DOI: 10.1002/cbic.200800454

Chemical Biology Approaches to Probe the Proteome

Huib Ovaa*[a] and Fred van Leeuwen[b]

Understanding disease-associated cellular defects at a molecular level is critical for the development of pharmacological intervention strategies. Recent breakthroughs in microarray and sequencing technologies have provided powerful tools to rapidly reveal the cellular defects caused by alterations in the genome or transcriptome. However, the picture of how the cellular proteome is affected in a disease state and how changes in DNA and RNA affect protein function is often incomplete. This is perhaps not surprising because the functions of proteins are not just determined by primary sequence and abundance, but are under the control of many regulatory mechanisms. Here, we highlight several recent advances in proteomics technologies that are being developed to generate comprehensive human proteome maps and discuss them in the context of strategies that have been developed in simple model organisms. Chemical biology will play a critical role in drafting a map of the proteome with functional information. Chemical genetic approaches that use high-throughput small molecule screening have resulted in the public availability of small molecule datasets through web interfaces such as PubChem. With such approaches, the opportunities to investigate disease and to explore the proteome with chemistry are rapidly increasing. In addition, new tools are being developed to probe protein function. Here we highlight recent developments in chemical biology and the exciting opportunities that are arising with them.

Introduction

With the advent of high-throughput technologies, enormous progress has been made over the last decade in the understanding of human disease. Defects in the cell that lead to disease include alterations in the genome, changes in messenger RNA (mRNA) expression or processing, and changes at the protein level. Changes in the genome and transcriptome have been best characterized. The recent breakthroughs in high-throughput sequencing and micro-array technologies have enabled efficient and accurate analysis of changes in the genome such as DNA mutations, deletions, amplifications, and even translocations.[1] The same technologies have allowed for the systematic analysis of mRNA expression levels and alternative splicing. The combination of microarrays and sequencing technologies provides a toolbox with a dynamic range that seems sufficient to analyze in parallel the sequence and expression level of both low- and high-abundance transcripts.[2,3] Diseases that are caused by a DNA mutation or by the misexpression of certain genes will most likely give away their secrets to the brute forces that can now be used to interrogate the genome and transcriptome. Translation of the changes in DNA or RNA into phenotypes is not always straightforward. For example, genome-wide studies in yeast have shown that a change in mRNA levels alone may be a poor predictor of functional relevance[4] and that changes in the transcriptome show a poor correlation with changes in the proteome.[5] In addition, translation of mRNAs can be modulated by microRNAs (miRNAs) without affecting mRNA levels.[6] Finally, there seem to be many opportunities for a cell to change its identity or fate without mutations to the genome. Such "epigenetic" changes can be caused by changes of the chromatin, the packaging material of the genome, or by feedback loops in transcriptional circuits.[7,8] Once initiated, such epigenetic events can be propagated in dividing cells like DNA mutations.

Complexity of the Functional Proteome

Thus, the next challenge will be to develop technologies and tools to determine the disease-related changes that occur at the protein level. This is not a simple task. Whereas genome and transcriptome analysis requires in essence only two pieces of information (sequence and abundance), protein function in a given cell or tissue can be affected by many different parameters, such as proteolytic processing (for example, of latent enzymes such as zymogens or of proteins with targeting signals), expression levels, localization, interacting proteins (inhibitors, activators, scaffolds), turnover, catalytic activity, co-factors, proline isomerization, and post-translational modifications (PTMs) including phosphorylation, acetylation, methylation, ubiquitination, sumoylation, lipidation, glycosylation and others (Figure 1). These interactions and modifications determine important biophysical parameters such as affinity for potential protein interaction partners, membrane anchoring, trafficking, stability and overall catalytic activity in the case of enzymes. Here we discuss recent developments in model organisms and

[a] *Dr. H. Ovaa*
Division of Cellular Biochemistry, Netherlands Cancer Institute
Plesmanlaan 121, 1066 CX Amsterdam (NL)
Fax: (+31) 20-5121989
E-mail: h.ovaa@nki.nl

[b] *Dr. F. van Leeuwen*
Division of Gene Regulation, Netherlands Cancer Institute
Plesmanlaan 121, 1066 CX Amsterdam (NL)

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

WILEY InterScience® DISCOVER SOMETHING GREAT

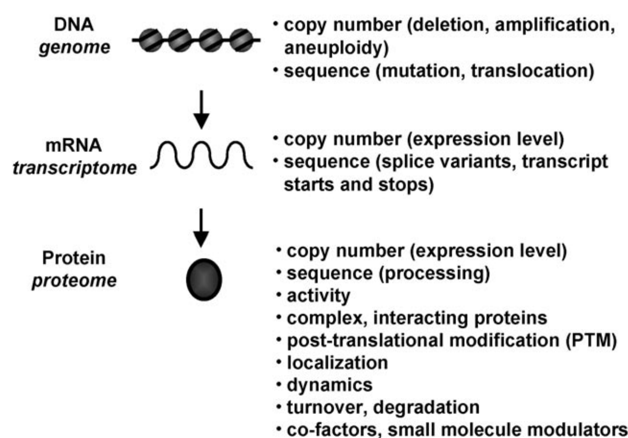

Figure 1. From genome to proteome. Analysis of cellular defects at the level of the proteome requires the integration of many different parameters. Measurement of the different aspects of proteome function will rely on the development of novel chemical tools.

human cells to systematically probe the various and diverse aspects of the proteome and illustrate how chemical approaches will be indispensable to obtain further understanding of proteome function.

