Functional proteomics in drug discovery

The critical bottleneck of the drug discovery process is developing the right drug for the right target. Today proteins represent more than 90% of druggable targets and will most likely remain the single most important class of molecules targeted by pharmacological agents as more novel targets are identified. However, the current preponderance of potential drug targets and the hype about target validation has led to confusion over how to determine the optimal molecular target for pharmacological intervention. Functional proteomics represents a powerful approach of providing valuable information on target druggability and simultaneous rapid access to relevant therapeutic leads, creating a new paradigm that will accelerate downstream drug development.

The shift from chemistry-driven to biology-driven drug discovery was triggered by the realisation that a better understanding of the genetic basis of diseases would ultimately lead to more efficient drugs. It was assumed that the sequencing of the human genome would eventually result in a wealth of new drug targets and subsequently also lead to more efficient and safer drugs. Genetic methods have indeed been used extensively for the identification and characterisation of potential novel drug targets, but it soon became apparent that the information provided by the genome is not sufficient to assess the suitability of a target for pharmacological intervention.

As the main executors of biological function, proteins have been estimated to be magnitudes greater in number, dynamics and complexity than genes. They thus generate the basis for the functional diversity in living systems and at the same time provide the classical molecular target for drug intervention. Indeed, more than 95% of the 500 drugs in the market today target proteins[1], whereas many of the small number of drugs that target DNA are toxic due to lack of specificity[2]. For RNA-based drugs, so far there is only one anti-sense drug in the market targeting cytomegalovirus infections of the eye[3].

In a recent study, the total cost and time of developing a drug has been estimated to about $880 million and 14.7 years, respectively[4]. The early phase of this process, including target identification and validation, has the highest attrition rates and represents almost half of the total costs to bring a drug to market. Further decreasing time and increasing the predictability of this early phase will drastically reduce costs and improve the success rate of the overall drug development process. This article will discuss the impact of functional proteomics in optimising this early phase and will focus on emerging functional proteomics technologies.

Defining functional proteomics

The urgent need to assign function to novel protein sequences coupled with the recent hype about proteomics has led to the frequent misuse of the term 'functional proteomics'. Blackstock and Weir suggested subdividing the proteomics domain into different categories that are typically associated with a particular set of technologies used[5]:
1. Expression proteomics is mainly concerned with

By Dr Stefan W. Henning and Dr Leodevico L. Ilag

Proteomics

References
1 Drews, J (2000). Drug Discovery: A Historical Perspective. Science 287, 1960-1964.
2 Crul, M, van Waardenburg, RC, Beijnen, JH, Schellens, JH (2002). DNA-based drug interactions of cisplatin. Cancer Treat Rev 28(6), 291-303.
3 Jabs, DA and Griffith, AM (2002). Formivirsen for the treatment of cytomegalovirus retinitis. Am J Ophthalmol 133(4), 552-556.
4 Flanagan, A, Guy, P, Lubkeman, M, Steiner, M, Reeves, M and Tollman, P (2002). A Revolution in R&D: How genomics and genetics are transforming the biopharmaceutical industry. Boston Consulting Group Report 26 November 02, pp.1-64. www.bcg.com/publications/files/eng_genomicsgenetics_rep_11_01.pdf
5 Blackstock, WP and Weir, MP (1999). Proteomics: quantitative and physical mapping of cellular proteins. Trends in Biotechnol 17, 121-127.
6 Jeffery, JC (1999). Moonlighting proteins. Trends in Biochem Sci 24, 8-11.
7 Henning, SW, Ilag, LL and Rudert, F (2001). Functional Proteomics. Inno Pharm Tech 1(16), 62-70.
8 Ilag, LL, Ng, JH, Beste, G and Henning, SW (2002). Emerging high-throughput target validation technologies. Drug Disc Today 7(18) Suppl., S136-S142.
9 Collins, FS and McKusick, VA (2001). Implications of the human genome project for medical science. J Am Med Assoc 285, 540-544.
10 Schreiber, SL (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. Science 287, 1964-1969.
11 Haggarty, SJ, Mayer, TU, Miyamoto, DT, Fathi, R, King, RW, Mitchison, TJ and Schreiber, SL (2000). Dissecting cellular processes using small molecules: Identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. Chem Biol 7, 275-286.

