DRUG DISCOVERY TODAY TARGETS

Finding new drug targets in the 21st century

Mark A. Lindsay

The past 30 years have witnessed a steady decline in the number of new drug targets. This review concentrates on the initial process of target identification and argues that current problems have resulted from a decrease in clinical research, an overemphasis on the discovery of new targets through an understanding of the molecular causes of disease and the adoption of cell and animal models that are poor predictors of human disease. To resolve this situation, we argue for increased clinical research and show that an intervention at the physiological level, using drugs to target at the extracellular signalling pathways, will facilitate identification of novel drug targets in the 21st century.

Despite the enormous amount of public and private investment in biomedical research, there has not been the expected increase in new treatments for human disease. This is most visible in the area of drug development, where the number of innovative new drugs has remained relatively constant at ~20 per year between 1990 and 2004 (www.fda.gov/cder/rdmt/pstable.htm), despite the amount of money being spent on research and development having increased at a rate of 13% each year since 1970 [1].

The reason for this decline in new drugs has been widely debated and is likely to be multifactorial [2,3]. However, it is generally agreed that the initial identification and validation of disease-relevant targets, a process entitled target discovery, is an essential first step in the drug discovery pathway. Because there are already several excellent general overviews in this field [4–6], this article will focus specifically upon the process of target identification and other potential problem areas. In particular, it will concentrate upon issues such as the decline in clinical medicine, the strengths and weaknesses of different approaches to target identification and the limitations associated with existing cell and animal models of disease. In the final section, I will argue that these issues could be substantially overcome by adopting a clinical- and physiological- orientated approach to the identification of novel drug targets.

The decline of clinical medicine

Clinical research is crucial to drug discovery because all drug discovery programmes begin with the identification of disease relevant phenotypes. In the past, these studies established the importance of high blood pressure in stroke, increased blood cholesterol in atherosclerosis and smoking in causing lung cancer. Therefore, it is not surprising that the steady decline in number of novel targets has been correlated with an overall drop in the clinical research disciplines such as physiology, anatomy and histology [7]. Several factors have been implicated in this fall, including financial disincentives, the increased legislation surrounding human studies and the rise of evidence-based research with its emphasis upon

Mark A. Lindsay
Biopharmaceutics Research Group,
Airway Disease,
National Heart and Lung Institute,
Dovehouse Street,
Imperial College,
London SW3 6LY,
UK
e-mail:
m.lindsay@imperial.ac.uk

1359-6446/05/$ – see front matter PII: S1359-6446(05)03670-6

expensive large multi-centred clinical trials. Thus, although the latter has been crucial in confirming the role of individual phenotypes in disease, often this has been at the expense of the patient-orientated observations, which provided the original hypothesis. However, arguably the most important influence has been the shift in focus towards an understanding of the genetic and mechanistic basis of disease. The meteoric rise of this approach has attracted large numbers of clinicians (and scientists), drawn by the technological innovations, high-impact papers and large research funding.

Target identification

Following the identification of disease-relevant phenotypes, the subsequent identification of novel drug targets that modulate or inhibit these responses can be broadly divided into studies at the physiological, mechanistic or genetic level [5]. Whenever possible, target discovery programmes will attempt to use information derived from all these approaches to provide convincing evidence of the potential benefit of developing a drug against a specific protein target.

Physiological approach

The physiological approach attempts to identify novel targets through studies in whole organisms. Because most diseases, including psychiatric, metabolic, cardiovascular, gastrointestinal disorders, are only manifest at the level of the organism, this approach is the preferred and often the only approach directly applicable to target identification.