Measuring Protein Abundance, Modification, and Interactions

Protein abundance and biophysical properties are key parameters of the proteome. In yeast, the abundance of nearly every protein has been determined by epitope tagging of most of the protein-coding genes. By using immunoblots, the expression of each protein can be quantitated by comparing signals to a reference sample. This approach has led to a comprehensive proteome-scale dataset of steady-state protein expression levels.[9] In addition, epitope-tag strain collections have been used to systematically examine protein half-lives in yeast.[10] For human cells, which are more difficult to manipulate genetically and which come in different cell-types, this approach is unfeasible. In addition, epitope tags may interfere with protein function. Mass spectrometry provides a powerful and more universal approach to measure protein abundance and modification. Advances in mass spectrometry have rapidly increased our understanding of the functional human proteome and allowed for quantitative comparison of proteins and their modifications in different complex samples. Quantitative proteomics takes advantage of differential isotopic labeling techniques. Heavy isotopes can be introduced metabolically into cells or even whole animals using isotopically labeled amino acids,[11,12] or they can be introduced after sample preparation by chemical reactions using designed heavy weight or control reagents.[13] Alternatively, specific (subsets of) proteins can be differentially labeled with dyes and subsequently quantified by fluorescence scanning after enrichment and protein separation.[14] While the dynamic range of recent DNA and RNA technologies is sufficient for genome-wide measurements,[1] analysis of the complete proteome has so far proven too complex to study directly in all of its details. One solution for this problem is to select subproteomes through enrichment procedures that remove abundant, non-specific proteins, which would otherwise interfere with measurements. The combined use of designer reagents and mass spectrometry is a particularly powerful strategy to study enriched proteome subsets. Organic synthesis can provide the tools that help to select smaller subproteomes. Proven strategies are the chemical modification of a drug at predefined positions or its random crosslinking to a retrieval tag or resin (Figure 2). When the drug finds its target,

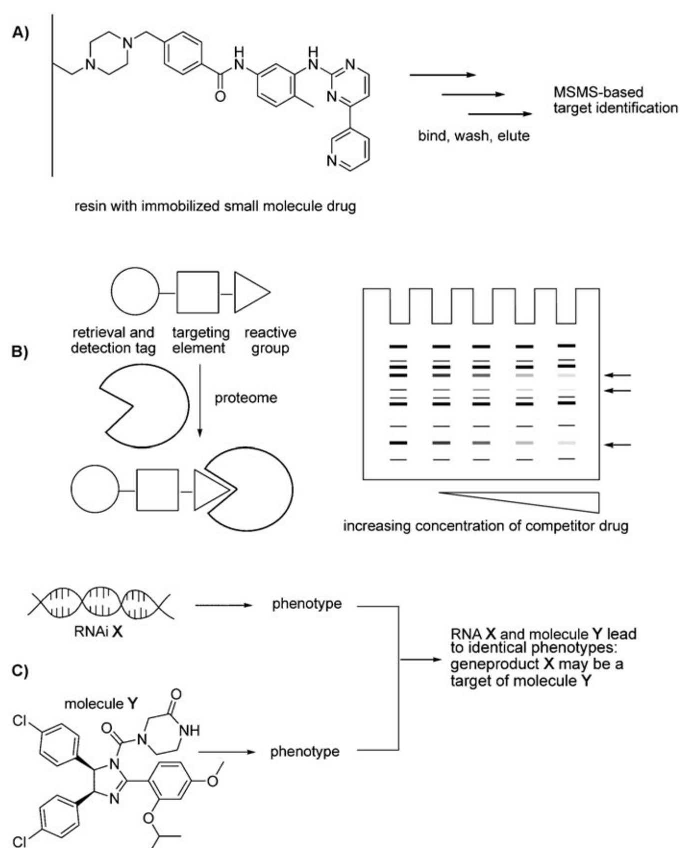

Figure 2. Target identification A) using immobilized drug chromatography or B) using chemical reagents (activity-based probes) to label proteins that can be used in a competition format with small molecules. Competition for labeling reveals molecular targets. C) Selecting target candidates by RNAi.