the analysis of changes in protein expression levels in different cellular situations (eg normal vs diseased) or subcellular localisations by the use of 2D-PAGE, mass spectrometry and protein arrays. 2. Cell map or interaction proteomics involves the determination of protein-protein interactions and the description of protein networks, using yeast 2-hybrid techniques, interaction mapping and affinity capture combined with mass spectrometry as typical tools. 3. Structural proteomics uses NMR, x-ray crystallography as well as homology and molecular modelling to determine or predict the three-dimensional structure of proteins.

However, these approaches only provide circumstantial evidence or hints of association of a protein to a physiological state or biological pathway. Only technologies that are typically associated with functional proteomics such as chemical genetics, analog-sensitive enzyme alleles, the use of blocking antibodies or aptamers and other loss-of-function approaches like chromophore-assisted laser inactivation (CALI) provide a direct link to the protein's role in the physiology of the cell or organism. Protein function cannot be determined by examination of the individual components of a cell but most be assessed in the context of individual cell type, localisation and interaction partners[6]. Thus, functional proteomics takes advantage of the systematic perturbation or functional inactivation of a protein within its physiological environment to address the potential role of the target protein in a cellular process, akin to gene disruption to elucidate the function of a gene[7]. The parallel evaluation of protein function at the cellular and molecular level allows efficient decision-making for target validation and lead discovery. Selecting the right therapeutic strategy at an early stage of the drug discovery process will significantly save time, reduce costs and improve the success rate of drug development.

The right drug for the right target
The key to the development of a therapeutic strategy is to understand how the pathway from gene to phenotype is modulated at the molecular level and how this modulation results in particular phenotypic consequences[8]. A gene can be transcribed or differentially spliced into several mRNAs that are subsequently translated into proteins. In addition, proteins may in turn be post-translationally modified, leading to multiple forms that may have distinct functional roles. This explains how functional complexity can be created from a relatively few number of coding elements. Different stages in the pathway from DNA to RNA to protein can be

basis for the development of therapeutics. However, a particular therapeutic approach requires an accompanying target validation strategy that should introduce a perturbation at the same level as the final drug. In other words, the type of target molecule (DNA, RNA or protein) eventually to be modified by the drug dictates the appropriate technology for determining the physiological role of the target. For example, if a mRNA is prematurely degraded or blocked by antisense or RNAi approaches, the protein(s) encoded will not be produced. This will lead to subsequent inhibition of multiple biological functions and result in a certain phenotype. Nevertheless, if in this case the resulting cellular phenotype is the desired effect of the drug, an antisense drug would be appropriate. However, if the desired effect of a novel drug is the molecular modulation or inactivation of a protein, it will be necessary to validate the target's function at the protein level and not at the mRNA level. With complex targets such as proteins that have multiple structural and functional domains, targeting the appropriate molecule down to the specific functional epitope will substantially improve a drug lead's success rate.

Forward and reverse approaches
Genetics has been the classical approach in determining gene function and has been used to extrapolate protein function and the understanding of the molecular basis of disease[9]. However, genetics may be limited because they do not provide relevant pharmacological information on the target. The modifications at the gene level in most instances do not reflect how a drug will work by direct modulation of the protein product. Nevertheless, the principles of genetics can be adapted to protein-based methods for modulating protein function and can also be applied in forward or reverse approaches.

Similar to the classical forward genetic approach, the starting point of a forward 'proteinetic' approach is the identification or induction of a biological phenotype of interest and the subsequent deconvolution of the proteins that were modulated leading to the change in phenotype. One advantage of this method is that no prior knowledge of a target is needed – it is only driven by the selection of the appropriate disease-relevant biological assays and the 'hits' screened as indicated by change in response. Furthermore, mutagenesis methods are not needed and maintenance of mutant lines that can be unstable is not required.

