Accelerating drug discovery

Although the evolution of ‘-omics’ methodologies is still in its infancy, both the pharmaceutical industry and patients could benefit from their implementation in the drug development process

Sandra Kraljevic, Peter J. Stambrook & Kresimir Pavelic

Drug development, from initial discovery of a promising target to the final medication, is an expensive, lengthy and incremental process. The ultimate goal is to identify a molecule with the desired effect in the human body and to establish its quality, safety and efficacy for treating patients. The latter requirements ensure that the approved medication improves patients’ quality of life, not only by curing their illness, but also by making sure that the cure does not become the cause of other problems, namely side effects (Snodin, 2002). It also means that this is a particularly costly and prolonged process. At present, bringing a single new drug to market costs around US$800 million, an amount that doubles every five years. According to the US Food and Drug Administration (FDA), it takes, on average, 12 years for an experimental drug to progress from bench to market. Annually, the North American and European pharmaceutical industries invest more than US$20 billion to identify and develop new drugs, about 22% of which is spent on screening assays and toxicity testing (Michelson & Joho, 2000). In addition to costs, administrative hurdles have become problematic, which contributes to the high failure rate of new drug candidates. Of 5,000 compounds that enter pre-clinical testing, only five, on average, are tested in human trials, and only one of these five receives approval for therapeutic use (Fig 1). It is not surprising that, while development costs have increased, the absolute number of newly approved drugs has constantly decreased for several years. These trends—increasing costs for drug development and testing and greater scrutiny of the approval process—create a growing problem both for the drug industry and for patients who are desperately waiting for new drugs to treat their illnesses. It is therefore timely to consider how new technologies, namely functional genomics, proteomics and the related field of toxicogenomics, can help to speed up drug development and make it more efficient.

The current process of identifying a new drug and bringing it to market involves several lengthy steps (Fig 2). It starts with the synthesis of small molecules to target specific proteins or enzymatic activities in living cells. The next step is to identify those compounds that have the best chance of survival in clinical trials. These drug candidates are then subjected to a battery of in vitro tests to investigate potential class- and compound-specific toxicity; it is in these early stages that most candidates fail. Compounds that make it through this stage are then subjected to acute and short-term in vivo toxicology studies. All information gathered in these pre-clinical stages is then used as a guide for subsequent clinical trials in human volunteers and patients. It is on these pre-clinical and clinical tests that new technologies could have the largest impact.

Functional genomics, which includes proteomics and transcriptomics, is an emerging discipline that represents a global and systematic approach to identifying biological pathways and processes in both normal and abnormal physiological
Functional genomics is a strategy to expand biological investigations beyond studying single genes and proteins to a comprehensive analysis of thousands of genes and gene products in a parallel and systematic way. Given that about 30% of the open reading frames in the human genome have as yet unknown biological functions, scientists have begun to shift from using genome mapping and sequencing for determining gene function towards using functional genomic approaches, which have the potential to rapidly narrow the knowledge gap between gene sequence and function, and thus yield new insights into biological systems.

In transcriptomic studies, DNA microarray analyses have already become standard tools to study transcription levels and patterns in cells (Gershon, 2002; Macgregor, 2003). Furthermore, advances in two-dimensional gel electrophoresis and mass spectrometry are providing new insights into the function of specific gene products (Banks et al, 2000; Jungblut et al, 2001; Lefkovits, 2003). Full understanding of the proteome, however, requires more than gene expression levels as many proteins undergo post-translational modifications that dictate intracellular location, stability, activity and ultimately function. Relying exclusively on mRNA levels to measure protein function can therefore be misleading (Choudhary & Grant, 2004), and thus requires additional information about protein levels and modifications as well as signalling pathways and metabolite concentrations and distribution. These large-scale approaches, aided by using bioinformatics to analyse the data, now generate more biological information than previously possible.

The application of functional genomics to drug discovery provides the opportunity to incorporate rational approaches to the process (Fig 2). Combinatorial chemistry—using high-throughput technologies to rapidly synthesize a huge range of new compounds—and computer-assisted drug design, together with information from emerging proteomics methodologies, are now being exploited to identify new drug targets. The expectation is that combinatorial chemistry, along with computer analysis of the 30,000 or so human genes and their protein products, will yield new
information on hitherto unidentified drug targets. Because traditional high-throughput screening of drug candidates is inherently inefficient, virtual screening of libraries of existing compounds should be an excellent method for in silico prediction for active therapeutics (Dolle, 2002; Jorgensen, 2004). Plexxikon, a drug company in Berkeley (CA, USA), is already exploiting this approach by synthesizing new low-molecular-weight ‘scaffold-like’ compounds that interact broadly with many members of a protein family and target their conserved regions. By combining low-affinity biochemical assays and high-throughput X-ray crystallography, the company identifies promising scaffold compounds for lead development. This platform is unique in that it combines high-throughput co-crystallography, parallel biochemical assays, informatics, screening of compound libraries and chemistry, all combined to accelerate the drug discovery process (Fig 3).