the resulting complex can be purified using the tag and further analysis by mass spectrometry. This approach is also extremely useful to identify cellular targets and may help to explain or predict the off-target effects of small molecule screening hits or bona fide drugs. Recent examples include the identification of targets of the BCR-ABL kinase inhibitors imatinib, nilotinib and dasatinib, which are used to treat leukemia.[15] Several of the identified targets of these kinase inhibitors include non-kinase substrates. Chemical reagents can also be used to label proteins to enrich proteome subsets for analysis. In these chemical strategies, functional proteins may be modified with a variety of the aforementioned post-translational modifications, allowing their identification and further study.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

An alternative method that has been successfully used to select subproteomes is the use of in vivo tags. For example, by expressing a tagged version of the posttranslational modifier ubiquitin, ubiquitylated proteins can be enriched by affinity chromatography.[16] However, this approach is currently only applicable to a limited set of genetically encoded PTMs. Identification of proteins that bind to other classes of PTMs can be facilitated by the use of protein chips, on which either a library of modified peptides or a collection of proteins (or sub-domains) with putative PTM binding modules are spotted. For example, this concept has been explored successfully to characterize the specificity of proteins that bind to PTMs on specific sites of histones. To find the enzymes responsible for PTMs, enzymes of interest can be assayed on proteome-scale protein chips. For example, protein chips containing most of the proteins encoded by the yeast genome have been used to systematically identify substrates of protein kinases.[17,18] The use of in vivo tags in combination with mass spectrometry has proven invaluable for the identification of proteins that co-purify with a target protein of interest. In yeast, genome-wide collections of strains with tagged proteins have been used to systematically purify each protein and identify protein complexes by mass spectrometry.[19–21] They have also been used to measure the activities of the different purified proteins (or protein complexes) in specific in vitro assays such as phosphodiesterase or kinase assays.[22] Protein–protein interactions have also been analyzed by systematic in vivo approaches, including yeast two-hybrid or protein-fragment complementation.[23–25] Together, these assays provide comprehensive proteome-wide protein–protein interaction maps. A better understanding of the interplay between individual proteins at a molecular level and the availability of molecules that interfere with this interplay will reveal new targets for the development of novel therapies and drugs, and will help us find new applications for existing drugs.

Measuring Protein Localization and Activity

One common way to examine protein localization is to fuse the protein of interest to a fluorescent reporter and determine its sub-cellular localization by immunofluorescence. By fusing nearly every gene of the yeast genome to GFP, a systematic proteome-wide picture of protein abundance and localization has been derived.[26] However, since this approach is not directly applicable to other organisms, examination of protein localization in human cells will require other tools such as the systematic production of good antibodies against each human protein.[27] While protein abundance and localization are obviously important, the biological activity of proteins is of greater significance. Here, chemical biology may provide novel and powerful tools to address this critical aspect of protein function. Chemical reporters offer a direct way to study specific cellular events. Such molecules can report enzymatic activities, protein–protein interactions, the presence of ions or small molecules (cations for example),[28] or act as receptors[29] and phenotypic sensors,[30] among others. Most chemical reporters take advantage of an event-induced change in their spectral characteristics (like an emission spectrum) to allow a readout of action. Some of these reagents may ultimately find use as diagnostic tools or imaging agents to sense protein malfunction and/or disease. Such reporters have also been developed for use in chemical genetic approaches in combination with high-throughput small molecule screening. Active-site directed probes form a special class of sensors that can be used to study various activities simultaneously; this technique is also known as activity-based protein profiling. These reagents have been extensively reviewed in this journal[31] and elsewhere.[32–34] Briefly, reagents for activity-based protein profiling generally contain a reporter module (such as a dye, biotin or radiolabel), a targeting device (a specific peptide or small molecule) that has high affinity for the proteome subset or specific target, and a reactive group or crosslinker that is used to link the targeted reporter module to the receptor or enzyme of choice (Figure 2 B). Depending on the reagents used, this strategy might allow for the retrieval and study of proteins from complex mixtures (cell or tissue lysates) and create opportunities for direct readouts such as microscopy on cells[35,36] or animals.[37] The design of molecular sensors and active-site directed probes has been catalyzed by the development of a number of bio-orthogonal reactions, which are well-suited for use in the presence of complex biological matrices. Notable reactions that have been exploited extensively are the Staudinger ligation[38] and the Huisgen cycloaddition[39,40] as well as improvements on these themes.[37,41] Such bio-orthogonal reactions are of considerable importance as they bring synthetic chemistry closer to biochemical action and facilitate the further development of various chemical probes. Probes based on this theme have, for example, enabled the study of glycosylation patterns in zebra fish[37] and provided many improved alternatives to existing techniques (such as an improved "click" staining as an alternative to bromodeoxyuridine (BrdU) incorporation.[42] Chemists have exploited various combinations of chemistry, purification and mass spectrometry-mediated protein identification techniques to provide novel tools and standards to the scientific community.