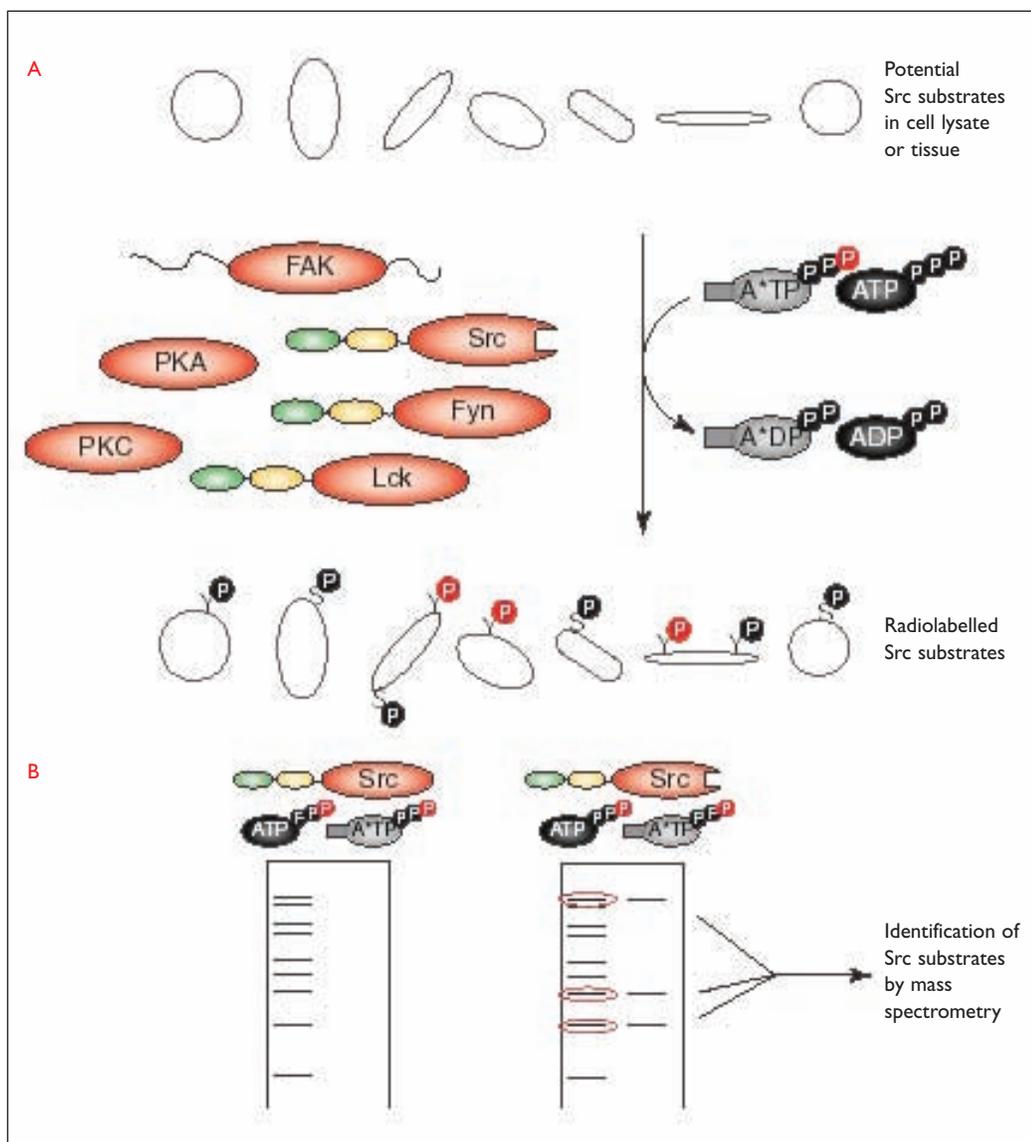In analogy to reverse genetics, whereby candi-

date genes are directly mutated and the biological consequence scored, in a reverse 'proteinetics' approach, molecules can be selected for binding to pre-characterised targets hypothesised to play a role in a particular disease of interest. These binders are then tested with the appropriate biological assays to determine whether the binder can modulate the protein target. Such an approach is useful towards validating protein targets identified from first generation genomics and proteomics technologies such as differential display, protein-protein interactions and *in silico* predictions. These methods have already provided a deluge of candidate proteins – using reverse approaches each one of these putative target proteins can be validated through determination of its causative role in a disease phenotype.

