Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Chemical Biology

Proteomics and genomics: perspectives on drug and target discovery

Editorial overview

Natalie G Ahn and Andrew H-J Wang

Current Opinion in Chemical Biology 2008, **12**:1–3

Available online 11th March 2008

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2008.02.016

Natalie G Ahn

University of Colorado, Department of Chemistry and Biochemistry, Campus Box 215, Boulder, CO 80309-0215, United States
E-mail: natalie.ahn@colorado.edu

Natalie G Ahn is a Professor of Chemistry and Biochemistry and Investigator of the Howard Hughes Medical Institute at the University of Colorado at Boulder. She received her B.S. degree in Chemistry from the University of Washington, and her Ph.D. in Chemistry from the University of California, Berkeley. Her research examines mechanisms by which signal transduction pathways control cancer, by integrating technologies of proteomics and mass spectrometry with biochemistry and molecular biology.

Andrew H-J Wang

Institute of Biological Chemistry, Academia Sinica, 128 Academia Road, Nankang, Taipei, Taiwan

Andrew H-J Wang was educated in Taiwan (B.S., M.S. in Chemistry, National Taiwan University) and received his Ph.D. in Chemistry, University of Illinois (1974). He worked in the Department of Biology, MIT, 1974–1988; was a Professor of Biochemistry/Biophysics/Chemistry at the University of Illinois (Urbana), 1988–2000. He is presently Vice President of Academia Sinica; Distinguished Research Fellow, Institute of Biological Chemistry; elected Academician of Academia Sinica, Taiwan and Fellow, Third World Academy of Sciences. He uses structural proteomics to understand the functions of important bio-systems, for example, structural studies of enzyme targets for drug discovery and state-of-the art proteomic studies of post-translational protein modifications.

Genomics and proteomics technologies together with computational methods have led to powerful new strategies in basic and clinical research. It is useful to reflect on the success of these methods toward developing therapeutic strategies for human diseases. We focus in this issue on recent progress and innovations utilizing 'omics' technologies to identify and validate drug targets, discover disease biomarkers, and design more effective drugs.

Proteins involved in signal transduction pathways have emerged as candidate drug targets in many pathologies, and it is not surprising that leading candidates are regulatory enzymes which catalyze protein post-translational modifications. Proteins targeted by these enzymes in turn provide further candidates for drug therapeutics and potential cellular markers of disease. Prevalent in many diseases, the large class of protein kinases accounts for ~25% of research and development projects in biotechnology and pharma. B. Turk summarizes advancements in chemical and protein analyses to understand the range of substrates recognized by different protein kinases, and what controls substrate recognition. Enzymatic features governing specificity for kinase substrates and inhibitors underlie key aspects of drug design. Sirtuins represent another important protein class with diverse functions, including family members with deacetylation activity or ADP ribosylation activity. SIRT1 has been implicated in several age-related diseases such as type 2 diabetes. Its expression increases in animals following caloric restriction, with therapeutic potential indicated for drugs that elevate catalytic activity. Milne and Denu present a detailed perspective on small molecules that allosterically activate or inhibit SIRT1 and SIRT2, and their use as probes of enzyme mechanism and function.

Nonenzymatic protein oxidation is a widespread oxidative stress response, demonstrated in many disease states. Poole and Nelson discuss the chemistry of reactive cysteine oxidation residues, catalyzed by hydrogen peroxide generated as a second messenger in response to localized reactive oxygen species (ROS) production. Key targets for cysteine oxidation to sulfenic acid and further oxidized derivatives include transcription factors and protein tyrosine phosphatases, as well as antioxidant enzymes that themselves catalyze reduction of cysteine-oxidized substrates. However, the breadth of protein oxidation targets is still difficult to assess, because of the transient and localized nature of ROS production. The authors describe exciting new chemical strategies for directly monitoring oxidized cysteine in proteins, which promises to expand our understanding of this large class of targets.

Disorders ranging from Alzheimer's disease to cystic fibrosis arise when proteins become denatured or misfolded, which have serious consequences

relating to the proteins themselves or their environment. In the review by Luheshi *et al.*, issues associated with protein misfolding, particularly gain of toxicity, are elegantly discussed. The authors try to link the protein misfolding and resulting amyloid fibrils with the more qualitative observations made in disease animal models, and discuss how biophysics and computational methods can be integrated to develop underlying physicochemical models of aggregation.

The availability of structural information on proteins, especially in complex with small-molecule ligands, is key to defining the requirements for drug specificity and selectivity. New technologies have accelerated protein structural analysis, including the availability of powerful synchrotrons to enable the collection of high-resolution diffraction data from small crystals, and methods for high-throughput protein production and crystallization. In two papers by Weigelt *et al.* and Marsden and Knapp, the authors review the current status of structural genomics, with particular attention to its impact on drug design. Weigelt *et al.* document current progress of public structural genomics centers, highlighting their importance in generating structures of new targets and coverage of druggable protein families, as well as providing information about solubility and crystallization conditions. Marsden and Knapp focus specifically on protein kinases, illustrating how data mining approaches are now possible with the large structural database, which includes 118 distinct kinase domains and 762 kinase-ligand cocrystal structures. It is pointed out that structural coverage is still rather poor for many kinase families and cocrystal complexes with diverse inhibitors are available in only a few examples. Nevertheless, studies by the authors show how computational approaches can be effective to examine the diversity of kinase-inhibitor structures and discern unique modes of drug interactions.