Manipulating the Proteome

The availability of yeast deletion mutant collections and the more recent advent of RNA interference (RNAi) technology have allowed for the systematic evaluation of the contribution of individual genes and transcripts to disease, and the expectation for the years to come is high. Systematic genetic knockdown has become routine and is carried out on genomic subsets such as the kinome,[43] comprising all kinases, the "drugable genome," [44] comprising all targets that are likely drugable (for example, enzymes, (ion)channels and receptors) or simply "genome-wide" and "unbiased." Thus, the advent of RNAi technology provides a powerful and general entry to identify genes that play roles in normal physiology and disease. Chemical genetics, a fashionable term for drug screening that has been routine in industry for decades, has become an exciting academic discipline. It typically brings together small molecule libraries and in vitro or cellular assays, as well as informatics

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

and data mining tools for analysis of the data generated. Chemical genetics applies the same principles as traditional genetics, replaces the toolbox with molecules, and focuses on proteins rather than genes. Chemical genetics complements classical genetics by providing probes that can be used to study not easily transfected cells. Individual chemicals can often influence single proteins in a way that gene knockouts and transgenes cannot (Figure 3). Chemical genetics can be divided into forward and reverse modes. In a forward chemical genetics approach, the phenotype induced by a molecule is studied in a cell or tractable organism. Its ultimate target is discovered at a later stage. This approach is identical to classical drug discovery. In a reverse chemical-genetic screen, an enriched or homogeneous protein is tested against a large number of small molecules, which are subsequently selected for further chemical optimization to improve their chemical and biophysical properties. Thus, this is identical to modern high-throughput target-oriented drug discovery and analogous to reversed genetic approaches that use cells or animals, which have been designed to lack a gene of interest to study the genes biological function.

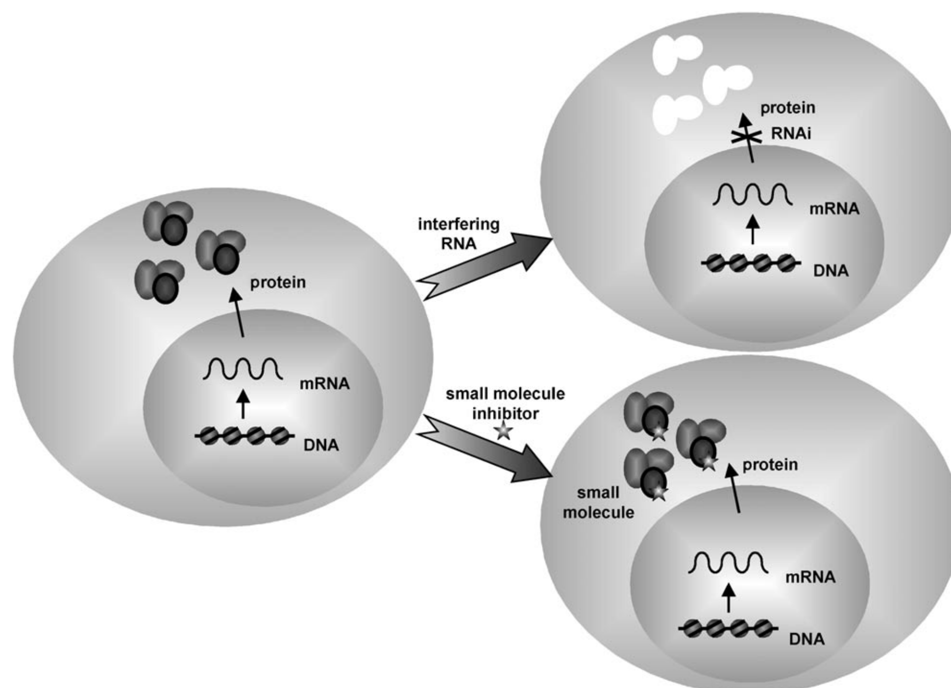

Figure 3. Modulating protein function: RNAi knockdown versus small molecule inhibition. Knockdown of protein expression by RNA interference and inhibition of protein function by small molecules can be used to verify specificity of each method. However, due to mechanistic differences, the two methods can also lead to different phenotypic outcomes.

