

UNIVERSITY OF WESTMINSTER



WestminsterResearch

<http://www.wmin.ac.uk/westminsterresearch>

Proteomics: more than just biochemistry.

Pamela Greenwell
Sanjiv Rughooputh

School of Biosciences

This is an electronic version of an article published in *Biomedical Scientist*, 9 (6). pp. 459-460, May 2005.

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners. Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch. (<http://www.wmin.ac.uk/westminsterresearch>).

In case of abuse or copyright appearing without permission e-mail wattsn@wmin.ac.uk.

Modern dictionaries of science would be incomplete without reference to terms that end in the suffix 'omics'. Genomics, proteomics, metabolomics and associated bioinformatics tools are set to achieve a kind of biological alchemy, as Pamela Greenwell and Sanjiv Rughooputh explain in a brief introduction to the subject.

Proteomics: more than just biochemistry

Growth of the so-called 'omics' technologies may appear to be simply a case of giving new, trendy names to boring, old-fashioned technology. After all, who would dispute that genomics is just genetics and molecular biology, glycomics is the study of glycosylation or that metabolomics is simply the study of metabolic pathways? Surely, then, proteomics is just biochemistry. Indeed, students are now told that they are studying proteomics, as there is a tendency to turn off at the mere mention of biochemistry.

Biochemistry and bioinformatics

So, is proteomics simply biochemistry? The answer is a resounding no! Proteomics is a mixture of traditional biochemistry and the use of web-based bioinformatics tools. It enables us to understand and visualise protein structure, function, interaction and expression. This is really the traditional biochemistry aspect. However, using bioinformatics, it provides the opportunity to predict the biochemical properties, structure and function of proteins that have been derived from translation of the cloned nucleotides or polymerase chain reaction (PCR) products. Indeed, the output of most genome projects has been random gene sequences, the function and identity of which have been derived using bioinformatics tools.

For example, logging on to the Genome Japan website (www.genome.ad.jp) and opening the Basic Local Alignment Search Tool (BLAST) will permit comparison of an isolated small segment of DNA or a protein sequence with every one of the billions of known sequences lodged in databases around the world. What's more, the results are available in minutes! Analysis of homology with other proteins then helps to identify the protein.

Go to the ExPasy Proteomics Server (<http://us.expasy.org/>), type in the name of a protein and it is possible to access the TrEBML data set for that protein. Here, it is possible to find links to the protein sequence, the nucleotide sequences (with any known variants), its function, subcellular location, tissue specificity, polymorphisms, disease associations, similarities to other proteins, and, for some proteins, a specific website (eg www.albumin.org). Scrolling down reveals access to projected two-dimensional (2D) polyacrylamide gel electrophoresis (PAGE) analysis, 3D structures, domain structure, post-translational modifications, the protein sequence, all known variants (with references), and the ability to calculate pI and mass.

Far from straightforward

It is not always straightforward, however. Some organisms have genes and proteins that show no significant homology to any other identified. In such cases it is possible to ask questions via web-based tools. For example, is the sequence likely to be membrane-bound? Does it have motifs like any other type of protein? With some proteins, however, properties and function will only be known when they are purified and assayed by traditional technology.

So, what else is available? If you teach, the Kyoto Encyclopedia of Genes and Genomes (KEGG), which is available at Genome Japan, gives links to the KEGG protein network, and, in turn links to

metabolic pathways, regulatory pathways, molecular complexes, network-network relations, network-environment relations and diseases. The metabolic pathway pages are fascinating and allow comparison of generic pathways with those in your organism of choice, highlighting the differences.

A visit to the Human Protein Reference Database (www.hprd.org/) ultimately will provide information on every known human protein, protein localisation and tissue-specific expression. This site is still collating information, but it is an impressive resource, nonetheless. Protein microarray databases are now available at the European Bioinformatics Institute site (www.ebi.ac.uk/Databases/microarray.html), where it is possible to visualise some of the many microarray datasets available. There are also pages with links to all the important databases (eg www.biol.rug.nl/mbp/ListDatabases.htm#PD).

Microarray technology

Proteomics and bioinformatics are being used extensively by pharmaceutical companies interested in drug targets and modelling potential protein-drug interaction. Microarray technology is proving to be a fast primary screen for drugs and ligands. In vaccinology, potential targets are being found using a combination of traditional biochemistry and proteomics. Modelling studies allow drug and ligand refinements to be made, in order to improve binding characteristics.

'Proteomics is a mixture of traditional biochemistry and the use of web-based bioinformatics tools'

'Proteomics will enhance conventional science, providing rapid assessments, screening and new technologies

Protein microarrays are also finding a place in the analysis of antibodies in, for example, the understanding of antibody-antigen binding, the analysis of antibody variants to determine and improve specificity, and in autoimmune disease for the identification of antigen targets. In microbiology, comparative genome analysis (www.webact.org/) makes it possible to visualise potential areas associated with pathogenicity by comparing pathogenic and non-pathogenic strains.

In the hunt for surrogate disease markers of cancer, a range of traditional biochemical techniques, such as chromatography and 2D PAGE, makes it possible to discover, for example, whether the production of individual proteins is up- or down-regulated, or whether or not some proteins are expressed after mutation.

Time to explore

So, will proteomics replace 'wet' science? The answer must be no, as the databases referred to above rely on information produced by the traditional scientific approach. There are also drawbacks to predicting structure, pI and mass. For example, if a protein is heavily glycosylated, the real values will not reflect those obtained for the simple polypeptide chain. However, proteomics will enhance conventional science, providing rapid assessments, screening and new technologies. Clearly, now is the time to start exploring! Those wishing to download links, exercises or ask questions should go to www.mydocsonline.com/pub/greenwp/proteomics.

Pamela Greenwell and Sanjiv Rughooputh are members of the Molecular and Medical Microbiology Research Group at the School of Biosciences, University of Westminster, London.