

# Hitting the target

The route to new therapeutics often ends in costly failure. The secret of success is the rapid and accurate identification of drug targets with true potential, says Caitlin Smith.

For the pharmaceutical industry, the Human Genome Project has proved to be both a blessing and a curse.

Where potential drug targets were once hard to come by, the industry is now awash with them. This has left researchers with the unenviable challenge of sifting through the data in search of the elusive proteins that are instrumental in human disease.

Akin to seeking a needle in a haystack, this herculean task has boosted the importance of rapid screening technologies. Of the roughly 35,000 genes in the human genome, only a few have known functions. So the task of identifying and verifying a positive lead is key to effective drug development.

Drugs fail in the clinic for two basic reasons: they either don't work or they prove to be unsafe. "Both of these are often the direct result of sloppy early target validation," says David Szymkowski, director of biotherapeutics at the biopharmaceutical company Xencor in Monrovia, California.

Validation is a crucial step in the drug-discovery process. Most drugs are inhibitors that block the action of a particular target



David Szymkowski says drug failures can be due to poor target validation.

protein. But the only way to be completely certain that a protein is instrumental in a given disease is to test the idea in humans. Obviously such clinical trials cannot be used for initial drug development, which means that a potential target must undergo a validation process — its role in disease

must be clearly defined before drugs are sought that act against it, or before it is used to screen large numbers of compounds for drug activity.

Deciding to develop a drug against a particular target is a big commitment in terms of time and money. Once a target enters a pharmaceutical company's pipeline, it can take about 12 years to develop a marketable drug. Each new drug that reaches the market represents research and development costs of close to US\$1 billion.

"Reducing failures early in development is far more important than filling a pipeline with poorly chosen late-stage products likely to fail, and fail expensively," says Szymkowski.

## Model interactions

Computer models are a fast, relatively cheap option for initial screening of both targets and potential drugs. These models usually focus on how the two types of candidate structures interact with each other.

De Novo Pharmaceuticals in Cambridge, UK, has a suite of software for the process, covering virtual screens,

## A QUESTION OF FORM

For any protein, it is not guaranteed that all of its isoforms will have the same function, so it is important to work out which forms are valid drug targets. Many proteins are expressed as different isoforms as a result of alternative splicing of the precursor messenger RNAs (mRNAs) and post-translational modifications. Techniques such as gene knockouts effectively remove all isoforms, but manipulation of mRNA with antisense technology or RNA interference (see 'The silent treatment', page 343) cannot distinguish between isoforms that differ in their post-translational modifications.



Transcription factor production at Sangamo.

To overcome this problem, biotechnology company Sangamo BioSciences in Richmond, California, has developed a system that allows the expression of endogenous genes to be altered in cells or whole-animal models. It uses zinc-finger transcription factors in which the DNA-recognition domain is coupled to a functional domain that allows the expression of the target gene to be up- or downregulated. The advantage of this method over gene knockouts or transgenes is that the expression levels of all the isoforms of an endogenous target gene can be specifically manipulated.

This technique has been used by a group including researchers at Sangamo and Frank Giordano at Yale University to investigate the growth of blood vessels by activating the endogenous expression of vascular endothelial growth factor (VEGF). The team found that all isoforms of VEGF are needed to stimulate the growth of blood vessels that do not leak (E. J. Rebar *et al. Nature Med.* **8**, 1427–1432; 2002).

Casey Case of Sangamo is excited by what he sees as the growing awareness of the biological importance of alternative mRNA splicing. "Conventional overexpression methods for target validation miss the biological consequences of this important phenomenon," he says.

Another approach to the problem has been developed by Xerion Pharmaceuticals in Martinsried, Germany. It uses a system that can remove specific proteins by laser inactivation rather than through genetic manipulation. This route to studying protein function is more direct than genetic means, and may circumvent potential artefacts of overexpression; but, unlike genetic manipulations, Xerion's methods are still limited to cells in culture.

C.S.

docking programs and ligand-based design.

If the structure of the target protein's active site is known, the company's SiteExplorer can predict potential drug-binding sites, and can evaluate interactions between these sites and the drug candidates. If the structure is unknown, then its Quasi2 software will produce a virtual protein based on molecular features known to be important in binding in other targets. Drugs can then be designed against the model.

De Novo also offers software to aid the design of chemical probes used in target validation. The SkelGen suite of programs can then use these data to generate new chemical structures optimized for interaction with a target's active site.