Chemical genetics

Instead of genetic mutations, the concept of chemical genetics is based on the use of small molecules to modulate protein function rapidly and conditionally and thus elucidate their biological function[10]. Chemical genetics applied in a forward approach involves phenotypic screens of structurally diverse and complex small molecule libraries that can provide individual probes of biological function as well as information about interaction properties of small molecules with biological systems. Phenotypic screens have been used in the past to identify active compounds but have regained popularity with the emergence of chemical genetics. The difference today is that proteomics technologies are being used to determine the actual molecular target of the active

Proteomics

12 Lenz, GR, Nash, HM and Jindal, S (2000). Chemical ligands, genomics and drug discovery. Drug Disc Today 5, 145-156.

13 Schreiber, SL, Nicolaou, KC and Davies, K (2002). Diversity-oriented organic synthesis and proteomics. New frontiers for chemistry and biology. Chem Biol 9, 1-2.

14 Shimizu, N, Sugimoto, K, Tang, J, Nishi, T, Sato, I, Hiramoto, M, Aizawa, S, Hatakeyama, M, Ohba, R, Hatori, H, Yoshikawa, T et al (2000). High performance affinity beads for identifying drug receptors. Nat Biotechnol 18, 877-881.

15 Shokat, K and Velleca, M (2002). Novel chemical genetic approaches to the discovery of signal transduction inhibitors. Drug Disc Today 7, 872-879.

16 Hwang, YW and Miller, DL (1987). A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein. J Biol Chem 262, 13081-13085.

17 Bishop, AC, Ubersax, JA, Petsch, DT, Matheos, DP, Gray, NS, Blethrow, J, Shimizu, E, Tsien, JZ, Schultz, PG, Rose, MD, Wood, JL et al (2000). A chemical switch for inhibitor-sensitive alleles of any protein kinase. Nature 407, 395-401.

18 Kraybill, BC, Elkin, LL, Blethrow, JD, Morgan, DO and Shokat, KM (2002). Inhibitor scaffolds as new allele specific kinase substrates. J Am Chem Soc 124, 12118-12128.

19 Lin, Q, Jiang, F, Schultz, PG and Gray, NS (2001). Design of allele-specific protein methyltransferase inhibitors. J Am Chem Soc 123, 11608-11613.

20 Holt, JR, Gillespie, SKH, Provance, DW Jr, Shah, K, Shokat, KM, Corey, DP, Mercer, JA and Gillespie, PG (2002). A chemical-genetic strategy implicates myosin-1c in adaptation by hair cells. Cell 108, 371-381.

21 Jay. DG (1988). Selective destruction of protein function by chromophore assisted laser inactivation. Proc Natl Acad Sci USA 85, 5454-5458.

compounds. The critical components in this approach are the relevance and robustness of the phenotypic screens and the diversity of the probe molecules. Forward chemical genetics has led to the identification of several interesting molecules in phenotypic screens of small model organisms and cell-based assays[11]. In addition, chemical genetics has been successfully used in a reverse approach by direct screening of small molecules that bind to a putative target of interest. These molecules are then tested for biological activity in appropriate cell-based assay systems. The major selling point of both forward and reverse approaches is that once the target has been validated it comes with a potential small molecule lead compound[12].

However, one of the current limitations of chemical genetics is the low diversity of small molecule libraries compared to biological libraries of antibodies or aptamers. A diversity-oriented organic synthesis (DOS) approach to specifically address this limitation and make new chemical libraries more suitable for forward chemical genetic screens has recently been described[13]. Infinity Pharmaceuticals is one of the pioneers which has made significant progress in this matter through a facile approach using high-capacity macrobeads as single microreactors. Another limitation of the forward chemical genetics approach is the difficulty in determining the target of the active molecule, which provides a hit from the phenotypic screen. It is not only technically rather challenging to use small molecules for affinity purification methods, but they also have a tendency to bind to multiple proteins making it difficult to draw conclusions as to their molecular specificity. However, new methods using small molecules to isolate their corresponding targets are currently being developed[14].

