Chemical genomics - a new drug discovery paradigm

Integrating the chemistry of small synthetic molecules with the available arsenal of tools in genetics, biology and pharmacology will create a more powerful and relevant drug discovery process.

Dr Lutz Weber, Morphochem

The discovery of novel medicines is the science of investigating, understanding and exploiting the interaction of defined molecules with complex biological systems in humans or other living organisms. Historically, three major approaches to the discovery and development of new drugs can be roughly distinguished. The first, driven by chemistry and pharmacology, studied the result of exposing specific molecules to whole organisms and was both successful and rapid. A range of well-known drugs - such as aspirin or Valium (diazepam) - were discovered using this “phenomenological” rather than mechanistic method.

Subsequent progress in biochemistry allowed the isolation and characterisation of proteins as drug targets, and introduced a sequential discovery process aimed at finding molecules that bind tightly and specifically to target proteins. Both random screening and rational drug design have been used to find such molecules, with pharmacology being established at the end of this process. Typical drugs that emerged from this approach are, for example, the currently marketed HIV protease inhibitors.

A third approach - genetic analysis - has recently been introduced; this aims to understand the building blocks of life and the foundation for disease. Consequently, current “third phase” drug discovery uses genetics tools to identify targets and then validate them, express the corresponding proteins, develop screening assays, identify binding molecules and move up the discovery chain until, finally, molecules are pharmacologically tested in animals and later in humans. The anticancer drug, Herceptin, provides an initial example of how this pioneering approach can be used to develop a product successfully.

Now, the tools that have been developed by geneticists in the recent past may have the potential to dramatically change the lengthy and sequential drug discovery chain of the future. The purpose of this article is to describe some of these new ideas, and to show some initial examples of what we believe will be a fourth, more rapid, yet more effective paradigm towards the discovery of novel medicines.

The common theme of this novel paradigm is to integrate the chemistry of small synthetic molecules with the available arsenal of tools in genetics, biology and pharmacology to create a more powerful and relevant drug discovery process. This new concept arose in 1994 when Tim Mitchison described “pharmacological genetics” as a discipline that would use small molecules as probes for complex, pharmacological events. Later on, he and Stuart Schreiber started the Institute of Chemistry and Cell Biology (ICCB) to follow up on what Schreiber also called “chemical genetics”. An ambitious goal was set for chemistry - to discover a highly-specific and tightly-binding small molecule for each expressed protein, with a view to using such molecules as tools for functional genomics, cell-cycle regulation, pharmacology and, finally, as drug prototypes. A novel chemistry would be required - directed at the creation of highly diverse, more complex, small synthetic molecules bearing a high potential to bind specifically to protein targets, and rather different from previous, low-diversity combinatorial chemistry efforts.
Having obtained a desired phenotype, the next step was to identify the target protein that mediated this pharmacological effect. In the genetic experiment, this would be achieved by identifying the mutated target. In the chemical pharmacology experiment, the biologically active compound could be utilised through affinity chromatography to extract and purify the corresponding binding protein. An advantage of the latter method is that it results in an identified, validated target protein - and also a possible drug candidate - in just one experiment. In a striking example, a small natural product found in chilli peppers - capsaicin - was used to identify the pain mediating vanilloid receptor 1 (VR1), giving rise to a novel understanding of the biological basis of and possible treatments for pain (2).

**Chemical genetics**

The problem of identifying a specifically binding molecule for each expressed protein puts a new demand on chemistry that - until now - has not been satisfied by current, established methods. Such molecules would have the properties of antibodies in terms of affinity and specificity, allowing the dissection of one particular biological pathway without affecting others. Natural products have been a good source of molecules that exhibit such properties. Comparing the structural features of these natural molecules, it becomes apparent that they provide a much greater structural variability and diversity than traditional synthetic "drug-like" molecules. Therefore, Schreiber put forward the idea of a novel, diversity-orientated chemistry that yields molecules similar to those found in nature (3).
The first examples of efficient synthesis strategies towards “natural-like” chemical scaffolds have now been established (4). Molecules created to dissect or accelerate biological pathways can be used to bring cells into, for example, a conditional knock-out state - a method that traditionally has been the domain of genetic engineering. Chemical genetics methods can also address biological questions that are not amenable to genetic manipulation, by exploring and controlling cellular processes with chemical probes in a more generalised way (5).