The company is collaborating with GeneFormatics of San Diego, California, in a programme focused on inhibitors of the M10 family of matrix metalloproteinases, enzymes that are involved in cancer and inflammatory disorders. GeneFormatics is using proteomics to identify the target enzymes and characterize their active sites, while De Novo is running docking models and virtual screens of small molecules against the proteins.

Software that can model drug-receptor interactions is available from a number of companies including Tripos of St Louis, Missouri; Accelrys in San Diego, California; and Metaphorics in Mission Viejo, California. In addition, some software is available free to researchers

at non-profit organizations, such as AutoDock 3 made by the Scripps Research Institute in La Jolla, California, and GOLD from the Cambridge Crystallographic Data Centre, UK. Molsoft in La Jolla, California, which makes the ICM molecular modelling software, last month released an ICM browser for the Apple Macintosh.

The Accelrys suite of structural homology programs identifies the possible function, fold family and three-dimensional structure of target proteins by comparing them with sequences and structural homologues of known function. Once the protein's structure is determined, functional information can be gleaned using different modules within Accelrys's Insight II program, which supports a number of processes including X-ray crystallography, nuclear-magnetic-resonance studies and protein engineering.

Target Engine from LION Bioscience in Heidelberg, Germany, aids target selection by offering the ability to analyse gene sequence and expression data, find homologous structures, map potential functional features onto protein structure, view related gene annotation and protein pathway information and use text mining to find functional relationships.

In biotherapeutics, proteins themselves are developed as active drugs. One software suite designed to help optimize protein function is Protein Design Automation (PDA) produced by Xencor. "We don't



Purifying proteins for functional assessment at Xencor

screen DNA sequences," says Szymkowski. "More specifically, PDA computationally screens massive numbers of amino-acid changes in a known protein structure." It then derives functional information from the three-dimensional protein structure and designs novel features into the protein to optimize its function.

#### Sense reversal

Another route to target validation hinges on disrupting gene expression to reduce the amount of the corresponding protein, and so identify the physiological role of the target. Examples of this technique include gene knockouts, antisense

## THE SILENT TREATMENT

MARLENA EMMONS

RNA interference (RNAi) is a selective method of silencing protein expression at the post-transcriptional level. Double-stranded RNA specific to the gene to be silenced is introduced into the cell, where it is processed into short single-stranded RNA fragments. Antisense strands, complementary to the fragments, bind to the target RNA and prompt enzymes to disable it. The effect is to destroy all the target messenger RNA, effectively halting production of the protein.

Genetica, a start-up company in Cambridge, Massachusetts, co-founded by Gregory Hannon at Cold Spring Harbor Laboratory, New York, is developing RNAi technology for high-throughput target validation in mammalian cells. Hannon and his colleagues in New York have done much of the recent work in the stable RNAi suppression of gene expression in mammalian cells.

Scientists at Bristol-Myers Squibb, based in Princeton, New Jersey, are trying to create RNAi reagents for the entire proteome, to allow analysis of all expressed genes. But there are still problems with the method, says Pam Carroll, senior research investigator in applied genomics at the company. "Most researchers use synthetic RNA double-stranded oligos that are expensive and there is still variability in response. Nonetheless, it has been an amazing year for RNA interference as



Gregory Hannon has established RNA interference in mammalian cells.

a mammalian target validation technology," she says. For example, only recently has RNAi been successfully used in mammalian model systems.

Another exciting development in RNAi is the use of small interfering RNA (siRNA) compounds. These are "efficient, specific and relatively non-toxic", says Dmitry Samarsky, manager of technology development at functional-genomics firm Sequitur in Natick, Massachusetts. But he notes that getting these compounds into cells is still a challenge. For researchers interested in a convenient approach to RNAi, companies such as OligoEngine in Seattle, Washington, and Ambion in Austin, Texas, offer siRNA kits and expression vectors, as well as custom siRNA synthesis. Custom siRNAs are also now supplied by many RNA companies such as Dharmacon in Lafayette, Colorado; Proligo in Hamburg, Germany; and QIAGEN in Venlo, the Netherlands.

The biotech firm Benitec, in St Lucia, Australia, has circumvented the difficulties of introducing double-stranded RNA into cells by developing a DNA template that produces double-stranded hairpin-loop molecules within the cell, which trigger RNAi. Last September, Benitec announced that it had made the first transgenic mice in which targeted endogenous genes were suppressed using RNAi.

C.S.

technology and RNA interference (RNAi).