Chemical Genetics and RNAi Screens

If RNAi and small molecules are directed against the same gene product, then independent methods can often be used to verify whether phenotypes are caused by specific or off-target effects. However, RNAi and small molecules each have their specific advantages and disadvantages (Figure 3). Whereas RNAi is slow due to its mode of action, small molecules normally provide effects that are near instantaneous. Effects are generally quick and reversible when a small molecule is withdrawn, which is a useful trait, especially when essential genes are the topic of study. As RNAi acts on the transcriptional level, its effect depends on the turnover of the protein targeted and on the dilution rate due to cell division. Hence, long-lived proteins in non-cycling cells are not ideal RNAi targets. In addition, application of RNAi often results in a modest knock-down of gene expression only. The final outcome of RNAi is reduced protein production. If the target protein is part of a protein complex, this may in turn affect architecture of the remaining complex and can affect the stability of its protein partners. Many small molecules typically affect the activity of a protein without changing its abundance and thereby leave protein complexes unaltered (Figure 3). Thus, reducing protein activity by using small molecules or RNAi may not always give the same outcome. Which of the two strategies is most useful to achieve the desired phenotype may be different for each protein. When RNAi and small molecule screening are deployed simultaneously, the same cell-based readout can be used, while small molecule hits may eventually be correlated with the outcome of the parallel RNAi screen (Figure 2C). The RNAi screen may provide an initial clue as to whether a target is likely drugable, provided that targets of these molecules can be found among the hits from the parallel RNAi screen. The chemical screen may directly provide a universal entry point for chemical interference in cell types or tissues that are difficult to target with RNAi-based methods. A recent approach[45] used a chemical genetic synthetic lethal screen to identify chemosensitizer loci in cancer cells. In this study, responsiveness to paclitaxel, which is an agent that targets tubulin, was investigated in a one-well, one-gene RNAi knockdown setup to identify genes that affect sensitivity to this agent; this uncovered both known and novel target proteins. Another recent example of forward genetics is the recent discovery of aminoacetonitrile derivatives that function as resistance-breaking antinematode drugs. They were identified to target a nematode-specific subtype of nicotinic acetylcholine receptor (nAChR).[46] Definition of this nematode-specific target with a molecule-initiated approach might pro-

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

vide new solutions to tackle nematode infections. Another recent example of forward genetics is the use of kinase inhibitors to identify kinases that reduce bacterial replication in host cells. In combination with RNA screens, this approach has led to the identification of a network of kinases that is abused by both *Salmonella* and *Mycobacterium tuberculosis* to facilitate infection and replication in the host. Interestingly, PKB/Akt, which is a central player in this network, is also being exploited for the development of PKB/Akt-targeted cancer therapies by many pharmaceutical companies. Thus, although classical genetics and more recently systematic RNAi-mediated knockdown strategies are by far superior to identify gene products that regulate different biological processes, chemical genetic approaches are often a great help to understand in molecular detail how proteins perform their biological functions. Importantly, chemical genetics identifies small molecules that might be developed into a drug of medical value. It is also important to note here that most post-translational events cannot easily be investigated with common genetic techniques, as the post-translational machinery is complex and dynamic. Often a great redundancy in PTM-modifying enzymatic activities are present in a given cell. That is, a particular phosphorylation event often cannot easily be attributed to a single kinase. The action of opposing phosphatases should also be taken into account. While this complexity is true for most PTMs, this problem can typically be addressed by chemical approaches.

High-Throughput Small Molecule Screening

With high-throughput screening platforms available at many academic scientific institutes, massive amounts of small molecule screening data are being generated. These chemical genetics approaches are often publicly available for further exploration by organic synthesis. The PubChem project, launched only a few years ago, in 2004, has already resulted in an impressive set of publicly available small molecule information and screening data. PubChem is a component of the NIH's Molecular Libraries Roadmap Initiative and provides information on biological activities of small molecules and can be accessed through PubMed. The PubChem interface is not yet as user-friendly as some of the other PubMed tools, but the initiative is still in development. These publicly available datasets are increasing in size at rapid pace. Both quality and chemical synthesis are crucial to turn these screening data into successful chemical probes and drugs that can chemically tamper with and understand the proteome. However, the trend of "unbiased high-throughput testing" has some disadvantages. First of all, no method is truly unbiased, as assay conditions inevitably influence the outcome of any experiment. Sometimes the rational approach is not so bad; many false positives are frequently encountered in high-throughput screening campaigns. Often, these false positives are not apparent at first glance but turn out to interfere directly with assay readout or prove to be efficient protein precipitants, while some small molecule "frequent hitters" seem to score well in any assay.[47] Even a validated hit in a screen does not necessarily lead to a useful molecule. Occasionally, molecules identified as hits in screens are no longer commercially available or only available in limited amounts. In addition, sometimes hits are caused by impurities in the library or by structures that were wrongly assigned, which underscores the importance of thorough quality control and resynthesis of hits. Finally, after sorting out molecules and addressing these issues, the chemist is often left with a molecule of little charm or with little chemical opportunity for further modification. Nonetheless, chemical opportunities from such screening efforts are emerging at a rapid pace and complementary techniques that include fragment-based screening and in silico screens are rapidly winning ground in the screening arena. Numerous examples are available of successful forward and reverse chemical genetics approaches, including inhibitors of non-replicative bacteria,[48] inhibitors of bacterial replication that act on the host,[49] small molecules that induce the conversion of procaspase-3 to active caspase-3 to promote apoptosis[50] and modulators of autophagy.[51]