Analog-sensitive enzyme alleles (ASEA)

An emerging powerful method of rational chemical genetics that is primarily applicable for reverse approaches is the use of orthogonal chemical genetics and more specifically of analog-sensitive enzyme alleles[15]. The concept is based on the premise that a modification of a small molecule ligand accompanied by the complementary modification of the target protein will facilitate a specific interaction between the ligand and receptor. The modified small molecule cannot interact with its natural protein target (it is 'orthogonal') and preferably any other target protein in the cell, whereas the target receptor protein is engineered to accept the newly modified compound but retains its normal physiological functions. Hwang and

Miller[16] pioneered the concept of ASEA with their studies on the GTP-depended elongation factor Tu.

Protein kinases represent about 2% of all human genes and are in involved in the regulation of virtually all cellular processes and hence many disease mechanisms. Kinases are particularly amenable for ASEA approaches since they do not require individual chemistries and structural engineering for each protein of interest. A panel of analog-sensitive kinase alleles (ASKA) was developed by creating a subtle but unique structural distinction in the catalytic domain of one kinase that distinguishes it from all other kinases (**Figure 1**). The additional creation of a wide variety of inhibitor scaffolds for kinase alleles leads to an approach that can systematically probe the physiological roles of the more than 500 individual members within the kinase superfamily[17]. Indeed, a new class of triphosphate substrate analogs for highly divergent analog-sensitive kinase alleles has been developed and is expected to be applicable across the entire kinase superfamily[18]. Furthermore, the concept of ASEA is not limited to kinases but has also been applied to other enzyme families such as methyltransferases[19].

Another convincing demonstration of the potential of the ASEA approach is its application in defining the specific role of myosin-1c among all 40 different myosins in the adaptation of sensory hair cells[20]. An exciting further development of ASEA would be the introduction of mutant alleles into an animal model to determine the *in vivo* role of the target protein upon addition of the small molecule analog. The combination of ASEA with different approaches such as gene and protein expression profiling will lead to a better understanding of how a drug modulates its target and the evaluation of its impact on the whole physiology of the organism, eventually resulting in a better appreciation of drug toxicity. However, only a limited number of functional domains with catalytic activity might be amenable for manipulation by ASEA and other functional domains might simply not be accessible to small molecule analogs.

Chromophore-assisted laser inactivation (CALI)