In the realm of drug discovery, antisense technology — the use of short oligonucleotides to target specific messenger RNAs for destruction — was developed as a way of finding oligonucleotide-based drugs that interfere with gene expression, rather than with protein function. But the technology is currently enjoying greater success as a high-throughput method of target validation because it offers a highly specific and efficient way to inhibit the expression of potential target proteins *in vitro* and *in vivo*.

GeneTrove, the genomics division of Isis Pharmaceuticals in Carlsbad, California, is one of the companies active in this field. It is focusing on the untapped pool of potential therapeutic target RNAs for both target validation and drug discovery, says Nicholas Dean, GeneTrove's vice-president of functional genomics. It offers custom target-validation packages that include optimized antisense inhibitors against any target of interest and control oligonucleotides for testing in cell-culture model systems. It also applies antisense technology to target validation *in vivo* in animal models.

Biognostik, a biotechnology company in Göttingen, Germany, offers a drug-target validation kit that can be used *in vitro* or *in vivo*. It includes five target-specific phosphorothioate antisense inhibitors and two random-sequence oligonucleotides to control for nonspecific



GeneTrove analyses the effects of antisense on gene expression.

effects. It has also developed a sequence-design system called RADAR, which determines antisense oligonucleotides based on specificity, minimal nonspecific effects or protein binding, and the ability to be taken up into cells.

Sequitur, a functional-genomics company in Natick, Massachusetts, has a slightly different approach to rapid target validation. It combines an antisense library with high-throughput DNA microarray assays to test the effects of the antisense molecules on gene expression. The company's technology was used recently to validate a major therapeutic target for Alzheimer's disease. Sequitur also carries out target validation based on RNAi (see 'The silent treatment', page 343).

Custom phosphorothioate antisense oligonucleotides for research are available

from firms such as Sigma-Genosys at the Woodlands, Texas; atugen in Berlin, Germany; and Integrated DNA Technologies in Coralville, Iowa. Gene Tools in Philomath, Oregon, offers morpholino antisense oligonucleotides, and Danish companies Cureon in Copenhagen and Exiqon in Vedbæk offer modified oligonucleotides based on 'locked nucleic acid' technology that can be used for antisense.

### The proteomics approach

One disadvantage of doing target validation at the genetic level is that many genes produce several different protein isoforms, which can have subtly different functions (see 'A question of form', page 341). Post-translational modifications can also give protein variations. As a result, a developing approach in target validation is to focus on manipulating the activity of the potential target protein itself. "As the vast majority of drugs target proteins, validating targets is best done by modulating protein activity, not expression levels," Szymkowski says.

Proteomics — the study and manipulation of the protein make-up of a cell — is making it easier to distinguish and target just one specific form of a protein. This allows researchers to avoid unwanted changes in the expression of other proteins — another potential drawback of genetic manipulations.

Stefan Henning, director of functional biology at Xerion Pharmaceuticals in

## A WHOLE PICTURE

Beyond Genomics in Waltham, Massachusetts, is taking a broad view for target validation — its approach hinges on systems biology. The company first characterizes potential targets within their normal biochemical pathways. "We measure cells, tissues and body fluids at multiple bioanalytical levels, including transcript, protein, and endogenous metabolite and enzyme product," says Aram Adourian, the company's senior director of advanced technologies and strategic development.

It then compares these profiles with similar profiles from disease states, and correlates features of the profile with the disease in question using pattern-recognition algorithms and bioinformatics techniques. The result is a set of potential targets, which can then be perturbed by RNAi or transgenics, for example. The subsequent biological changes are monitored by the same systems-biology approach.

As a target-validation tool, systems biology is "a truly exciting method for illuminating causative relations between target modulation and effects on disease", says Adourian. By learning more about the biological pathways in which the target is involved, it is possible to predict the potential toxicity of perturbing its normal function, he says.

Adourian believes that an integrative approach to target validation is more likely to yield the highest-quality targets: "One ultimately needs to examine the system as a whole, with its interacting networks and components of genes, proteins, metabolites and other molecules, to begin to assemble a unified and contextualized perspective on the role, function and relevance of a potential target."

Another firm taking a wide view of target validation is MDS Proteomics in Toronto, Canada, which uses bioinformatics, gene-expression information, high-throughput protein analysis and protein-pathway biology to identify prospective targets and gain all-round evidence of their cellular role. **C.S.**



Combining many techniques is key to systems biology.

Martinsried, Germany, agrees that validation at the proteomics level is a powerful approach. “On a technical level, the development of protein microarrays, multidimensional liquid-based protein separation and technologies that manipulate protein expression and protein–protein interactions will have their impact,” he says.