Prior to the advent of molecular biology, novel drugs (and their targets) were commonly identified through the 'blind' screening of large numbers of either biologically derived or chemically synthesized compounds against the relevant disease phenotypes in cell and animal models [8]. Often, it then took many years to identify the actual drug target and mechanism of action. However, this approach has been highly successful and targets identified through this approach in the period between World War II and the late 1970s still provide the pharmacological basis for the treatment of cancer, infectious, inflammatory, cardiovascular, respiratory and psychiatric disease. Today, for example, corticosteroids, which are now used in the treatment of over 200 diseases, were first identified in a protein fraction isolated from adrenal glands. Similarly, the prototypical nonsteroidal anti-inflammatory drugs (NSAIDs) indomethacin and ibuprofen were discovered by screening many hundreds of compounds derived from serotonin and carboxylic acid in a rabbit model of paw inflammation. In many other cases, new drugs and in particular those used for the treatment of psychiatric disease were identified by serendipity or chance. This includes those used in depression, such as the selective serotonin uptake inhibitors, which were identified during a screening programme for anti-histamine drugs [9], and the monoamine oxidase inhibitors, which were originally licensed for the treatment of tuberculosis.

Throughout the post-war period, the steady improvements in scientific techniques resulted in an increased understanding of the biochemical mechanism underlying physiological responses, which, in turn, permitted a more-targeted approach. An early example was the development of the angiotensin converting enzyme (ACE) inhibitors such as captopril, which are now commonly used in the treatment of hypertension. Unlike the earlier 'random' biological screens, these drugs were developed through selective screening against the renin-angiotensin system, following clinical demonstration of its central role in the control of blood pressure in the early 1970s [10]. Importantly, this trend was accelerated by the advent of molecular biology, which has permitted the identification and manipulation of potential drug targets (see Section 'mechanism-driven approach'), as well as examination of the role of genetics (see Section 'gene-driven approach').

In the case of the physiology-driven approach, the most important impact of molecular biology has been the emergence of transgenic and knockout mouse models. These are now commonly used to identify the role of individual genes in normal and pathological responses and in the development of models for diseases, such as diabetes, atherosclerosis and obesity [11,12]. Unfortunately, a number of factors, such as species and strain differences, as well as embryonic lethality and the induction of compensatory mechanisms during development in transgenic and knockout animals, means that these data are often difficult to interpret and can produce misleading conclusions (see Section 'animal models').

Mechanism-driven approach

The mechanistic approach attempts to identify drug targets by comparing the intracellular pathways that regulate biological responses in primary cells obtained from normal patients with those from diseased patients. Although this has been useful in the elucidation of the basic pathways and mechanisms that regulate cellular responses, the complexity of the interactions and the variation between individuals have made it difficult to pinpoint specific targets involved in disease. This task is compounded by several experimental difficulties and, in particular, by the rapid biochemical and structural changes that occur in cells *ex vivo*. Thus, investigation of the profile of protein expression in the plasma membrane of endothelial cells showed that of the 450 proteins expressed in rat lungs, 41% were found to be absent in the equivalent cultured cells [13]. The limitations of the mechanistic approach to target identification are best exemplified in oncology. In this area hundreds of billions of dollars have been spent elucidating the genetic and molecular mechanisms responsible for cell proliferation, leading to the identification of numerous oncogenes, tumour suppressors and repair genes. However, with the exception of some minor success in the treatment of rare cancers, such as the treatment of chronic myeloid leukaemia using Gleevec™, this enormous

investment has had little impact on overall patient prognosis. Thus, in spite of the development and testing of many hundreds of targeted treatments, the survival rate of patients with solid tumours, such as colon, lung, breast and prostate, which make up 90% of all cancers, has not significantly increased [14,15].

Despite these problems, an understanding of the basic cellular interactions has been useful for identification of novel targets in infectious disease. In particular, knowledge of the life cycle of the human immunodeficiency virus (HIV) has resulted in the development of inhibitors of viral entry, reverse transcription and proteolytic processing of the viral proteins. These drugs, when used in combination, have profoundly reduced AIDS related mortality. In addition, this mechanistic approach has been useful in identifying additional targets in pathways with proven clinical efficacy. Thus, following the discovery of an additional isoform of cyclooxygenase (COX), named COX-2, it quickly became evident that it might be possible to avoid the gastrointestinal side effects associated with the existing NSAIDs using selective COX-2 inhibitors. This eventually led to the development of compounds such as rofecoxib (Vioxx™) and celecoxib (Celebrex™) [16].