Although ASEA-based approaches as exploited by Cellular Genomics, Inc will undoubtedly have a substantial impact on drug discovery, today small molecule modulators are available only for a few protein families that all relate to existing drug targets. Antibodies provide an efficient alternative way to acquire specific protein binders, but the success rate of obtaining neutralising antibodies that block protein function is unsatisfactory low.
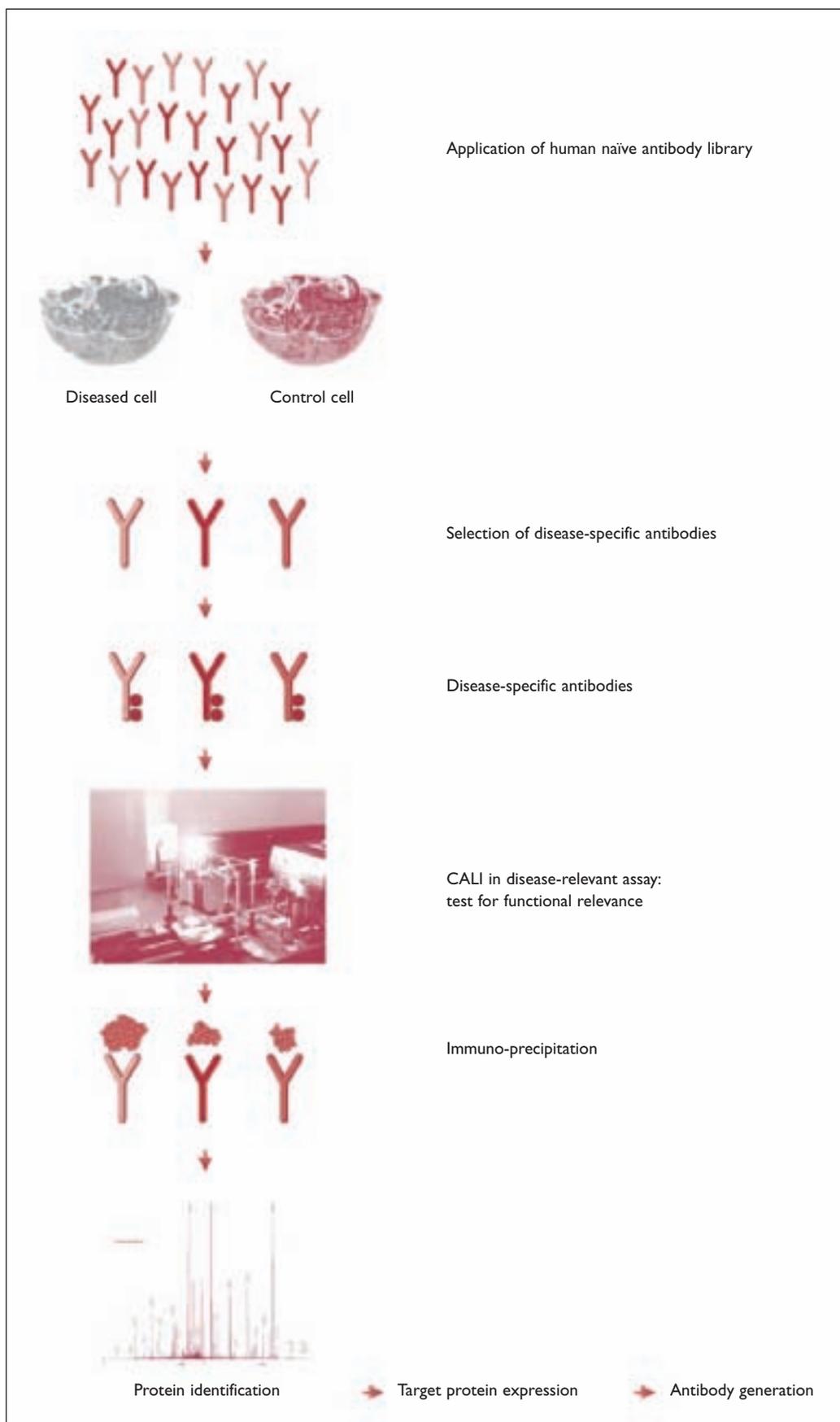
Application of human naïve antibody library

Diseased cell   Control cell

Selection of disease-specific antibodies

Disease-specific antibodies

CALI in disease-relevant assay:
test for functional relevance

Immuno-precipitation

Protein identification   ➤   Target protein expression   ➤   Antibody generation

Figure 2
Simultaneous identification and functional validation of novel disease relevant drug targets via Xstream. In this approach antibodies are selected from human naïve single chain antibody phage display libraries against surface proteins of cells representing a particular phenotype, eg a diseased state. The antibodies are then systematically screened for their ability to mediate the inhibition of a disease relevant biological function via chromophore-assisted laser inactivation (CALI), thus converting the diseased phenotype back to normal. Once a functional hit is generated, indicating that the corresponding antibody has enabled the inactivation of a target protein essential for the diseased phenotype, the antibody is used to immunoprecipitate its cognate target antigen. After separation of the immuncomplexes by one-dimensional gel electrophoresis the isolated protein is identified via mass spectrometric analysis. The identified protein represents a functionally validated target in the disease indication for which the cell type and the functional assay used is pertinent

Proteomics

Continued from page 20

22 Ilag, LL, Ng, JH and Jay, DG (2000). Chromophore-assisted laser inactivation (CALI) to validate drug targets and pharmacogenomic markers. Drug Dev Res 49, 65-73.
23 Beck, S, Sakurai, T, Beste, G, Eustace, B, Schier, R, Rudert, F and Jay, DG (2002). Fluorophore-assisted light inactivation (FALI): A high-throughput tool for direct validation of proteins. Proteomics 2, 247-255.
24 Henning, SW and Beste, G (2002). Loss-of-function strategies in drug target validation. Current Drug Disc 2, 17-21.