Notwithstanding these novel approaches, large-scale methodologies will become an indispensable tool for understanding drug responses and will provide a rational basis for predicting toxicological outcomes. These new tools should therefore reduce the time and costs required for identifying mechanisms of drug action and possible toxic effects, thereby facilitating the speed with which a new potential drug reaches the market. Better understanding the processes by which drug candidates affect the human body and identifying the cellular factors and processes with which these compounds interact will be the key to improved therapeutics. This particular application of functional genomics to toxicology is defined as toxicogenomics. It allows researchers to identify the toxic effects of a given compound at the level of mRNA translation and gather additional valuable information on protein function and modifications as well as metabolic products (Aardema & Lindon, 2002; Boorman et al., 2002; Newton et al., 2002; Robosky et al., 2002).

Microarray-based toxicogenomic experiments to describe changes in gene-expression profiles induced by a toxic compound may help to establish signature markers of toxicity that are characteristic for a given compound. Recent studies have shown that chemicals with similar mechanisms of toxicity induce characteristic gene-expression profiles (Burczynski et al., 2000; Waring et al., 2001). The microarray data may also provide supporting evidence for potential mechanisms of toxicity (Amin et al., 2004; Hamadeh et al., 2002; Newton et al., 2004; Waring et al., 2001). Two related approaches have been used to classify toxicants on the basis of changes in expression profiles. The first focuses on identifying specific genes whose expression is altered by exposure to a toxicant, so that these can be used as a standard for toxicity tests. The second aims to classify chemicals on the basis of their capacity to alter transcriptional profiles similarly to known toxicants. These strategies may eventually lead to targeted, specific toxicity arrays, which could lower experimental costs and provide better mechanistic data. As public gene-expression databases grow, more toxicological markers will be added and will contribute to greater predictive capacity.

There is considerable interest in using gene-expression profiling to define markers both for desired pharmacological activities and for toxic effects. Such markers can be used to characterize drug candidates and select those with optimal properties for further development. Similarly, proteomics offers a comprehensive overview of the cellular protein complement and can provide useful data about alterations in protein expression after exposure to a toxicant (Fountoulakis & Suter, 2002; LoPachin et al., 2003). A toxicant can act on proteins at many levels: by affecting gene expression, it can induce changes in protein levels, and toxicant-induced oxidative stress can cause secondary damage to proteins. Furthermore, toxicants acting directly or indirectly on their protein targets can alter important post-translational modifications or enhance or decrease stability. All these processes individually or collectively can lead to the disruption of normal protein function in a cell (LoPachin et al., 2003).

Toxicogenomics is already moving from being a purely descriptive science towards being a predictive tool (Fig 4). The identification of more genetic, protein and metabolic toxicity markers allows predictive models of toxicity. Furthermore, these can be grouped into one or several experiments to test which markers are modified by exposure to the compound under investigation. Administration of several doses of a toxicant at different intervals then allows for the separation of pharmacological effects from toxic responses. But to achieve a level of predictability and reliability that is acceptable for drug development and testing, it will require identifying more true markers for toxic response and/or induced toxicity. Such a high confidence in marker prediction will be achieved only by comparing data from large reference databases, multiple doses, different treatment periods, post-exposure points and biological models for each condition.
The implementation of toxicogenomics in toxicology and eventual drug development depends on several factors. The first requires further advances in bioinformatics. Analysing and interpreting expression changes in hundreds of genes and modifications of proteins and metabolic pathways is a daunting task, even when dealing with a small number of samples. Biological pathways are highly complex and interconnected, and high-throughput experiments commonly generate many false-positive and false-negative signals. Advances in bioinformatics and new analytical tools, however, are already improving the interpretation of large-scale expression data and contribute to mechanistic and predictive information that is indispensable for drug discovery and development. The second factor concerns the proprietary issues that result from costly large-scale studies on toxic effects performed by pharmaceutical companies. It is important that this information is made freely available for other companies and researchers to enable them to develop new predictive tests and models. The third challenge is the standardization of raw data deposition in data banks (Kramer & Kolaja, 2002). The minimum information content for microarray experiments, for instance, is already a topic for debate (Ball et al, 2002; Brazma et al, 2001). In brief, the application of functional genomics methodologies to toxicology should optimize the prediction of drug responses.