Gene-driven approach

The final approach to the identification of potential targets has been to compare the differential expression of genes (genomics) and proteins (proteomics) in normal and disease cells and tissues. In general, as well as being technically challenging, this approach produces large numbers of potential targets and it has proven difficult to identify those that are causally involved in the disease. Indeed, a recent investigation by Miklos *et al*. [17] of different experimental approaches to the identification of genes involved in schizophrenia showed little correlation between data generated by DNA microarrays and other non-microarray-based clinical and biological approaches. This realization has resulted in the increased use of genomics and proteomics approaches for the identification of disease biomarkers rather than novel drug targets.

An alternative approach has been to investigate difference in gene sequences and, in particular, to correlate disease phenotype with single nucleotide polymorphisms. As with genomics studies, these genetic association studies only identify correlations and, in addition, have been difficult to reproduce because they are crucially dependent upon the availability of large, well-characterized patient populations.

Disease models

As previously stated, target identification is also dependent upon the availability of relevant disease models and would therefore ideally be undertaken in human subjects. However, ethical implications, as well as expense and technical difficulties often prevent human-based studies. To overcome this bottleneck, medical scientists have traditionally turned to cell and animal models. Significantly, the advent of molecular biology and the need to manipulate genes has encouraged the widespread use of immortalized cell lines and mouse models. Although these studies are useful for basic mechanistic studies, several issues have emerged that limit their utility as predictive models of human disease.

Cellular models

At present, the functional role of a protein is investigated either through overexpression of the wild-type and/or dominant-negative gene or small interference RNA (siRNA) mediated knockdown [4]. However, because the use of these experimental approaches is dependent upon plasmid and siRNA delivery, this has resulted in the widespread use of easy-to-transfect immortalized cell lines in preference to primary cells. Fortunately, many of the intracellular pathways appear to be evolutionarily conserved and these studies have elucidated important basic mechanistic interactions. Unfortunately, because of the genetic differences between immortalized and primary cells, the difficulty in simulating the *in vivo* milieu and the 'noise' that is inherent in all biological systems, it is virtually impossible to 'model' physiological and pathological responses by extrapolation from cell-based studies [18].

Animal models

Historically, examination of medical textbooks will show that animal studies have provided valuable information on human physiology and pathology. However, as a result of crucial differences in areas such as immunology, metabolism and anatomy, these investigations were often undertaken in a variety of species (e.g. mouse, rat, guinea-pig, rabbit, pig, dog) to identify those that were most analogous to humans. Today, the ease in breeding and maintenance, the availability of the genome sequence and our ability to produce transgenic and knockout animals, means that mice are almost exclusively employed, irrespective of whether they represent the most suitable species. In addition, unlike earlier investigations that employed 'out-bred' and 'wild-type' animals, these mouse colonies are highly 'in-bred' and, therefore, genetically homogenous. Although this often means that smaller numbers of animals are required to obtain statistically significant results, the final results are often highly strain dependent. As an example, a recent publication examining the induction of emphysema in five mouse strains, following chronic exposure to cigarette smoke, showed a wide range of responses [19]. Similarly, contradictory results are often produced when using transgenic and knockout mice of different strains to study the role of individual genes [20]. These species and strain differences have brought into question the utility of mouse models even for simple physiological studies [18,21]. Furthermore, these problems are compounded when developing models of chronic disease, where the long-term environmental factors and the influence of

ageing need to be modelled. In most cases, the combination of time constraints, expense and experimental complexity, means that these important influences are simply ignored. This is particularly true in the case of ageing, where, despite its strong correlation with the development of chronic disease, the majority of studies are still performed in young animals. Interestingly, this problem was highlighted in a recent publication that examined the development of pulmonary emphysema using transgenic mice and found that characteristic phenotypic changes were seen in 6- and 12- but not in 2-month-old mice [22].