Thus, a technology that could add a neutralising activity to the high specificity and affinity of modern recombinant antibodies would make a valuable addition to the tool kit of functional proteomics.

Chromophore-assisted laser inactivation or CALI is such a technology that takes advantage of highly specific recombinant antibodies and converts these antibodies with high probability into a blocking agent[21]. To perform CALI, an antibody that binds to the target protein of interest is labelled with a chromophore of particular characteristics, which make it suitable for the CALI process. When irradiated with light of an appropriate wavelength, the dye moieties are excited and generate highly reactive oxygen species. These further modify amino acid residues of the target protein in the vicinity of the binding site of the dye-labelled antibody. These modifications typically result in the functional inactivation of the target, which can be monitored in a relevant biological assay system. CALI has been developed as a method to determine the *in situ* function of proteins in biological processes and has been extensively tested against a diverse array of proteins in cellular systems as well as small animal model organisms[22].

Recent efforts to define the function of drug target candidates derived from genomics and proteomics approaches have further increased the potential of CALI by the use of phage displayed recombinant single-chain Fv (scFv) antibody fragment libraries as an unlimited source of target specific ligands[23]. Furthermore, the combination of CALI with antibody phage display and mass spectrometry has generated a powerful forward proteomics screen (Xstream[TM]) (**Figure 2**) to identify novel drug targets validated on the basis of their functional involvement in a particular disease relevant biological function[24]. ScFv antibodies specific for a cellular phenotype of interest (eg diseased or malignant state) are systematically screened with CALI for functional hits that convert the diseased phenotype to normal and are further used to isolate and identify the cognate target by immunoprecipitation and mass spectrometry, respectively. Because of their higher specificity and affinity as well as established immunoprecipitation techniques antibodies have proven to be much more advantageous in this type of forward proteomics screens compared to small molecules.

Conclusions

For the pharmaceutical industry the choice of the right drug for the right target has become a critical success factor for filling their dwindling drug pipeline. Genomic and proteomic technologies have generated a deluge of novel potential drug targets with little or no functional annotation or link to a disease relevant phenotype. Since proteins will remain the most important class of molecules targeted by pharmacological agents, new technologies that unravel the complex role of proteins in normal and diseased phenotypes will play an increasingly important part in the comprehensive characterisation and prioritisation of proteins as drug targets. In addition, advances in the design of phenotypic screens predictive of human disease in combination with forward proteomics approaches will lead to the discovery of novel proteins that are accessible to therapeutic intervention.

The loosely defined term 'target validation' is in fact a continuous process and in the extreme case a target might only be 'validated' once a drug has been proven to be efficacious and safe after several years of clinical trials. Such retrospective definition will change as soon as one target validation technology has demonstrated high predictive value. However, it is clear that a better understanding of the physiological function of a target is required and the crucial link between type of therapeutic molecule and the appropriate target validation strategy needs to be better appreciated. Direct protein modulation strategies will continue to evolve and may even outperform the impact of genetics on the drug the discovery process.   DDW

Dr Stefan W. Henning is Director of Functional Biology at Xerion Pharmaceuticals AG, Martinsried, Germany. He previously held post-doctoral positions at the Howard Hughes Medical Institute at the University of California, Berkeley, USA, the Imperial Cancer Research Fund, London, UK and at the MRC Laboratory for Molecular and Cell Biology London, studying signal transduction events in T-lymphocytes. He was a graduate student at the German Cancer Research Centre in Heidelberg and received his PhD degree in Biochemistry and Immunology from the University of Tübingen in 1993.

Leodevico L. Ilag is Chief Scientific Officer of Xerion Pharmaceuticals AG. Prior to his current position and co-founding Xerion Pharmaceuticals AG, he was a Senior Scientist, Group Leader and Member of the R&D Directorate at Morphosys AG. He was a postdoctoral fellow at Purdue University, where he used structural methods to study virus assembly and disassembly processes. He obtained a PhD from the University of Tennessee Medical School at Memphis for work on bacteriophage ɸX174.