Clinical Proteomics Journal

ISSN 1542-6416/04/01:3–5/$25.00

Viewpoint

Moving Forward With Clinical Proteomics

*Sam Hanash**

Department of Pediatrics, Univeristy of Michigan Medical School,
University of Michigan, Ann Arbor, MI, 48109-0656

A remarkable development in the post-genome era is the emergence of proteomics as a new discipline that is rooted in old fashioned chemistry and biochemistry, as well as molecular biology and genomics. Chemistry has provided basic strategies for sorting proteins into classes with different properties and for elucidating their posttranslational modifications, and biochemistry has contributed decades of know-how in elucidating functional aspects of proteins and their involvement in different pathways. Genomics has provided a sequence-based framework for mining proteomics data. Molecular biology has contributed numerous tools for manipulating proteins and for elucidating their interactions and their occurrence as part of complexes.

What about clinical proteomics? The past few years have seen a tremendous interest in the potential ability of proteomics to address many unfulfilled needs in clinical research. Such unmet needs include more effective strategies for early disease detection and monitoring, more effective therapies, and developing a better understanding of disease pathogenesis. Proteomics is particularly suited for investigating biological fluids to identify disease related alterations and to develop molecular signatures for disease processes. A mobilization effort is under way involving academia, governments, industry, and philanthropy to develop agendas for medical proteomics. This is reflected for example in the prominence of proteomics in the National Institues of Health (NIH) roadmap, in the interest on the part of major medical centers to develop plans for faculty recruitment, for investing in proteomics resources, and for developing a focus on specific diseases based on areas of strengths at their institutions. It is also reflected in the prominence given to proteomics at annual meetings of societies that have a disease focus, from heart disease to cancer.

*Author to whom all correspondence and reprint requests should be addressed:
E-mail: shanash@umich.edu

With heightened expectations that proteomics will deliver and overcome some of the limitations of other approaches, or at least complement them, it is important to keep in mind that a solid foundation for the field first needs to be developed Such a foundation includes the development of adequate resources and technologies as well as training and education programs. There is also a need for vehicles for information dissemination, as *Clinical Proteomics* intends to provide, and judging from the contents of this first issue, it is evident that the scope of *Clinical Proteomics* is well suited for the field.

What kind of resources and technologic innovations are needed in clinical proteomics? Proteomics involves the analysis of the global proteome of a cell, a tissue, or a biological fluid. An added requirement for clinical proteomics is the need to analyze large numbers of samples, given the substantial variability one encounters with disease samples. The variability results not only from heterogeneity that characterizes disease states but also from numerous others sources such as the sample procurement process itself. Consequently, there is a need for high-throughput to analyze samples in sufficient numbers to reach statistical significance. A further challenge is to figure out how to develop proteome-scale quantitative read-outs to fulfill expectations of advancing our understanding of basic biology and of disease processes. There is currently no single technology that allows such analyses and much effort is intended to further develop proteomics technologies to expand their reach. Nevertheless, even with the limited scope of current proteomics technologies, numerous disease related investigations may be envisaged. For example, subproteomes are more amenable to comprehensive profiling than the whole proteome and investigations of alterations involving specific cell compartments or specific subsets of proteins are quite feasible. Also, classifying disease based on proteomic profiling does not necessitate analysis of all protein constituents. Identifying novel protein targets for therapy is also feasible with current technologies. It should, however, be emphasized that no matter how limited the current technologies might be, expediency is not a substitute for rigor in scientific inquiry, from experimental design to data interpretation to validation, the latter being particularly crucial to determine clinical relevance.

Besides technology development, there is a need for developing resources to facilitate the application of proteomics to disease investigations. A case in point, the Human Proteome Organization (HUPO.org) is engaged in an effort to develop such resources. There is a substantial need for informatics resources in proteomics for practically every aspect of the field. For example, the current approaches to the analysis of protein data are highly informal and nonstandardized. An important informatics related effort that HUPO is involved in is aimed at developing and adopting standardized approaches that facilitate the cross-lab analysis of proteomics data through a Proteomics Standards Initiative (PSI). HUPO hopes that publishers will adopt recommendations resulting from this initiative. The initial focus has been on protein–protein interaction data and mass spectrometry data. HUPO PSI has proposed a community standard data model for the representation and exchange of protein interaction data. This data model is supported by major protein interaction data providers. Progress has also been made in the development of common standards for data exchange in the field of mass spectrometry. HUPO informatics efforts will also extend to other aspects of proteomics such as quantitative protein expression analysis. They will also address the needs of HUPO's targeted tissue and disease proteomics initiatives with respect to data collection, storage, and dissemination, which will be broadly relevant to clinical proteomics.

Another major resource related initiative is the development of genome scale protein capture agents. The plans under development consist of initially making antibodies for the proteins identified as part of the HUPO proteome projects and eventually making antibodies to all the proteins encoded in the human genome. These antibodies will allow proteome scale studies to be done as in the case of antibody microarrays that assay proteins in a tissue or biological fluid.

An issue that has become quite important in this era of genome and proteome scale investigations is data sharing. The capacity to generate data far exceeds the ability of one group to fully mine such data. Furthermore, it is advantageous for data mining to have access to multiple sets of data. It is therefore crucial that published data is accessible to other investigators. In this respect, investigators, funding agencies, and publishers share a responsibility to facilitate access to data.

Clinical Proteomics will undoubtedly chronicle the progress made in this growing field through its articles and special features.