Such a global analysis will lead to a better understanding of biological mechanisms that cause toxic responses. As Castle and colleagues (Castle et al, 2002) argued, these global approaches will provide a better insight into human toxicology than current developments and have the potential to identify a toxicant earlier and faster in drug development.

Further down the development pipeline, toxicogenomics could also help to make clinical trials safer and more efficient by identifying either poor responders or those who are at particular risk of adverse side effects. One of the main functions of clinical research is to assess possible deleterious properties and side effects in humans of the drug under investigation. A central role in how humans react to a drug is played by the drug-metabolizing cytochrome P450 (CYP) enzymes in the liver. Patients with non-functional CYP alleles are at particular risk for adverse side effects, whereas those with additional copies respond poorly or not at all. The variability of CYP genes thus underlies the variable intensity of drug effects, adverse side effects, toxicity and duration of the toxic response for identical drug doses. In addition, many adverse drug effects are not due to single gene modifications but are polygenic in nature, and different combinations of haplotypes may thus exacerbate or attenuate a toxic response. Again, a toxicogenomic approach to identifying deleterious polymorphisms and the use of RNA expression profiles should help to overcome such problems. In this context, pharmacogenetics, the study of inherited variations in drug metabolism and drug response, could be used as a tool in clinical trials, either prospectively or retrospectively. Prospective genotyping may be used to include or exclude poor metabolizers or those at risk of adverse side effects. Retrospective genotyping can help to generate new hypotheses for further testing or explain unexpected events, such as outliers or adverse drug reactions. As the field of pharmacogenomics is relatively new, most experimental results are not yet suitable for regulatory decision-making; however, efforts to standardize methods and assays are already under way.

In addition, advances in toxicogenomics will also benefit patients in predicting the efficiency and side effects of existing drugs. It has been known for some time that different people in a population respond differently to a given drug. Genetic polymorphisms in genes that encode drug-metabolizing enzymes, transporters, receptors and other proteins are abundant and cause these individual differences in drug responses. For instance, specific variations in the gene that encodes thiopurine methyltransferase (TMPT)—the primary enzyme that metabolizes 6-mercaptopurine and a standard therapeutic for childhood leukaemia—may cause a life-threatening toxic reaction. Although these adverse reactions are well documented and understood, a recommendation for genetic testing before therapy has been vigorously opposed for several reasons: the tests are still rather complex and expensive, and their reliability needs to be improved. Also, training and familiarization of oncologists with genetic testing is needed to achieve a consensus on mandatory testing. Another
example of drug specificity is the use of Herceptin® to treat breast cancer, an effective drug for the 25% of patients who have a mutation in the HER2 receptor gene. A diagnostic test for mutations of the gene now helps to identify those patients who will respond positively to treatment with Herceptin.

There are numerous other benefits of using genetic markers, not only as a guide during drug development but also in treatment. Pharmacogenetics, for instance, promises a rapid elucidation of genetic inter-individual differences in drug disposition, thereby providing a stronger basis for optimizing drug therapy to each patient's genetic makeup. This will lead to individualized therapies in which risks are minimized and desired drug effects are maximized. Although it is financially impractical to design a drug specifically targeted to each patient's genetic constitution, it should be possible to target particular haplotypes and to increase a drug's efficacy or decrease its toxicity across a wider patient population (Evans & Johnson, 2001; Goldstein, 2003). This personalized approach would be based on molecular profiling and would thereby maximize benefit for the patient. Given that adverse drug reactions are the fifth leading cause of death in the USA, causing more than 100,000 fatalities each year (Lazarou et al., 1998), the application of pharmacogenomics to identify those at risk before treatment has huge potential for using existing drugs more safely and efficiently.

But we are not there yet. Large-scale approaches using microarray data analysis have come under criticism because of inter-laboratory, and sometimes even intra-laboratory, variability. This is mainly caused by the difficulties in identifying uncontrolled or unknown variables. Tissue heterogeneity and sampling error introduce additional variability to expression profiling. Tissues from individuals of different ethnicities lead to significant polymorphic noise between individuals, unrelated to the direct effect of the toxicant under study. The relative effect of these experimental variables on expression profiling in humans, including tissue source and patient ethnic background, is an important challenge for the design of better diagnostics (Novak et al., 2002).