Challenges in Chemical Genetics and Proteomics

Various chemical genetics screening options are available to the investigator, but all have their specific advantages and limitations. In vitro screens that use (semi)purified protein components are straightforward in practice but can lead to the identification of promiscuous inhibitors and many other false positives. Cell based screens avoid many of these problems. Microscopy-based phenotypic screens are becoming routine, facilitated by the commercial availability of microscopy-based screening platforms. In vivo screens have the disadvantage that targets remain unknown until a serious effort is undertaken to identify them. In some cases, the homozygous and heterozygous yeast knock-out collections have been useful in the identification of specific and non-specific targets of a given drug or small molecule.[52–55] It is expected that next to the drugable genome that includes receptors, ion channels and enzymes, proteomics strategies will uncover a range of important protein–protein interactions. The number of important interactions likely exceeds the number of currently drugable genome components by far. Although biopharmaceuticals almost exclusively bind to protein surfaces that are required for protein–protein interactions, small molecules are not easily designed to do so. Protein–protein interactions in general cover large surfaces areas compared to the traditional small interaction surfaces (normally specific receptors or enzymatic catalytic clefts or pockets) of routine drug targets. Most commercially available compound collections have been designed to fit small clefts and to obey strict biophysical properties, which in principle makes them less suited to bind larger surfaces. Despite these difficulties in targeting protein–protein interaction surfaces, great progress has been made over the last years, but it is still a long way to go before such molecules can be discovered routinely mainly due to the limits set by current molecule collections. Systematic knowledge of 3D structures of target surfaces will greatly improve such designs.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Outlook

The available proteomics strategies, which so far have resulted in maps of pathways, will eventually afford a near complete proteome atlas of proteins. This atlas will include information about protein interactions and the pathways they participate in as well as their levels of expression, tissue distribution, activity, and disease related data. Although it may take some time to get to this point, ambitious projects such as the Human Proteome Research Project aim exactly at this: a complete proteome map including validated antibodies directed against all human proteins is being pursued. This information is becoming available through efforts such as the human Proteinpedia, a web interface to which contributors add proteomics data accompanied with experimental evidence.[27] However, proteomics strategies remain challenging because of difficulties with quantification and the occasional lack of appropriate internal standards, which makes comparisons of different studies difficult. A clear advantage of the public availability of all the proteomics data sets is that they will allow researchers to determine how proteins interact and to cluster information into pathways. This will enable the design of strategies to target multiple pathways with combinations of pathway-specific drugs, which might increase chances of success and reduce the chance of drug resistance. Chemical genetics and proteomics approaches are expected to dramatically increase the number of potential drug targets and lead compounds. Further development of the new hits will depend on the rapidly advancing field of structural biology to provide a molecular understanding of protein function and protein-small molecule interactions, while the design of novel molecules will likely always remain the exclusive area of organic chemists.

Selected small molecule screening and chemical bioinformatics databases:

PubChem is a database that contains small molecule information linked to biochemical data. The database is maintained by the National Center for Biotechnology Information (NCBI) of the National Library of Medicine, part of the United States National Institutes of Health (NIH). Millions of structures and associated descriptive datasets as well as high-throughput screening results can be accessed for free through a web interface. PubChem: http://pubchem.ncbi.nlm.nih.gov/.

ChemBank is a public, web-based informatics environment created by the Broad Institute's Chemical Biology Program, part of the National Cancer Institute Initiative for Chemical Genetics. ChemBank data are also available through PubChem. ChemBank intends to guide chemists who are synthesizing novel compounds or libraries, to assist development of small molecule probes that perturb specific biological pathways, and to catalyze drug development. ChemBank: http://chembank.broad.harvard.edu/.

DrugBank is a database available at the University of Alberta that combines detailed chemical, pharmacological and pharmaceutical data combined with drug target information (that is, sequence, structure, pathway). The database contains roughly 4800 entries that include FDA-approved small molecule drugs, FDA-approved biotech (protein/peptide) drugs, nutraceuticals and experimental drugs. In addition, more than 2500 protein sequences are linked to these drug entries. DrugBank: http://www.drugbank.ca/.

Chemspider is a search engine for chemistry that contains chemical structures and their associated information, including vendors and links to other databases such as PubChem into a single database that can be searched by anyone. Chemspider: http://www.chemspider.com.

Acknowledgements

The Ovaa and van Leeuwen laboratories are supported by grants from the Netherlands Organization for Scientific Research, the Dutch Cancer Society and the European Union. We thank Anna Houben and Kitty Verzijlbergen for critical reading of the manuscript.