**Gene arrays and chemical genomics**

The post-genomic challenge is to discover the molecular mechanisms of disease. Gene arrays have become a common tool for functional genomics research in order to identify a particular disease state that has an altered expression pattern of individual genes, compared with the normal state found in healthy individuals or isolated tissues and cells. With these powerful tools to hand, the next step in what is now called “chemical genomics” is to identify a changed expression profile after the application of a particular compound within the complex biological system.

An initial example of this novel integration of genomics tools with chemistry is provided by the treatment of *Mycobacterium tuberculosis* (Mt) with the known TB drug, Isoniazid (6). Using a glass chip with 4,000 DNA fragments of Mt, a novel cluster of five genes of the Fas-II complex was identified by showing differential expression. Chemical genomics is, therefore, able to provide not only novel validated drug targets but also chemical drug prototypes. Using collections of diverse compounds of known and - maybe later - unknown biological activity, this approach can provide a shortcut from genomics to drug discovery.

**Protein arrays**

Proteomics is believed to be one of the most important technologies of the future - deriving from and making use of genomics sequencing efforts. However, the creation of protein chips is still in its infancy compared with gene chips, although initial examples have already been successfully identified. Proteins immobilised on chip arrays and combined with small molecules may be able to interfere with the binding of soluble proteins to the immobilised layer, and can be used to find specific binding reactions (7). Screening interesting molecules against thousands of proteins provides a number of valuable small molecules that can be used as tools to validate the role of interacting proteins in a cellular *in vitro* - or alternatively in an *in vivo* - environment.

**Animal arrays**

A step beyond gene and protein arrays has become possible through the use of whole organisms that have previously been sequenced. Examples include bacteria and multicellular eukaryotes - such as the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and Zebra fishes. These organisms are small enough to be assembled in arrays such as multi-well plates. Initial efforts have been made to establish a systematic relationship between genes and disease-relevant phenotypes (8). Initial experiments are also underway that use small...
molecules to reverse such genetically provoked phenotypes - once again arriving at validated targets and drug molecule prototypes, both at the same time and in one experiment. Several new biotechnology companies are also pursuing this approach using mice, the next highest organism to be sequenced.

**Site-directed mutagenesis**

The ideas currently being formed from the merging of genomics and chemistry will depend on finding small molecules that can interact specifically with a given protein target molecule. This goal might be difficult to achieve when dealing, for example, with proteins that are rather similar in their sequence, three-dimensional structure or function - as are the kinases. To allow for the fast selection of suitable molecules, Shokat changed the three-dimensional structure of the v-Src kinase through site-directed mutagenesis, whilst still retaining the enzyme's catalytic properties (9). He then showed that an ATP analogue, created through introducing a side chain by chemical means into ATP, resulted in a highly specific substrate that is only converted by the mutated v-Src kinase and not by other members of this large enzyme family. It was therefore possible to study the pharmacological role of v-Src kinase, without affecting other kinases.

**How to create new drugs more efficiently**

The examples described illustrate a new quality of investigating the interaction of small molecules within complex biological systems. How can one use the answers obtained to design better drugs, more rapidly and more efficiently? Typically, the biological responses towards a chemical entity are complex and multi-dimensional, being apparent in the patterns of gene expression, protein binding ranges, phenotype reversals and phenotype inductions. Multi-dimensional optimisation algorithms will be required to translate such biological feedback for the creation of better drugs. In low-dimensional optimisation spaces, traditional structure-activity relationships (SARs) work well. Higher dimensional search spaces, as provided by chemical genomics, need heuristic optimisation procedures, as exemplified by neuronal networks or genetic algorithms (10). Genetic algorithms utilise biological genomics information in combination with a recently introduced DNA-like description of small synthetic molecules to provide an opportunity for optimising small molecules more efficiently using Nature's evolutionary principles.

Ultimately, one could envisage a co-evolutionary drug discovery process that would enable the discovery of novel drug targets by using small molecules, which would then be improved by respective biological feedback-loops into highly potent and selective new drug candidates.

**References**