Ultimately, it will be market forces that decide whether the pharmaceutical industry will start using the large-scale ‘-omics’ approaches

In addition, the pharmaceutical industry is concerned that clinical trials could become even more costly if clinical pre-testing is required to determine who should or should not participate. Identifying non-responders, however, has the potential to reduce the cost of drug development by making clinical trials more focused. It should be emphasized that the pharmaceutical industry is a profit-making industry, and that pharmaceutical companies are intent on reaching as many consumers as possible with an approved drug. Because only about one-third of patients benefit from any given prescription drug, companies have little incentive at present to develop tests that alert the remaining two-thirds of their customers to the fact that they are not benefiting. But we would argue that linking a new drug to a pharmacogenomic trait and implementing new functional genomics methods in drug discovery and drug development would ensure profit, while drug discovery and pre-clinical studies should be affected only minimally, if at all. First, true responders would be identified prospectively and properly dosed, which would also save healthcare money spent on adverse effects. It would also lower the risk of the ultimate and most damaging failure: that a company has to pull a drug from the market when serious side effects become known after approval, which not only creates huge losses in monetary terms but also in consumer trust and credibility, notwithstanding the threat of lawsuits. Second, toxicogenomic-guided pre-clinical studies and subsequent pharmacogenomic-focused clinical trials would shorten the drug development process and significantly lower costs (Fig 2). Despite these advantages, pharmaceutical companies are still hesitant to integrate these methodologies because they fear that their use will engender new regulations for clinical trials (Eisenberg, 2002; Lesko & Atkinson, 2001). Nevertheless, many pharmaceutical companies have joined the Single Nucleotide Polymorphisms Consortium, which will determine the frequency of certain disease-linked single-nucleotide polymorphisms (SNPs) in three major world populations. The aim is to draw a map of disease SNPs to improve the understanding of disease processes and thus facilitate the discovery and development of safer and more effective therapies. GlaxoSmithKline (Uxbridge, UK) has formed a partnership with Affymetrix (Santa Clara, CA, USA) for its GeneChip technology for the development of genechips for HIV to correlate virus variants with the efficacy of antiviral drugs and drug combinations. In addition, GlaxoSmithKline now uses genotyping in 50 clinical trials in the development of 15 compounds worldwide. This clearly shows that the pharmaceutical industry is responsive to the reality of inter-individual variability in its development of new drugs. Ultimately, it will be market forces that decide whether the pharmaceutical industry will start using the large-scale ‘-omics’ approaches. If it leads to cost savings, as we believe it will, pharmaceutical companies will inevitably adopt them.

From the patients’ and regulators’ points of view, does the pharmaceutical industry have an obligation to adopt the new ‘-omics’ methodologies? So far, their use is not required in seeking approval of a new drug, although the FDA is already drafting ‘Guidance for Industry: Pharmacogenomic Data Submissions’. But before forcing companies to adopt these new technologies in pre-clinical research and clinical trials, it would be prudent to pause and take stock. So far, there is not sufficient assurance that these new methodologies and procedures are able to meet the requirements of safety, accuracy and clinical validity.
accuracy and clinical validity. The new techniques are in fact still inadequate to ensure safety and accuracy, because of a lack of uniformity in the use of new technologies between different laboratories, a lack of uniformity of data and a large variability in the interpretation of these data (Eisenberg, 2002). Before they can be implemented in standard drug development and testing, it is important to achieve consensus on, or at least acceptance of, issues such as standardized materials, standards for assay validation and specific regulatory guidelines for the validation of test results.

The evolution of ‘omics’ methodologies is still in its infancy, and it is important that these approaches are further developed and standardized before they are implemented in drug development for the benefit of the patients and the pharmaceutical industry alike. Nevertheless, they are powerful tools for understanding signalling and biochemical pathways and for elucidating the mechanisms in disease and drug disposition. For that reason, they will eventually facilitate the development of new drugs and the better use of existing ones. More importantly in the short term, they will help to make the drug development process faster and more efficient by eliminating flawed drug candidates early on and thus making sure that when drugs fail, they fail ‘cheaply’ and not after a long and expensive process of pre-clinical and clinical testing. This alone would mean a huge improvement in light of ever increasing costs for drug development, decreasing drug approvals and the fact that many diseases, cancers and others, cannot yet be treated efficiently and safely.

REFERENCES

Sandra Kraljevic and Kresimir Pavelic (bottom) are in the Laboratory of Functional Genomics at the Rudjer Boskovic Institute’s Division of Molecular Medicine in Zagreb, Croatia. Peter J. Stambrook is in the Department of Cell Biology, Neurobiology and Anatomy at the University of Cincinnati Medical Center in Ohio, USA.

E-mail: skraljevic@irb.hr
doi:10.1038/sj.embr.7400236

842  EMBO reports  VOL 5 | NO 9 | 2004

©2004 EUROPEAN MOLECULAR BIOLOGY ORGANIZATION