Keywords: chemical genetics · drug development · proteomics · RNA interference · screening

[1] E. R. Mardis, *Trends Genet.* **2008**, *24*, 133–141.
[2] B. T. Wilhelm, S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C. J. Penkett, J. Rogers, J. Bähler, *Nature* **2008**, *453*, 1197–1198.
[3] U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, M. Snyder, *Science* **2008**, *320*, 1344–1349.
[4] G. W. Birrell, J. A. Brown, H. I. Wu, G. Giaever, A. M. Chu, R. W. Davis, J. M. Brown, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8778–8783.
[5] E. J. Foss, D. Radulovic, S. A. Shaffer, D. M. Ruderfer, A. Bedalov, D. R. Goodlett, L. Kruglyak, *Nat. Genet.* **2007**, *39*, 1369–1375.
[6] O. Hobert, *Science* **2008**, *319*, 1785–1786.
[7] P. A. Jones, S. B. Baylin, *Cell* **2007**, *128*, 683–692.
[8] S. R. Bhaumik, E. Smith, A. Shilatifard, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1008–1016.
[9] S. Ghaemmaghami, W. K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea, J. S. Weissman, *Nature* **2003**, *425*, 737–741.
[10] A. Belle, A. Tanay, L. Bitincka, R. Shamir, E. K. O'Shea, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13004–13009.
[11] S.-E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376–386.
[12] M. Kruger, M. Moser, S. Ussar, I. Thievessen, C. A. Luber, F. Forner, S. Schmidt, S. Zanivan, R. Fassler, M. Mann, *Cell* **2008**, *134*, 353–364.
[13] R. Raijmakers, C. R. Berkers, A. de Jong, H. Ovaa, A. J. R. Heck, S. Mohammed, *Mol. Cell Proteomics* **2008**, *7*, 1755–1762.
[14] B. M. Riederer, *J. Proteomics* **2008**, *71*, 231–244.
[15] U. Rix, O. Hantschel, G. Durnberger, L. L. Remsing Rix, M. Planyavsky, N. V. Fernbach, I. Kaupe, K. L. Bennett, P. Valent, J. Colinge, T. Kocher, G. Superti-Furga, *Blood* **2007**, *110*, 4055–4063.
[16] J. Peng, *BMB. Rep.* **2008**, *41*, 177–183.
[17] J. Ptacek, G. Devgan, G. Michaud, H. Zhu, X. Zhu, J. Fasolo, H. Guo, G. Jona, A. Breitkreutz, R. Sopko, R. R. McCartney, M. C. Schmidt, N. Rachidi, S. J. Lee, A. S. Mah, L. Meng, M. J. Stark, D. F. Stern, C. De Virgilio, M. Tyers, B. Andrews, M. Gerstein, B. Schweitzer, P. F. Predki, M. Snyder, *Nature* **2005**, *438*, 679–684.
[18] J. Ptacek, M. Snyder, *Trends Genet.* **2006**, *22*, 545–554.
[19] A. C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, G. Superti-Furga, *Nature* **2002**, *415*, 141–147.
[20] N. J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M. D. Robinson, A. Paccanaro, J. E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Staros-

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tine, M. M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H. Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J. S. Weissman, C. J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, J. F. Greenblatt, *Nature* **2006**, *440*, 637–643.

[21] A. C. Gingras, M. Gstaiger, B. Raught, R. Aebersold, *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 645–654.

[22] M. R. Martzen, S. M. McCraith, S. L. Spinelli, F. M. Torres, S. Fields, E. J. Grayhack, E. M. Phizicky, *Science* **1999**, *286*, 1153–1155.

[23] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. M. Rothberg, *Nature* **2000**, *403*, 623–627.

[24] T. Ito, K. Ota, H. Kubota, Y. Yamaguchi, T. Chiba, K. Sakuraba, M. Yoshida, *Mol. Cell. Proteomics* **2002**, *1*, 561–566.

[25] K. Tarassov, V. Messier, C. R. Landry, S. Radinovic, M. M. Molina, I. Shames, Y. Malitskaya, J. Vogel, H. Bussey, S. W. Michnick, *Science* **2008**, *320*, 1465–1470.

[26] W. K. Huh, J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, E. K. O'Shea, *Nature* **2003**, *425*, 686–691.

[27] The Human Proteinpedia Consortium, *Nat. Biotechnol.* **2008**, *26*, 164–167.

[28] B. Tang, L. J. Cui, K. H. Xu, L. L. Tong, G. W. Yang, L. G. An, *ChemBioChem* **2008**, *9*, 1159–1164.

[29] F. Krieger, A. Mourot, R. Araoz, F. Kotzyba-Hibert, J. Molgo, E. Bamberg, M. Goeldner, *ChemBioChem* **2008**, *9*, 1146–1153.

[30] B. K. Wagner, H. A. Carrinski, Y. H. Ahn, Y. K. Kim, T. J. Gilbert, D. A. Fomina, S. L. Schreiber, Y. T. Chang, P. A. Clemons, *J. Am. Chem. Soc.* **2008**, *130*, 4208–4209.

[31] M. Uttamchandani, J. Li, H. Sun, S. Q. Yao, *ChemBioChem* **2008**, *9*, 667–675.

[32] K. T. Barglow, B. F. Cravatt, *Nat. Methods* **2007**, *4*, 822–827.

[33] M. Fonovic, M. Bogyo, *Curr. Pharm. Des.* **2007**, *13*, 253–261.