Classic pharmacological approaches are equally important for discovery, and complementary to target-directed small-molecule screening. Starting with small molecules that elicit cellular responses, screening methods are developed to identify the targets of these compounds. A comprehensive review by Sleno and Emili describes recent strategies for identifying protein targets of bioactive molecules, ranging from capture methods based on affinity binding and activity-based enzyme identification, to protein microarrays for screening small-molecule-binding interactions, and arrayed small-molecule 'ligand chips' for identifying binding proteins. Also discussed are methods for target identification, using global proteomics to profile proteins differentially expressed between normal versus disease states, which when validated serve as candidates for drug development.

In contrast to protein-based screening methods, where protein sequencing directs gene identification in target discovery, Yashiroda *et al.* discuss the inverse strategy of starting with large-scale 'ORFeome' libraries of protein expression plasmids. These enable systematic expression of proteins, which can be used in arrays or pooled formats in order to screen for binding to biomolecules or small-molecule inhibitors, or to screen based on enzymatic activity or protein post-translational modifications. Likewise, tagged ORFs can be expressed in cells and probed by epifluorescence or immunofluorescence, in order to understand subcellular localization and compartmentalization. Ultimately, protein expression studies aim not only to characterize protein location but also to be able to tell us where and when a protein is active. VanEngelenburg and Palmer review genetically encoded biosensors developed to function in living cells as readouts of protein activity. Single fluorescence and FRET-based sensors have been successful in reporting small molecule and ion concentration, protein–ligand binding, and protein–protein interactions. Particularly exciting are protein biosensors that act as surrogate substrates for enzymes, reporting localized activities of kinases, protease, and small GTPases.

Complementary to genomics and proteomics are emerging technologies for metabolite profiling, identifying small molecules as phenotypic readouts of cell perturbations or diseases. In the article by Shyur and Yang, new technology platforms and their applications are surveyed. For example, comparative metabolomics are used for monitoring disease development, drug metabolism, and chemical toxicology, particularly in phytomedicine research. An important focus is on herbal medicine using traditionally tested natural products, where metabolomes of medicinal plants are a rich resource for the development of new phytotherapeutics and nutraceuticals. Given the importance of plant secondary metabolites as inhibitors of therapeutic targets, there is clear value in understanding metabolite signatures in extracts of herbal remedies or traditional Chinese medical plants.

Critically needed for medicine are biomolecules that report disease states with specificity and sensitivity, and can serve in detecting disease and/or assessing drug efficacy. Biomarkers of specific diseases are rare and require costly effort to find. Simpson *et al.* discuss the special case of investigating proteins in blood and other fluids for biomarker discovery. The review covers the many technical challenges of profiling proteins in fluids, and discusses strategies for detection and validation. Not only is biomarker discovery difficult, assays to detect biomarkers that can be used clinically are extremely challenging to develop, because of low abundance, degradation, and binding competition from other proteins in fluids. One hope is that protein arrays will eventually provide a rapid and specific assay to screen multiple biomarkers in patient fluids. Zichi *et al.* reviews the current state of protein microarrays for diagnostics, dis-

cusses strengths and limitations of current arrays, and compares technologies which employ antibody probes to promising aptamer-based probe technologies.

Glycan arrays represent another emerging technology for diagnostics and discovery. Glycoconjugates play a central role in molecular recognition at the cell surface, mediating cell–cell adhesion, host–pathogen interactions, and immune responses. Liang *et al.* reviews recent advances in this technology, from basic aspects of developing carbohydrate libraries attached to solid supports, to quantifying carbohydrate–protein interactions and profiling cancer cell markers. Recent successes include determining the carbohydrate specificity used by viruses to recognize and bind to host cells through their surface glycoproteins. Unique glycoepitopes of key glycoproteins involved in the life cycle of viruses are signatures that can be exploited for viral detection and vaccine development.

Mechanisms underlying host–pathogen interactions are an active research area, with the goal of developing new vaccines and antimicrobial drugs. Wu *et al.* review genomics and proteomics strategies to discover virulence factors of pathogenic bacteria, including cell-surface membrane proteins that mediate host–pathogen interactions, and proteins secreted from pathogen to host. Comparative genomics is used to identify virulence factors with conserved function between pathogens, while proteomics approaches are used to identify bacterial proteins and their post-translational modifications within different bacterial compartments. In recent years, infectious diseases derived from high profile viruses (HIV, SARS-CoV, and H5N1) have had widespread impact around the world. Fast and reliable methods for the early detection of viruses are important for the accurate diagnosis of associated diseases. M. von Itzstein reviews recent advances in the area of influenza virus, highlighting drug discovery and development for viral sialidase enzymes. The highly pathogenic avian influenza virus H5N1 is reviewed, with particular attention to understanding the determinants for human-to-human transmission and new directions for therapeutic intervention.

Last but not least is a most intriguing article by Zaneveld *et al.* on metagenomics analysis of organisms (such as human and mouse) that host microbial symbionts. The host environment, mainly gut, influences the genetic composition of microbiota, which in turn influences organism responses to metabolites and drug treatment. With new methods for assaying microbial compositions, the authors and other researchers are currently addressing the hypotheses of coevolution — reciprocal adaptation between bacterial symbionts and host in response to each other. These are supported by recent experimental findings, which demonstrate clear links between mouse models of human diseases, such as obesity, and genes enriched in gut microbial communities.

Significant advances have been made in recent years in developing new experimental strategies toward drug target discovery, biomarker discovery, and drug design. While some approaches are still restricted to the realm of basic research, many of these technologies are now applied to organismal models of disease as well as human pathology specimens. The integration of chemical biology and large-scale analyses by genomics and proteomics together with methods in computational biology and structural determination provide new perspectives for addressing problems in clinical medicine, and will continue to evolve rapidly to improve drug and target discovery.