Xencor has developed ProCode, which enables researchers to study the functions of a cell’s protein make-up. A ProCode library is a protein-expression library from any cell or tissue of interest, in which every protein (after translation) is tagged with a plasmid, a small circular piece of DNA containing its corresponding complementary DNA (cDNA). The library can be expressed in cultures of the appropriate mammalian cells so that proper protein folding and processing are retained. The expressed proteins can be screened for their interactions with potential drugs, and the cDNA tags allow easy identification of any protein that gives a positive reaction.

Xerion’s XCALibur carries out simultaneous identification and functional validation of potential drug targets. Using target-specific antibodies to identify the proteins and chromophore-assisted laser inactivation (CALI) to ‘switch off’ target proteins by photochemically modifying their functional sites, XCALibur can validate specific targets for particular diseases or find new potential targets with disease-associated functions.

XCALibur is incorporated into the Xstream platform, which takes a disease-based approach to target validation. It searches for hits from a suite of antibodies specifically created against the proteome of a diseased cell. The antibodies bind near to functional sites of proteins and contain dyes that are released by CALI, thereby inactivating the proteins’ functional sites. If this inactivation has an effect on the disease, the protein is precipitated by the attached antibody and analysed by mass spectroscopy and database searches.

### Validation *in vivo*

One of the most important tests for a potential drug is an assessment of its role in disease in an animal model. But animal models for certain diseases, such as psychiatric illnesses, are extremely difficult to develop.

“The greatest challenge in target validation is the procurement or development of the correct animal models for the human disease in question,” says Bob Gordon, De Novo’s vice-president of biology. “For example, there are few, if any, reliable animal models for stroke. So validation is effectively done in phase III trials in the clinic. Progress in this disease



Xerion’s XCALIBUR switches off target proteins by laser.

area is understandably slow and expensive.”

*In vivo* target validation using gene knockouts, in which genes are deleted or disrupted to halt their expression, is a powerful method of predicting drug action. “Many of the targets for the top-selling drugs of the biopharmaceutical industry have been knocked out,” says Arthur Sands, president and chief executive of Lexicon Genetics in the Woodlands, Texas.

This kind of target validation is based on the assumption that knocking out the gene for the potential target has the same effect as administering a highly specific inhibitor of the target protein *in vivo*.

“With the effective use of mouse knockout technology, expensive drug-discovery activities can be focused on the drug targets that are most likely to lead to breakthrough therapeutics,” says Sands. Furthermore, target-specific side effects can be discovered before time and money are invested in drug design.

But mammals are not the only creatures in use — zebrafish have recently entered the fray as a model animal for some human diseases. The fish are more affordable, easier to keep, and faster to raise than mammals, giving a higher-throughput system. Drugs can also be tested for toxicity and their potential therapeutic activity against the target more easily than in mammals.

Perhaps surprisingly, genes that cause disease in zebrafish are similar to those in humans, for example in angiogenesis, inflammation and insulin regulation. The transparency of zebrafish embryos also makes them suitable for large-scale, high-throughput genetic and drug screens.

Zygon in Atlanta, Georgia, has developed a transgenic zebrafish system called Z-Tag which can be used for target validation. The company can also make

various zebrafish organs visible by tagging the tissues with fluorescent markers.

One of the most widely used models of human disease is the mouse, but working with mice can be both time-consuming and expensive. Lexicon Genetics has met this challenge by industrializing the generation of mouse knockouts, using gene targeting, gene trapping and mouse embryonic-stem-cell technologies. The result is the company’s Genome 5000 programme, which aims to analyse 5,000 genes over the next five years — over 750 have already been done. Custom transgenic and knockout mice are also available from Deltagen in Redwood City, California, and memorec stoffel in Cologne, Germany.

Researchers are slowly but surely making progress in validating the targets revealed by the Human Genome Project. But proving that a target protein has a causative role in human disease remains a real challenge. “The most exciting technologies are, and will be, those that address the issue of elucidating causative roles of targets in human disease, as opposed to simple associations,” says Aram Adourian, a senior director at the biopharmaceutical firm Beyond Genomics in Waltham, Massachusetts (see ‘A whole picture’, page 345).

More challenging tasks lie in discovering the effects of interactions between newly validated targets in both healthy and diseased models. Such complex information will require not only information systems to correlate multiple variables and outcomes, but also a sophisticated knowledge of protein–protein interactions under a variety of conditions. But for now, researchers are taking it one target at a time. ■

Caitlin Smith is a science writer based in Portland, Oregon.