Finding new drug targets

Emphasis on human-omics

To address the problems outlined in the previous sections, it is important to understand that disease is manifest at the level of the organism. For this reason, target identification is crucially dependent upon clinical research and the identification of disease relevant phenotypes. Using modern terminology, this might be best named a 'human-omics' approach. Unfortunately, these more-traditional approaches involving the investigation changes in structure (anatomy, histology) and function (physiology) have been generally sidelined by the emergence of molecular sciences. This is particularly evident in academia. Therefore, it is essential that steps are taken to reverse the ongoing fall in basic clinical research. As in the past, this is likely to occur through the development of new techniques for measurement of biological responses and parameters in human patients. Thus, the use of noninvasive techniques, such as the stethoscope, blood pressure monitoring, ECHO and electrocardiogram, revolutionized our understanding of cardiovascular disease and provided the basis for the development and assessment of novel drugs. In the near future, modern imaging techniques, such as magnetic resonance imaging (MRI) and positive emission tomography (PET), might permit exciting new insights into the aetiology of human disease.

When undertaking clinical research, it is also essential to understand that this is an ongoing process since the changing environment (e.g. diet, housing and infectious agents), increased life expectancy and even the impact of new drug regimes are changing the causes and profile of disease. As an example, our reduced exposure to certain pathogens, resulting from immunization and the clean modern environment, has been implicated in the increase in allergic diseases, such as asthma and rhinitis. Similarly, the advent of treatments for high blood pressure and elevated cholesterol has contributed to a >50% decrease in the incidence of stroke and heart disease, meaning that cancer is now the largest killer in the 45–64 age group.

Targeting extracellular pathways

Following the identification of a clinically relevant phenotype, there are two general approaches to the identification of potential drug targets. The first approach involves the direct targeting of the disease phenotype whereas the second involves indirect targeting through modulation of interrelated physiological pathways. Broadly speaking, uncovering targets with direct involvement in the phenotypic response requires mechanistic studies using relevant cell models. Unfortunately, with most complex diseases and, in particular, psychiatric disorders such as depression and schizophrenia, it is not possible to study the phenotype in isolation from the organism. Furthermore, even where suitable cell models are available, as outlined earlier, there are several technical problems associated with the implementation and interpretation of these data. In spite of these difficulties, this approach has been proven useful in the identification of novel targets in 'relatively simple systems', such as infectious disease.

Indirect targeting of the disease phenotype through modulation of physiological systems has historically been a rich source of drug targets. Thus, although β_2-adrenergic and glucocorticoid signalling have not been implicated in the aetiology of asthma, β-agonists and corticosteroids have been proven highly effective in the treatment of the airway contraction and inflammation associated with this disease. Similarly, although the underlying cause of essential hypertension is unknown, this can be effectively treated using ACE inhibitors, Ca^{2+} channel blockers and α_1-adrenoreceptor antagonists. Interestingly, these examples highlight one of the important benefits of this approach – that it is not necessary to understand the underlying cause of the disease phenotype. This is important because, with the exception of infectious and some cancer and genetic disease, this is still unknown. By contrast, this approach is clearly dependent upon a basic knowledge of physiological systems, which in a similar vein to clinical research, has been sidelined by the emergence of molecular sciences. Furthermore, even when this occurs, these investigations often involve the use of transgenic and knockout mouse lines, which are often poor model of human physiology.

In light of the importance of the physiological approach, the remaining question is how might this knowledge guide the identification of novel targets. In general, physiological responses are mediated through extracellular mediators that act upon plasma membrane or intracellular receptors, including G-protein-coupled and cytokine receptors. Importantly, the recent completion of the human genome project has identified a swathe of orphan receptors and extracellular mediators. It is, therefore, the characterization of the physiological role of these new (as well as existing) extracellular signalling pathways that is likely to produce novel drug targets. Interestingly, indirect support for this contention comes from the recent therapeutic successes using recombinant proteins and therapeutic antibodies. These include the use of extracellular mediators, such as erythropoietin and interferon in the treatment of anaemia and multiple sclerosis, as well as the

use of blocking antibodies for tumour necrosis factor α as a therapy for rheumatoid arthritis.

A new philosophical approach

Overall, identification of new drug targets in the 21st century will require a move from the existing reductionist approach that attempts to understand disease at the mechanistic (cellular) and genetic level, towards a dialectic approach that recognizes that disease is manifest at the level of the organism and, therefore, emphasizes the importance of clinical and physiological research [23]. This is essential because the causes of disease, particularly chronic disease, involve complex and changing interactions among genes, environment and ageing, which vary within individual patients and therefore cannot be simply broken down and understood at the molecular level. Finally, the complexity of disease means that biomedical and pharmaceutical research needs not always to be targeted, because many of our best drugs have been discovered by well-organized 'fishing trips'. This situation was aptly summarized by Sidney Farber when he said that 'the history of medicine is replete with examples of cures obtained years, decades, and even centuries before the mechanism of action was understood for these cures – from vaccination, to digitalis, to aspirin'.

References

1 Booth, B. and Zemmel, R. (2004) Prospects for productivity. *Nat. Rev. Drug Discov.* 3, 451–456
2 Miska, D. (2003) Biotech's twentieth birthday blues. *Nat. Rev. Drug Discov.* 2, 231–233
3 Glassman, R.H. and Sun, A.Y. (2004) Biotechnology: identifying advances from the hype. *Nat. Rev. Drug Discov.* 3, 177–183
4 Lindsay, M.A. (2003) Target discovery. *Nat. Rev. Drug Discov.* 2, 831–838
5 Sams-Dodd, F. (2005) Target-based drug discovery: is something wrong? *Drug Discov. Today* 10, 139–147
6 Knowles, J. and Gromo, G. (2003) A guide to drug discovery: Target selection in drug discovery. *Nat. Rev. Drug Discov.* 2, 63–69
7 Rosenberg, L. (1999) Physician-scientists–endangered and essential. *Science* 283, 331–332
8 Le Fanu, J. (1999) *The rise and fall of modern medicine*, Abacus
9 Nixon, J.S. (1982) *Ibuprofen: chronicles of drug discovery*, John Wiley & Sons
10 Smith, C.G. and Vane, J.R. (2003) The discovery of captopril. *FASEB J.* 17, 788–789
11 Tornell, J. and Snaith, M. (2002) Transgenic systems in drug discovery: from target identification to humanized mice. *Drug Discov. Today* 7, 461–470
12 Zambrowicz, B.P. and Sands, A.T. (2003) Knockouts model the 100 best-selling drugs–will they model the next 100? *Nat. Rev. Drug Discov.* 2, 38–51
13 Durr, E. *et al.* (2004) Direct proteomic mapping of the lung microvascular endothelial cell surface *in vivo* and in cell culture. *Nat. Biotechnol.* 22, 985–992
14 Kamb, A. (2005) What's wrong with our cancer models? *Nat. Rev. Drug Discov.* 4, 161–165
15 Gabor Miklos, G.L. (2005) The human cancer genome project - one more misstep in the war on cancer. *Nat. Biotechnol.* 23, 535–537
16 Flower, R.J. (2003) The development of COX2 inhibitors. *Nat. Rev. Drug Discov.* 2, 179–191
17 Miklos, G.L. and Maleszka, R. (2004) Microarray reality checks in the context of a complex disease. *Nat. Biotechnol.* 22, 615–621
18 Horrobin, D.F. (2003) Opinion: Modern biomedical research: an internally self-consistent universe with little contact with medical reality? *Nat. Rev. Drug Discov.* 2, 151–154
19 Guerassimov, A. *et al.* (2004) The development of emphysema in cigarette smoke-exposed mice is strain dependent. *Am. J. Respir. Crit. Care Med.* 170, 974–980
20 Pearson, H. (2002) Surviving a knockout blow. *Nature* 415, 8–9
21 Leaf, C. and Doris, B. (2004) Why we are losing the war on cancer (and how to win it). *Fortune* 149, 200–214
22 Morris, D.G. *et al.* (2003) Loss of integrin alpha(v)beta6-mediated TGF-beta activation causes Mmp12-dependent emphysema. *Nature* 422, 169–173
23 Lewontin, R. (2001) *The triple helix: Gene, organism, and environment*, Harvard University Press