[34] H. Ovaa, *Nat. Rev. Cancer* **2007**, *7*, 613–620.

[35] C. R. Berkers, F. W. van Leeuwen, T. A. Groothuis, V. Peperzak, E. W. van Tilburg, J. Borst, J. J. Neefjes, H. Ovaa, *Mol. Pharm.* **2007**, *4*, 739–748.

[36] M. Verdoes, B. I. Florea, V. Menendez-Benito, C. J. Maynard, M. D. Witte, W. A. van der Linden, A. M. van den Nieuwendijk, T. Hofmann, C. R. Berkers, F. W. van Leeuwen, T. A. Groothuis, M. A. Leeuwenburgh, H. Ovaa, J. J. Neefjes, D. V. Filippov, G. A. van der Marel, N. P. Dantuma, H. S. Overkleeft, *Chem. Biol.* **2006**, *13*, 1217–1226.

[37] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664–667.

[38] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.

[39] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

[40] R. Manetsch, A. Krasinski, Z. Radic, J. Raushel, P. Taylor, K. B. Sharpless, H. C. Kolb, *J. Am. Chem. Soc.* **2004**, *126*, 12809–12818.

[41] J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16793–16797.

[42] A. Salic, T. J. Mitchison, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2415–2420.

[43] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science* **2002**, *298*, 1912–1934.

[44] A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730.

[45] A. W. Whitehurst, B. O. Bodemann, J. Cardenas, D. Ferguson, L. Girard, M. Peyton, J. D. Minna, C. Michnoff, W. Hao, M. G. Roth, X. J. Xie, M. A. White, *Nature* **2007**, *446*, 815–819.

[46] R. Kaminsky, P. Ducray, M. Jung, R. Clover, L. Rufener, J. Bouvier, S. S. Weber, A. Wenger, S. Wieland-Berghausen, T. Goebel, N. Gauvry, F. Pautrat, T. Skripsky, O. Froelich, C. Komoin-Oka, B. Westlund, A. Sluder, P. Maser, *Nature* **2008**, *452*, 176–180.

[47] B. K. Shoichet, *Drug Discovery Today* **2006**, *11*, 607–615.

[48] R. Bryk, B. Gold, A. Venugopal, J. Singh, R. Samy, K. Pupek, H. Cao, C. Popescu, M. Gurney, S. Hotha, J. Cherian, K. Rhee, L. Ly, P. J. Converse, S. Ehrt, O. Vandal, X. Jiang, J. Schneider, G. Lin, C. Nathan, *Cell Host Microbe* **2008**, *3*, 137–145.

[49] C. Kuijl, N. D. Savage, M. Marsman, A. W. Tuin, L. Janssen, D. A. Egan, M. Ketema, R. van den Nieuwendijk, S. J. van den Eeden, A. Geluk, A. Poot, G. van der Marel, R. L. Beijersbergen, H. Overkleeft, T. H. Ottenhoff, J. Neefjes, *Nature* **2007**, *450*, 725–730.

[50] K. S. Putt, G. W. Chen, J. M. Pearson, J. S. Sandhorst, M. S. Hoagland, J. T. Kwon, S. K. Hwang, H. Jin, M. I. Churchwell, M. H. Cho, D. R. Doerge, W. G. Helferich, P. J. Hergenrother, *Nat. Chem. Biol.* **2006**, *2*, 543–550.

[51] S. Sarkar, E. O. Perlstein, S. Imarisio, S. Pineau, A. Cordenier, R. L. Maglathlin, J. A. Webster, T. A. Lewis, C. J. O'Kane, S. L. Schreiber, D. C. Rubinsztein, *Nat. Chem. Biol.* **2007**, *3*, 331–338.

[52] M. E. Hillenmeyer, E. Fung, J. Wildenhain, S. E. Pierce, S. Hoon, W. Lee, M. Proctor, R. P. St Onge, M. Tyers, D. Koller, R. B. Altman, R. W. Davis, C. Nislow, G. Giaever, *Science* **2008**, *320*, 362–365.

[53] K. Baetz, L. McHardy, K. Gable, T. Tarling, D. Reberioux, J. Bryan, R. J. Andersen, T. Dunn, P. Hieter, M. Roberge, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4525–4530.

[54] P. Y. Lum, C. D. Armour, S. B. Stepaniants, G. Cavet, M. K. Wolf, J. S. Butler, J. C. Hinshaw, P. Garnier, G. D. Prestwich, A. Leonardson, P. Garrett-Engele, C. M. Rush, M. Bard, G. Schimmack, J. W. Phillips, C. J. Roberts, D. D. Shoemaker, *Cell* **2004**, *116*, 121–137.

[55] H. Luesch, T. Y. Wu, P. Ren, N. S. Gray, P. G. Schultz, F. Supek, *Chem. Biol.* **2005**, *12*, 55–63.

Received: July 6, 2008
Published online on October 29, 2008

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim