun yet. Most likely, small molecules acting on the prospective target won’t be known either, and as we’ve seen, small molecule inhibitors or agonists or antagonists are very fine tools for biologists to have at their disposal when doing TV. It could be that such a tool molecule wouldn’t be available until the project is in full swing lead optimization (LO) mode, which a company is unlikely to enter until a target has been better validated, a sort of a Catch-22. Normally, this would mean that validation would need to progress without a small molecule tool. But there is one other way that’s based on the preference of target families to bind to structurally similar compounds, as noted earlier. This is called chemogenomics.

---

**Figure 5.9** The HIV RT inhibitor Retrovir (zidovudine, AZT) and the HIV protease inhibitor Invirase (saquinavir).
Dr Mark Murcko of Vertex Pharmaceuticals provides an example of how this might work for a real-world project: “In the area of kinases or ion channels or GPCRs we’ve already worked on other members of the family,” he notes. “We already have some know-how about how to set up an assay in the field. We may already have useful compounds—not drug candidates, perhaps—but pretty decent starting points. Let’s say we read a paper in *Nature* that suggests that kinase X is a great target for cancer. But it’s speculative.

“What are our choices? Our choices are, a) do nothing and wait a few years till there are more publications so there’s more validation, b) put a lot of biologists on the project to try to make sense out of it, c) do a collaboration with somebody in academia who’s an expert in that target and use them as our surrogate biologist—which is a sort of variation of b—or, d) we could say, ‘Well, gee, in three months we can get an assay up and running, screen our kinase deck, see if we have a 100 nM inhibitor of kinase X, and then take that into primary cells for that cancer type and see whether there’s anything interesting. If not, then we’re done.’ It’s actually less work to do it that way than to do all the laborious biology up front.

“And what if you’re lucky and you find a molecule that’s pretty good? Then you’ve given a tool to the biologists so that when they go do their basic biology studies of that target they now have a tool they wouldn’t have had otherwise. If we pursue the target with no tool compounds, they can use RNAi or other biological tools, but it would be easier and faster for the biologists to do what they need to do if they had a tool compound. My view is that small molecules, if they’re selective and cell permeable, are always better tools.”

Add to this the fact that if successful your 100 nM tool compound might also be a 100 nM lead compound for which the chemistry is worked out, IP issues have been looked at, and some structure–activity relationship (SAR) may already exist if other analogs were also in the screen, and you’ve got quite a good strategy. But keep in mind that this relies on you having screening compounds directed at that family of targets already available. If you only have their second cousins (something that looks a little like something that might possibly be active) or would need to start making compounds from scratch it’s a whole different story.

So if there’s a great potential market for drugs acting at new targets, if delays due to the need for validation can be somewhat ameliorated by front loading, and if there’s sometimes even a workaround for the problem of not having tool compounds, why doesn’t everyone work on novel targets? If you remember the major advantage of working on established targets, the fact that resulting drugs have a higher-than-average probability of making it to market, you can probably predict what the major disadvantage of working on novel targets is.

By simply using readily available online resources like BioSpace (www.biospace.com) and DrugResearcher.com (www.drugresearcher.com), without delving any deeper it’s likely that you’ll be able to locate an average of at least one report of an exciting potential new target *every day*. Today, for example, my usual sources alerted me to *Heme oxygenase-1 (HO-1)* and its potential role in Kaposi’s sarcoma,73 while yesterday brought reports on the role of *lysyl oxidase (LOX)* in cancer metastasis74 and the promise of drugs acting on *metabotropic glutamate receptor (mGluR)* in Fragile X syndrome.75 All three represent laudable accomplishments that hold promise for victims of cancer or the most common inherited form of mental impairment. But how many of these will eventually result in marketed drugs?

If we make the conservative estimate that such novel targets are reported at the rate of one per day and the guesstimate that for each of these another has been uncovered in industry but not talked about (secrecy being especially important in the early stages of a project) that gives
us 730 new targets every year. But in any given year only a few drugs which act on novel targets are approved. Seven or eight such approvals would make for a banner year. Glossing over factors like the time lag between TV and drug approval, unpredictable changes in the numbers of approved drugs, etc., you can see that at least 99% of all new targets do not lead to marketed drugs. A recent estimate states that although 17% of projects based on established targets are likely to enter preclinical development (that is, result in optimized leads evaluated in the preclinical stage), only 3% of projects based on new targets will ever get that far. All of this should in no way be seen as taking credit away from the dedicated and often brilliant researchers who did the hard work of uncovering them originally.

Instead, their “failure to launch” could be due to any number of things. As anyone who’s ever made compounds described in the literature knows, experiments sometimes end up being difficult for other researchers to reproduce. Or no real follow-up might have been done because companies believed the markets to be too small or too difficult or it just didn’t fit into their research portfolio or was lost in the shuffle of project competition and corporate mergers. Perhaps, they felt that the published work fell into the “NIH” (not invented here) category and wasn’t up to their own, internal standards. Perhaps for a variety of reasons no suitable compounds could be found, as they couldn’t for the then-new renin inhibitors for 30 years. For one reason or another, the proposed target just never was able to work its way up the rungs of the validation ladder. There might be many reasons for this, but the bottom line is that while novel targets are plentiful, novel approved drugs acting upon them are not. More’s the pity.

### 5.5 Targets Arising from Phenotype or High-Content Screening

#### 5.5.1 Phenotype Screening Versus Target Screening

Cellular assays play an important role in every project. For a project based upon a given target, initial in vitro assays are usually followed up by target assays which take place in cells. For example, a proteasome inhibitor project would probably begin by screening compounds for inhibition of isolated 20S proteasomes. This would be followed up by studies where the active compounds were incubated with cells and then the levels of some marker of cellular proteasome function like the protein p21, which tends to accumulate when proteasomes aren’t actively degrading it, are measured. If compounds active in the first assay work in cells everyone will feel more confident that they’re onto something. If not, the usual villains of cell impermeability or cellular efflux will have to be faced down if the project is to continue. Think of these kinds of simple cellular assays as a rite of passage for most target-driven projects.

But cellular assays can also be used in a totally different way. To understand this, we need to take a very brief look at genetics. A time-tested genetic technique for finding the roles of specific proteins in biochemical pathways is to use mutagenesis in some living organism to randomly interfere with gene products, to focus in on the mutations that cause observable effects (the phenotype), and then to go back and identify which genetic alterations and resulting perturbed proteins were responsible. This process goes by the name of *forward genetics*. Mutagenesis here is usually accomplished through the random mutations induced by a nasty chemical mutagen like ENU, ethylnitrosourea. If instead, we use as our
perturbogen\textsuperscript{77} a library of small molecules which, rather than causing mutations, might interact with cellular proteins through inhibition, antagonism, or agonism, the process becomes what is known as forward chemical genetics.\textsuperscript{78} Figure 5.10 is a simple schematic illustrating how this works.

Starting with a clean slate, that is, even before a target has been identified, cellular assays can be devised which measure observable phenotypes, like viability, morphological changes, or give some other readout like protein translocation which reflects a given disease state and not just the striking of one prespecified target. The most direct example of this would be for a type of pathogenic bacteria where compounds inducing death (bacteriocidal) or failure to proliferate (bacteriostatic) would be considered potentially useful, assuming they didn’t do the same to human cells—a very big assumption. In this case it wouldn’t matter, at least initially, what target was involved. Everything could be thrown into such an assay and the mechanism of action for interesting hits could be sorted out later. This general sort of strategy has been used for many decades in antibiotics, although not necessarily on a high-throughput basis, and is reminiscent in a way of those interesting Petri dishes of Sir Alexander Fleming which gave the world penicillin. Inhibition of proliferation in a cancer cell line, which can happen through any number of different mechanisms, is an example of a simple multitarget cellular assay such a strategy might use.

This approach stands the normal process completely on its head: instead of going from target to drug, it proceeds from “drug” (in the sense of hit or lead) to target. In a way, this represents a return to the older drug discovery methodology that prevailed in the days of broth extract screening and tissue or even whole animal primary assays for synthesized compounds, back before the new field of molecular biology shone the bright spotlight of scientific inquiry on specific molecular targets. Although nobody suggests that we turn that spotlight off, many people feel that illuminating the rest of the stage as well could only help. In that context, phenotype screening and other types of multitarget or high-content screening (HCS, which usually implies the use of subcellular fluorescence imaging) represent a valuable addition to the armamentarium of drug discovery.\textsuperscript{79}
There’s a lot of diversity in the kind of assays that can be carried out, some of them having only been enabled by recent technologies, such as the ability to detect cell morphology changes by microscopy in a high-throughput format. Cells (which of course don’t constitute an in vivo system except in the case of single-celled organisms) can also be engineered to express things like known chemotherapy resistance factors and then compounds screened for their ability to restore the antiproliferative effects of a cancer drug like doxorubicin. In some cases where traditional cellular models are hard to come by, such as HCV, which isn’t amenable to simple infection and reproduction in cultured hepatocytes, cells can be engineered to express self-amplifying “minigenomes” called replicons containing the non-structural viral proteins, which represent most (but not all) conceivable antiviral targets. A compound active in such an assay might be hitting any of them.

True organisms, rather than cells, such as nematodes (C. elegans) and zebrafish embryos (Danio rerio) can be used in multitarget assays. Readouts can be as simple as fluorescence or as complex as a detailed analysis of modified protein pathways. Even beginning to understand all the possible variations on this type of screening strategy is a daunting prospect, but however many there are and however complex they may be, they all have a couple of things in common. They allow one to screen for drugs in a complex living system, which is more “real world” than an isolated enzyme or receptor binding assay, and to “cast a much wider net than target-by-target approaches.”

An overall summary of the advantages and disadvantages of using this type of drug discovery methodology is shown in Box 5.4. The single biggest advantage of this method is that by definition hits give real effects in cells, rather than an effect on an isolated target that may or may not work in a cellular milieu when tested later. Note, however, that initial hits, which are defined by some arbitrary level of response in a high- or medium-throughput screen which has been validated by running a positive control, require a certain amount of follow-up before any conclusions can be drawn. It’s not as simple as looking at the raw data for the screen. Artifacts must be ruled out, identity, purity, and good dose–response need to be established, all of which are further discussed in Chapter 6.

But if all these follow-up experiments have been done and hits are confirmed, especially if additional evidences like confirmation in a different cell line or apparent non-flat SAR for structurally related compounds in the screening set are present, then what’s left at the end of the day (or month, more likely) is a hit that produces a biological effect in a disease-relevant cellular assay, and the value of that shouldn’t be underestimated. Because the confirmatory experiments have ruled out insolubility (by visual examination, nephelometry, etc.), toxicity (by cell viability experiments), and other bad things, the genuine hits should provide unusually good starting points for compound optimization. Furthermore, if the target of the hits turns out to be intracellular, which many are, you can be sure that the compounds have a certain degree of cell permeability. All in all, what you end up with is a starting point that’s far ahead of those you’d get in other types of projects.

Furthermore, with this approach it’s possible to go in “blind,” without prejudging which targets may be present and what type of compounds should hit them, which may be more likely to result in the discovery of new targets and novel compounds if it finds anything at all. In cases where such information isn’t known this sort of phenomenological approach may be the only way to go. But if target information is available, bias can still be built in if desired,
for example by selecting a library of fused nitrogen heterocycles which look like kinase inhibitors in a cancer cellular assay.

Another advantage of this approach that may not be obvious at first is that compounds whose effects depend on hitting not just one but several different targets, which may have been discarded in more “target-centric” approaches precisely because of non-selectivity or just because such combinations weren’t considered, can still be identified. In fact, cellular screens may be the best way to study things like kinase inhibitors, where one recent study concludes that “cellular responses induced by multiplex protein kinase inhibitors may be an emergent property that cannot be understood fully by considering only the sum of individual inhibitor–kinase interactions.” And a number of such multitargeted compounds do populate the world of known drugs as we’ll see later. A particularly appealing therapeutic area for such multitargeted compounds is oncology, where many different gene products are typically mutated and phenotype assays (antiproliferative, pro-apoptotic, etc.) are straightforward.

Yet another unique aspect of phenotype screening is that it might even detect prodrugs, where the active species is not the compound applied but some transformed version of it. Isoniazid (isonicotinyl hydrazide), a tuberculosis drug in widespread use for over half a century, requires reaction with a specific peroxidase followed by reaction with NAD\(^+\) and NADP\(^+\) to generate the nucleotide adducts that are the active species. A recent study using a chemical genetics approach identified one of the targets of a specific nucleotide adduct as bacterial dihydrofolate reductase. Isoniazid originally came out of a simple bacteriocidal
that is, phenotype) assay, but had it been discovered today, elucidating the multiple MOAs of its several important metabolites would represent quite a challenge to the drug’s development. More importantly, though, it would be virtually impossible to discover through a target-centered approach.

Using the phenotype or multitarget or high-content approach to drug discovery has its downside as well. First of all, the high-tech nature of many of the sophisticated assay techniques involved means that licenses need to be taken out and often expensive license fees paid before assays can even be done. Examples include the use of green fluorescent protein (GFP), a frequently employed reporter gene, and the previously mentioned HCV replicon assay. HCV is a particularly expensive area to do research in since fees will also need to be paid to Chiron (now Novartis) which, owing to the groundbreaking work done by its scientists, owns the viral genome patent. Other expenses associated with such a project will include specialized equipment as well as costs associated with the frequently lengthy assay development process itself. This can be trickier than the usual enzyme or receptor screens, cells being somewhat delicate and capricious.

The fact is that in some areas like CNS indications, good disease-relevant cellular assays may just not be available owing to the complex nature of the disorders, which may not be adequately modeled in a single, culturable cell type. Even in therapeutic areas where cellular models exist, like cancer, there’s no guarantee that the transformed cell lines typically used (like the NCI-60, discussed later) will respond to a drug in the same way as an actual tumor, which is composed of many primary cell types living and dying in a variety of different aerobic and anaerobic environments. Even cancer cell lines themselves exhibit different behaviors in the presence of other cells, which can secrete factors to influence their proliferation or death. So it should come as no surprise that a drug with Level I validation in such an assay might make it no further. It is, after all, just a model, not something you expect to get into and drive away.

But the biggest drawback in using this approach to drug discovery becomes obvious once hits have been identified and confirmed—what exactly are they hitting? There are widely used drugs (including, up until recent years, aspirin) which act upon unknown targets. Drug candidates can still enter the clinic that way, a recent example being InterMune’s Pirfenidone (Figure 5.11), an investigational new drug for idiopathic pulmonary fibrosis (IPF) for which a proposed MOA (p38γ MAP kinase inhibition) was first suggested during Phase III testing. Such drugs might even make it to market without a known MOA. But the reality is that a drug going into the clinic without a known target is at a serious disadvantage, and many companies will be reluctant to pursue it for that reason. Everyone involved, including FDA, feels more comfortable if the MOA is known. Clinicians in particular need a feeling for how it would fit in with existing therapies based on its mechanism. Although drugs with unknown MOAs have been and continued to be approved, going in without one makes for a rougher road.

Knowing the MOA at an early stage would be far more preferable. That might allow the chemists to use structure based drug design (SBDD) for more rapid optimization. Knowing what the target is can also give scientists an idea of what the antitargets, related proteins possibly responsible for toxicity, might be. So in the real world identification of hits using phenotype screening will inevitably be followed by attempts to determine what target or targets are giving the effect. Target identification, in this case based on real cellular effects instead of correlative evidence and theory, has not been
avoided but only postponed. When it comes, it will be because an exciting new lead has been found from phenotype screening. Timelines will therefore be short and resources will need to be poured into it. It will be a far from trivial effort and may even call for more biologists than a small company has available. Additionally, chemists will probably be needed early on. This could be for several reasons. The initial hit can often be too weak (say 50 μM) to be useful in determining the target so that a certain amount of blind optimization needs to be done up front. Structural modifications involving linkers and/or labels may need to be carried out to fish out the target protein, as discussed below. So much excitement can be generated that chemists become eager to just start making analogs. All of these cases require the commitment of at least a few chemists to a still highly speculative preproject. These FTEs will need to be planned for in advance to avoid premature (and predictable) project termination for lack of resources later on.

When all is said and done, there are three possible outcomes for TID in this approach. If the target for a phenotype lead is discovered and found to fit into the established target bin, which is often the case, the company will need to decide whether or not it should be pursued. For example, uncovering an antibacterial agent that acted by inhibiting DNA gyrase, a target of fluoroquinolones like Cipro (ciprofloxacin) (Figure 5.11), would put the company in the

![Tacrolimus](image1)

![Pirfenidone](image2)

![Ciprofloxacin](image3)

![Leptomycin B](image4)

**Figure 5.11** Tacrolimus (FK506), Cipro (ciprofloxacin), Pirfenidone, and Leptomycin B.
position of having to go over the pros and cons mentioned in Section 5.2. A business decision would be called for, and it might or might not be a favorable one. If it was, though, they’d at least be starting out with a cell permeable and hopefully novel lead that could jumpstart the project in the LO stage.

The second possibility is that a novel target has been uncovered. Here all of the advantages of working on novel targets apply and one of the disadvantages, that of requiring validation, has at least been mitigated by starting out with a compound that works in cells. Although the higher risk of failure due to lack of further validation remains for such projects, they will at least excite everyone involved.

But there’s a third possibility, too, as shown in Figure 5.1, that the timeline originally set for target determination runs out and the target has still not been identified. This is the most disturbing of the three outcomes, but is a real possibility owing to the often rate-limiting nature of the process. With all of the modern techniques (see below) and all the resources in terms of people, expertise, and funding that can be brought to bear it should be possible to identify the target or targets for any cellular leads. But that assumes unlimited resources including time, a situation that never exists in the real world. Determining the targets of compounds found through phenotype screening is never trivial and there’s a real possibility that the budget and time requirements could end up being insupportable. The decision about whether to continue the investigation, increase the resources, and push out the timelines will need to be made. A second decision about whether or not to go into serious LO “blind,” without that critical target information, will also need to follow. After all, it will be pointed out, drugs have made it to the clinic and even to approval this way!

5.5.2 Elucidation of Phenotype-Derived Targets

Whatever the outcome, we can see that TID plays a critical role in phenotype-based projects and it’s been the subject of recent reviews. It’s now time to look at some of the more common techniques used in this important step, as shown in Box 5.5.

In many cases some idea of what the potential targets might be will exist before compounds have even been screened. There are always “the usual suspects” based on known pathways—tubulin binders in antiproliferative assays, MDR1 and MRP1 gene products in cancer drug resistance, etc. An educated guess and some follow-up can often prove successful in such cases. A more elegant example of this kind of knowledge-based approach was reported by Kau et al. Library high-content screening for inhibitors of FOXO1a (a transcription factor) nuclear export turned up a number of small molecule hits that were Michael acceptors. A known compound, Leptomycin B (Figure 5.11), which is also a Michael acceptor, acts upon the nuclear export receptor CRM1, by forming a covalent adduct with the protein’s Cys528 sulphydryl group. The hypothesis that the new hits were also acting on this target via this mechanism was subsequently borne out by follow-up experiments.

A very direct method for identifying molecular targets of a phenotype screening hit involves affinity purification. A strong binding interaction between the target and the drug, at least in theory, means that if the drug is attached to a solid support through some part of the molecule unimportant to interaction with the target and an extract or lysate containing the target is passed over it, the target will bind sufficiently well that the many other cellular proteins in the “soup” can be washed away. Stronger conditions such as elution with buffers
of high-ionic strength or cleavage of the linker between the drug–protein complex and the solid support should allow for isolation of the drug–target complex which can then be characterized and analyzed, usually by LC/MS/MS, to reveal the identity of the target protein.

The poster child for this sort of approach was the groundbreaking work of Dr Stuart Schreiber and his group in determining the biological target of the important immune suppressant FK506 (Figure 5.11), the protein called FKBP.93 Not only did this work elucidate the MOA for a drug (called Tacrolimus) which allows organ transplant patients to avoid transplant rejection and live for many years, but also, according to Dr Thomas Kodadek, Professor of Medicine at the University of Texas Southwest Medical Center, it “had a lot to do with convincing biologists that there really might be something to the idea of using chemistry as a key tool in molecular biology, while also suggesting to a recalcitrant synthetic chemistry community that it might be OK to work with real biological systems.”94 In the case of the FK506–FKBP complex the receptor was relatively abundant in cells and the ligand possessed high (single digit nanomolar) affinity, two things that are not true in many other cases.

“The bad news,” he continues, “is that it created the false impression among chemists without much classical biochemical experience that the target identification problem could be solved in most cases using this simple affinity chromatography approach. This view soon crashed upon the rocks of reality.”94 In the case of the FK506–FKBP complex the receptor was relatively abundant in cells and the ligand possessed high (single digit nanomolar) affinity, two things that are not true in many other cases.

Phenotype screening more commonly turns up leads in the double or perhaps single digit micromolar range, and the problem on using affinity purification then becomes one of signal-to-noise: washing away “sticky” non-target proteins without also washing away the relevant, low affinity target. Additionally, having to chemically modify the original hit so that it can be linked to a solid support isn’t necessarily trivial, especially in the absence of knowledge as to what’s important to target binding. A kind of preliminary SAR study may need to be done

---

**Box 5.5 Some Methods Used for Target Identification in Phenotype Screening**

- Hypotheses based on known pathways
- Affinity purifications
- Screening of expressed protein libraries
- Phage display
- 3-Hybrid systems
- Cellular microarrays
- Differential protein expression
- Transcriptional profiling
- Modulation by or comparison with genetic methods
with linkers attached to various positions in the hit molecule in order to find an acceptable spot. One interesting approach is to build the linkers into the original screening library, as exemplified by recent work where the targets of a triazine hit in a zebrafish embryo morphology screen were identified as components of the 40S ribosome.\textsuperscript{84} Another nice example of the use of affinity purification for phenotype TID has been reported by Tanaka et al., where a morphological phenotype screen revealed that a hydroxyl-containing analog of the kinase inhibitor PP gave rise to unique cellular effects, which affinity purification traced to a non-kinase target, CBR1.\textsuperscript{95}

Thus affinity purification has proven to be successful from time to time but, like every other means of TID, it’s by no means universally so. It’s then not surprising that many different methodologies have been developed. Quite a few of these involve using the techniques of molecular biology to express large number of proteins, exposing these to tagged or tethered small molecule hits, detecting their interaction and determining the target involved. Among these are phage display,\textsuperscript{96} and three hybrid systems. The latter, which is normally used in yeast has also been used in mammalian systems,\textsuperscript{97} involves linking the hit to a known ligand for a protein receptor, then using this bifunctional molecule (the “bait”) to bring together both domains required for the activation of a reporter gene but only in the presence of the hit’s target.\textsuperscript{98}

Another technique involves microarrays of cells engineered with different plasmids to overexpress various proteins.\textsuperscript{99} Use of an appropriately labeled hit allows for TID, although so far literature examples involved only limited subsets of potential protein targets and not genome-wide arrays. The relatively more direct method, which makes use of arrays of proteins themselves bound to a glass slide, can also be used but likewise has its limitations.\textsuperscript{100} As has been noted for years, immobilized arrays of proteins are more difficult to make and are less stable than DNA microarrays, but in this case one potential advantage they have over some of the other identification methods mentioned is that a detection technique called \textit{surface plasmon resonance} (\textit{SPR}), which obviates the need for hit modification via a label or linker, can be used.

In theory, one of the most direct ways of finding a target for a hit acting through an unknown mechanism is to simply look at the effects on cellular proteins by observing the differences in protein expression or composition in the presence and absence of the hit molecule using 2D gels. In practice, this is anything but simple owing to many factors such as the sheer number of proteins involved, normal differences in abundance (perhaps a million fold!) and activities, isolation difficulties, experimental variability, and massive data handling. So many proteins are likely to be differentially expressed that analyzing the data often results in a “haystack of needles.”\textsuperscript{101} The real target could be indistinguishable in the crowd of differentially expressed proteins, hiding in plain sight. That being said, this technique has had some successes. In one such instance, a comparison of 2D gels for an antiproliferative natural product, LAF389, showed a small shift in a protein which was traced to the incorporation of a single N-terminal methionine residue into the treated cells, suggesting that the compound was acting as an inhibitor of \textit{methionine aminopeptidases} (\textit{MetAP}s).\textsuperscript{102}

A major simplification of this kind of proteomic technique involves limiting the number of proteins studied to a smaller subset of gene products such as cysteine proteases or metalloproteases where appropriate \textit{activity based probes} (\textit{ABPs}) exist. By virtue of their target binding mechanism, which depends on activity and not abundance, another obfuscating
factor is eschewed as well.\textsuperscript{103} The relevance of this kind of strategy is probably limited to cases where the gene family of interest has been targeted by a directed library in the first place and even then it couldn’t rule out effects on the other, non-family targets per se. But within these tight constraints it’s a powerful method for identifying targets.

A variation on the whole genome proteomics involves transcriptional profiling, a comparison of a hit’s pattern of gene expression, or “signature,” to that of other perturbogens such as known drugs, using DNA chips. Figure 5.12 illustrates this concept. A collection of known transcriptional profiles is required for comparison and this might be obtained from a public database or a library of several hundred to several thousand known drugs compiled and tested for this purpose.

An example involves the NCI-60, where researchers have analyzed the effects of more than 1400 compounds on protein expression in the 60 cell lines as determined by DNA microarrays for more than 8000 genes.\textsuperscript{104} An analysis of the truly mind-boggling amount of data revealed clusters of drugs which acted by similar mechanisms (tubulin stabilizers, DHFR inhibitors, etc.). Profiling a hit from anticancer phenotype screening on such a gene chip and comparing it to the known patterns might place it in the appropriate MOA bin.

Another approach would be not to use transcriptomics or proteomics per se, but to do a direct comparison of growth inhibitory values in some or all of the NCI-60 cell lines with those of known drugs using the online program COMPARE. As its name implies, drugs acting through the same mechanism tend to display similar profiles here.\textsuperscript{105} Figure 5.13

\textbf{Figure 5.12} Identifying molecular targets using expression profiling. (Adapted with permission from Butcher, RA, Schreiber, SL. Using genome-wide transcriptional profiling to elucidate small-molecule mechanisms. \textit{Curr. Opin. Chem. Biol.} \textbf{2005}, \textit{9}, 25–30, copyright 2005, Elsevier.)
shows part of a profile for one anticancer drug, Vincristine, with red bars to the left indicating less than average activity against that cell line, while those going to the right indicate greater than average potency in the corresponding cells. Regardless of the readout used, these signature-based approaches can provide valuable mechanistic insights, but have limited utility in identifying novel targets, which may have equally novel signatures. Still, a “unique” profile here is a data point well worth having.
Transcriptional profiling of an unknown against a custom-built signature database is illustrated by a study in which the profiles for a small set of antibacterial compounds were obtained and used to classify the relevant MOAs for others\textsuperscript{106} and another in which a 2036-member library of small molecules was developed and used to elucidate MOAs for 85 antiproliferative agents.\textsuperscript{107} Such comparison of databases are not only limited to small molecule perturbogens but can also make use of profiles generated through genetic methods such as deletion mutants in yeast, as an excellent recent review of the subject points out.\textsuperscript{108}

In fact, in an ironic twist of fate, true genetic methods like gene-deletion mutants in yeast are often helpful in elucidating the targets of hits obtained through chemical genetics. Here the phenotype, not the expression profile as in the case above, is the observed variable, and collections where almost every gene product in yeast has been knocked out have been generated.\textsuperscript{109} A mutation which mimics the phenotype induced by the hit might implicate its MOA. This is called \textit{loss of function (LOF)} screening. Alternatively, a coded library of heterozygous yeast mutants can be treated with the hit to look for mutations causing sensitization to its effects—in effect a dose–response experiment where half the “dose” is induced by deletion of a known gene and the other half by the small molecule, a process called \textit{haploinsufficiency profiling}.\textsuperscript{110}

When this was done with Cincreasin (Figure 5.14), a simple small molecule identified through phenotype screening, one of the deletions conferring the greatest sensitivity was Mps1, the protein product of which has a spindle-checkpoint function. Follow-on experiments demonstrated that this was in fact the target of the hit.\textsuperscript{111} The opposite kind of modification, making the cells less sensitive to drug which might indicate some compensatory or multidrug resistance mechanism, can also be useful. Yeast, of course, is not man, but it’s a lot easier to manipulate, and the pathways involved can be very similar in both.\textsuperscript{112}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure5_14}
\caption{The cancer drug Vincristine and the Mps1 inhibitor Cincreasin.}
\end{figure}
5.6 In Conclusion

This chapter has presented a brief overview of how new drug discovery projects begin, some of the science and business decisions that go into them, what types of projects researchers can expect to see, and a little about how targets are identified and validated. Although this knowledge hasn’t exactly been a state secret, trying to gather it up from several disciplines and a myriad of scattered information sources means that it’s often difficult for those new to the industry to get the big picture. This overview obviously can’t do justice to the depth and subtlety of knowledge involved in each of the fields mentioned. But the complexities of the drug discovery process are such that even with a profound, in depth understanding of all of these topics, which could only be the result of much effort over many years, the researcher would still understand only a part of the drug discovery process. For example, where exactly do screening compounds come from and what technologies are used to screen for hits?

Notes

1. DiMasi, J.A., Paquette, C. The economics of follow-on drug research and development. Trends in entry rates and the timing of development. *Pharmacoeconomics* 2004, 22, Suppl. 2, 1–14.
2. For more about the economics involved in drug entry order, see Cohen, F.J. Entry order as a consideration for innovation strategies. *Nat. Rev. Drug Discov.* 2006, 5, 285–294.
3. Knowles, J, Gromo, G. Target selection in drug discovery. *Nat. Rev. Drug Discov.* 2003, 2, 63–69.
4. Source: *MedAd News*. See Appendix: “The Periodic Table of Drugs”.
5. Tobert, J.A. Lovastatin and beyond: The history of the HMG-CoA reductase inhibitors. *Nat. Rev. Drug Discov.* 2003, 2, 517–526.
6. An interesting account of this is provided in Chapter 9, “Atorvastatin Calcium (Lipitor)” in *Contemporary Drug Synthesis*, Li, J.-J, Johnson, D.J., Sliskovic, D.R., Roth, B.D., Eds. (Wiley Interscience, Hoboken, NJ, 2004).
7. See Marris, E. Diabetes drugs under scrutiny in a post-Vioxx world. *Nat. Rev. Drug Discov.* 2007, 6, 505–507.
8. Herrera, S. Price controls: preparing for the unthinkable. *Nat. Biotechnol.* 2006, 24, 257–260. For an interesting discussion of price controls and other concerns impacting drug discovery research, see Santoro, M.A., Gorrie, T.M., Eds. *Ethics and the Pharmaceutical Industry* (Cambridge University Press, Cambridge, 2005).
9. For an excellent review, see Palladino, M.A. et al. Anti-TNF-α Therapies: The Next Generation. *Nat. Rev. Drug Discov.* 2003, 2, 736–746.
10. For more on the thalidomide story, see Chapter 5 in Drews, J. *In Quest of Tomorrow’s Medicines* (Springer-Verlag, New York), 1999.
11. He, M.M. et al. Small molecule inhibition of TNF-α. *Science* 2005, 310, 1022–1025.
12. For more about small molecule TNF disruptors and small molecules interference with protein-protein interactions in general, see Wells, J.A., McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* 2007, 450, 1001–1009.
13. Kubinyi, H. Drug research: myth, hype, and reality. *Nat. Rev. Drug Discov.* 2003, 2, 665–668.
14. Gerhold, D.L. et al. Better therapeutics through microarrays. *Nat. Genet. Suppl.* 2002, 32, 547–552.
15. Gygi, S.P. et al. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 1999, 19, 1720–1730.
16. Quoted in Fitzgerald, K. Less correlation, more causation Curr. Drug Discov. January, 2002, 31–33.

17. Dr Szymkowski has presented these delightful analogies in slide presentations and in personal communications with this author.

18. See Szymkowski, D.E. Too many targets, not enough target validation. Drug Discov. Today 2001, 6, 398–399.

19. Kopec, K.K. et al. Target identification and validation in drug discovery: the role of proteomics. Biochem. Pharmacol. 2005, 69, 1133–1139.

20. See especially comments on target validation staging contained in Drews, J. Drug discovery: A historical perspective. Science 2000, 287, 1960–1964, comments on the definition of target validation by Dr Christopher Miller in Szymkowski, DE. Target validation joins the pharma fold. Drug Discov. Today 2003, 2, 8–9, and Winkler, H. Target validation requirements in the pharmaceutical industry Drug Discov. Today 2003, 2, 69–71.

21. With the exception of targeted chemotherapeutic agents like imatinib, where efficacy in models not involving BCR-ABL kinase wouldn’t be expected.

22. Zambrowicz, B.P. et al. Predicting drug efficacy: knockouts model pipeline drugs of the pharmaceutical industry. Curr. Opin. Pharmacol. 2003, 3, 563–570.

23. Saftig, P. et al. Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. Proc. Natl. Acad. Sci. USA 1998, 95, 13453–13458.

24. Smith, W.W., Abdel-Meguid, SS. Cathepsin K as a target for the treatment of osteoporosis. Expert Opin. Ther. Pat. 1999, 9, 683–694.

25. Zambrowicz, B.P., Sands, A.T. Knockouts model the 100 best-selling drugs—will they model the next 100? Nature Rev. Drug Discov. 2002, 2, 38–51.

26. See www.lexpharma.com/discovery/genome500.htm for the Lexicon effort and www.nih.gov/science/models/mouse/knockout/ for the NIH project.

27. Ferrara, N. et al. Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. Nature 1996, 380, 439–442.

28. See Pearson, H. Surviving a knockout blow. Nature 2002, 415, 8–9.

29. Kubota, N. et al. PPARγ mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. Mol. Cell 1999, 4, 597–609.

30. Yamauchi, T. et al. The mechanism by which both heterozygous peroxisome proliferator-activated receptor γ (PPAR-γ) and PPAR-γ agonist improve insulin resistance. J. Biol. Chem. 2001, 276, 41245–41254.

31. Knight, Z.A., Shokat KM. Chemical genetics: Where genetics and pharmacology meet. Cell 2007, 128, 425–430.

32. Zamecnik, P.C., Stephenson, ML. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc. Natl. Acad. Sci. USA 1978, 75, 280–284.

33. Crooke, ST. Molecular mechanisms of action of antisense drugs. Biochim. Biophys. Acta 1999, 1489, 31–44.

34. For a good review of AOs, see Dias, N., Stein, C.A. Antisense oligonucleotides: Basic concepts and mechanisms. Mol. Cancer Ther. 2002, 1, 347–355.

35. Myers, K.J., Dean, N.M. Sensible use of antisense: how to use oligonucleotides as research tools. Trends Pharmacol. Sci. 2000, 21, 19–23.

36. See www.isispharm.com/product_pipeline.html#301012.

37. See Rubinstein, M. et al. A review of antisense oligonucleotides in the treatment of human disease. Drugs Future 2004, 29, 893–909, and Potera, C. Antisense—down, but not out Nature Biotech. 2007, 25, 497–499.

38. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 1998, 391, 806–811.
39. Arts, G.-J. et al. Adenoviral vectors expressing siRNAs for discovery and validation of gene function. _Genome Res._ **2003**, 13, 2325–2332.

40. Paroo, Z., Corey, D.R. Challenges for RNAi _in vivo_. _Trends Biotechnol._ **2004**, 22, 390–394.

41. Chen, X. et al. Chemical modification of gene silencing oligonucleotides for drug discovery and development. _Drug Discov. Today_ **2005**, 10, 587-593.

42. For examples, see Grunweller, A. et al. Comparison of different antisense strategies in mammalian cells using locked nucleic acids, 2′-O-methyl RNA, phosphorothioates and small interfering RNA. _Nucleic Acid Res._ **2003**, 31, 3185–3193 and Vickers, T.A. Efficient reduction of target RNAs by small interfering RNA and RNAse H-dependent antisense agents. A comparative analysis. _J. Biol. Chem._ **2003**, 278, 7108–7118.

43. Jia, L. et al. siRNA targeted against matrix metalloproteinase 11 inhibits the metastatic capability of murine hepatocarcinoma cell Hca-F to lymph nodes. _Int. J. Biochem. Cell Biol._, 2007, 39, 2049–2062.

44. See Matsukura, S. et al. Establishment of conditional vectors for hairpin siRNA knockdowns _Nucleic Acid Res._ **2003**, 31, e71–e77.

45. Hoeflich, K.P. et al. Oncogenic BRAF is required for tumor growth and maintenance in melanoma models. _Cancer Res._ **2006**, 66, 999–1006.

46. Scacheri, P.C. et al. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. _Proc. Natl. Acad. Sci. U.S.A._ **2004**, 101, 1892–1897.

47. Jackson, A.L. et al. Expression profiling reveals off-target gene regulation by RNAi. _Nat. Biotechnol._ **2003**, 21, 635–637.

48. He, L., Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. _Nat. Rev. Genet._ **2004**, 5, 522–531.

49. Fitzgerald, K. RNAi versus small molecules: Different mechanisms and specificities can lead to different outcomes. _Curr. Opin. Drug Discov. Devl._ **2005**, 8, 557–566.

50. Cullen, B.R. Enhancing and confirming the specificity of RNAi experiments. _Nature Methods_ **2006**, 3, 677–681.

51. Bridge, A.J. Induction of an interferon response by RNAi vectors in mammalian cells. _Nat. Genet._ **2003**, 34, 263–264.

52. See Eggert, U.S. et al. Small molecules in an RNAi world. _Mol. BioSyst._ **2006**, 2, 93–96.

53. Whitehurst, A.W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. _Nature_ **2007**, 446, 815–819.

54. Rader, C. Antibody libraries in drug and target discovery. _Drug Discov. Today_ **2001**, 6, 36–43.

55. Kneuer, C. et al. Selectins—potential pharmacological targets? _Drug Discov. Today_ **2006**, 11, 1034–1040.

56. Biocca, S., Cattaneo, A. Intracellular immunization: antibody targeting to subcellular compartments. _Trends Cell Biol._ **1995**, 5, 248–252.

57. See Stocks, M.R. Intrabodies: production and promise. _Drug Discov. Today_ **2004**, 9, 960–966, which is also an excellent review of this subject.

58. Rousselet, N. et al. Inhibition of tumorigenicity and metastasis of human melanoma cells by anti-cathepsin L single chain variable fragment. _Cancer Res._ **2004**, 64, 146–151.

59. Zemel, R. et al. Inhibition of hepatitis C virus NS3-mediated cell transformation by recombinant intracellular antibodies. _J. Hepatol._ **2004**, 40, 1000–1007.

60. See Baines, I.C., Colas, P. Peptide aptamers as guides for small-molecule drug discovery. _Drug Discov. Today_ **2006**, 11, 334–341.

61. See Tuerk, C., Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. _Science_ **1990**, 249, 505–510 and Ellington, A.D., Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. _Nature_ **1990**, 346, 818–822.
62. For a recent review, see Djordjevic, M. SELEX experiments: New prospects, applications and data analysis in inferring regulatory pathways. *Biomol. Eng.* 2007, 24, 179–189.

63. Jenison, R.D. et al. High-resolution molecular discrimination by RNA. *Science* 1994, 263, 1425–1429.

64. For a review of spiegelmers, see Eulberk, D., Klussmann, S. Spiegelmers: Biostable aptamers. *Chembiochem* 2003, 4, 979–983.

65. Helmling, S. et al. Inhibition of ghrelin action in vitro and *in vivo* by an RNA-Spiegelmer. *Proc. Natl. Acad. Sci. USA* 2004, 101, 13174–13179.

66. Shearman, L.P. et al. Ghrelin neutralization by a ribonucleic acid-SPM ameliorates obesity in diet-induced obese mice. *Endocrinology* 2006, 147, 1517–1526.

67. Pestourie, C. et al. Aptamers against extracellular targets for *in vivo* applications. *Biochimie* 2005, 87, 921–930.

68. Famulok, M. et al. Intramers as promising new tools in functional proteomics. *Chem. Biol.* 2001, 8, 931–939.

69. Blank, M., Blind, M. Aptamers as tools for target validation. *Curr. Opin. Chem. Biol.* 2005, 9, 336–342.

70. Lee, J.F. et al. Aptamer therapeutics advance. *Curr. Opin. Chem. Biol.* 2006, 10, 282–289.

71. Capdeville, R et al. Glivec (STI571, Imatinib), a rationally developed, targeted anticancer drug. *Nature Rev. Drug Discov.* 2002, 1, 493–502.

72. For a recent review of progress in HIV drugs, see DeClerq, E. New approaches toward anti-HIV chemotherapy. *J. Med. Chem.* 2005, 48, 1297–1313.

73. Marinissen, M.J. et al. Inhibition of heme oxygenase-1 interferes with the transforming activity of kaposi sarcoma herpesvirus encoded G protein-coupled receptor. *J. Biol. Chem.* 2006, 281, 11332–11346.

74. Erler, J.T. et al. Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* 2006, 440, 1222–1226.

75. Miller, G. A fix for fragile X syndrome? *Science* 2006, 312, 518–521.

76. Source: Accenture news release, January 19, 2005. Pharmaceutical Companies Must Re-think Innovation Strategies to Address Pipeline Challenges, Accenture and CMR Research Finds. http://accenture.com/xd/xd.asp?it=enweb&xd=_dyn/dynamicpressrelease_796.xml

77. Szymkowski, D. Chemical genomics versus orthodox drug development. *Drug Discov. Today* 2003, 8, 157–159.

78. For timely and informative reviews of chemical genetics, see (a) Lokey, R.S. Forward chemical genetics: progress and obstacles on the path to a new pharmacopoeia. *Curr. Opin. Chem. Biol.* 2003, 7, 91–96, and, (b) Smukste, I., Stockwell, B.R. Advances in Chemical Genetics. *Annu. Rev. Genomics Hum. Genet.* 2005, 6, 261–286.

79. For more about HCS, see Kom, K., Krausz, E. Cell-based high-content screening of small-molecule libraries. *Curr. Opin. Chem. Biol.* 2007, 11, 503–510.

80. Smukste, I. et al. Using small molecules to overcome drug resistance induced by a viral oncogene. *Cancer Cell* 2006, 9, 133–146.

81. Bartenschlager, R., Lohmann, V. Novel cell culture systems for the hepatitis C virus. *Antiviral Res.* 2001, 52, 1–17.

82. Bartenschlager, R. Hepatitis C virus replicons: potential role for drug development. *Nat. Rev. Drug Discov.* 2002, 1, 911–916.

83. (a) Kaletta, T., Hengartner, M.O. Finding function in novel targets: *C. elegans* as a model organism. *Nature Rev. Drug Discov.* 2006, 5, 387–399. (b) Kwok, TC. A small-molecule screen in *C. elegans* yields a new calcium channel antagonist. *Nature* 2006, 441, 91–95.

84. Khersonsky, S.M. et al. Facilitated forward chemical genetics using a tagged library and zebrafish embryo screening. *J. Am. Chem. Soc.* 2003, 125, 11804–11805.
85. Butcher, E.C. Can systems biology rescue drug discovery? *Nature Rev. Drug Discov.* 2005, 4, 461–467.
86. Kung, C. et al. Chemical genomic profiling to identify intracellular targets of a multiplex kinase inhibitor. *Proc. Natl. Acad. Sci. USA* 2005, 102, 3587–3592.
87. Argyrou, A. et al. *Mycobacterium tuberculosis* dihydrofolate reductase is a target for isoniazid. *Nat. Struct. Mol. Biol.* 2006, 13, 408–413.
88. Khodarev, N.N. et al. Tumor-endothelium interactions in co-culture: coordinated changes of gene expression profiles and phenotypic properties of endothelial cells. *J. Cell Sci.* 2003, 116, 1013–1022.
89. Drews, J., Ryser, S. Molecular drug targets and genomic sciences. *Nat. Biotechnol.* 1997, 15, 1350–1359.
90. Hart, C.P. Finding the target after screening the phenotype. *Drug Discov. Today* 2005, 7, 513–519.
91. Terstappen, G.C. et al. Target deconvolution strategies in drug discovery. *Nature Rev. Drug Discov.* 2007, 6, 891–903.
92. Kau, T.R. et al. A chemical genetic screen identifies inhibitors of regulated nuclear export of a Forkhead transcription factor in PTEN-deficient tumor cells. *Cancer Cell* 2003, 4, 463–476.
93. Harding, M.W. et al. A receptor for the immuno-suppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature* 1989, 341, 758–760.
94. Burdine, L., Kodadek, T. Target identification in chemical genetics: The (often) missing link. *Chem. Biol.* 2004, 11, 593–597.
95. Tanaka, M. et al. An unbiased cell morphology-based screen for new, biologically active small molecules. *PLoS Biol.* 2005, 3, 764–776.
96. Smith, G.P., Petrenko, V.A. Phage display. *Chem. Rev.* 1997, 97, 391–410.
97. Liberles, S.D. et al. Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc. Natl. Acad. Sci. USA* 1997, 94, 7825–2830.
98. (a) Licitra, E.J., Liu, J.O. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc. Natl. Acad. Sci. USA* 1996, 93, 12817–12821. (b) Lefurgy, S., Cornish, V. Finding Cinderella after the ball: A three-hybrid approach to drug target identification. *Chem. Biol.* 2004, 11, 151–153.
99. (a) Zlauddin, J., Sabatini, D.M. Microarrays of cells expressing defined cDNAs. *Nature* 2001, 411, 107–110. (b) Bailey, S.N. et al. Applications of transfected cell microarrays in high-throughput drug discovery. *Drug Discov. Today* 2002, 7(Suppl.), 1–6.
100. (a) MacBeath, G., Schreiber, S.L. Printing proteins as microarrays for high-throughput function determination. *Science* 2000, 289, 1760–1763. (b) Mitchell, P. A perspective on protein microarrays. *Nat. Biotechnol.* 2002, 20, 225–229.
101. Baak, J.P.A et al. Genomics and proteomics—the way forward. *Ann. Oncol.* 2005, 16 (Suppl. 2), ii30–ii40.
102. Towbin, H. et al. Proteomics-based target identification: bengamides as a new class of methionine aminopeptidase inhibitors. *J. Biol. Chem.* 2003, 278, 52964–52971.
103. (a) Phillips, C.I., Bogyo, M. Proteomics meets microbiology: technical advances in the global mapping of protein expression and function. *Cell. Microbiol.* 2005, 7, 1061–1076. (b) Barglow, K.T., Cravatt, B.F. Discovering disease-associated enzymes by proteome reactivity profiling. *Chem. Biol.* 2004, 11, 1523–1531.
104. Scherf, U. et al. A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.* 2000, 24, 236–244.
105. (a) Weinstein, J.N. et al. An information-intensive approach to the molecular pharmacology of cancer. *Science* 1997, 275, 343–349. (b) Rabow, A.A. et al. Mining the national cancer institute’s tumor-screening database: Identification of compounds with similar cellular activities. *J. Med. Chem.* 2002, 45, 818–840.
106. Hutter, B. et al. Prediction of mechanisms of action of antibacterial compounds by gene expression profiling. *Antimicrob. Agents Chemother.* 2004, 48, 2838–2844.
107. Root, D.E. et al. Biological mechanism profiling using an annotated compound library. *Chem. Biol.* 2003, 10, 881–892.

108. Butcher, R.A., Schreiber, S.L. Using genome-wide transcriptional profiling to elucidate small-molecule mechanisms. *Curr. Opin. Chem. Biol.* 2005, 9, 25–30.

109. Giaever, G. et al. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 2002, 418, 387–391.

110. Lum, P.Y. et al. Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* 2004, 116, 121–137. For a brief, simplified explanation of the method in this paper, see Hughes, T. et al. *Cell* 2004, 116, 5–7.

111. Dorer, R.K. et al. A small-molecule inhibitor of Mps1 blocks the spindle-checkpoint response to a lack of tension on mitotic chromosomes. *Curr. Biol.* 2005, 15, 1070–1076.

112. For more on the utility of yeast in drug discovery, see Auerbach, D. et al. Drug discovery using yeast as a model system: A functional genomic and proteomic view. *Curr. Proteomics* 2005, 2, 1–13.