Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 65 (2004) 1443–1447
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Editorial

Proteomics: empowering systems biology in plants

Developing technologies being applied in the post-genomics era are not only allowing determination of global changes in plant gene expression but also the level and complexity of proteins and the resultant metabolite profile. The integration of such data forms the basis for modern systems biology, which aims to establish and model networks that are fundamental to understanding the evolution, development and adaptability of organisms. In plant breeding and biotechnology too, such data is serving to identify pleiotropic effects as a result of mutation and transgenesis allowing more effective rational improvement.

One major contribution of proteomics to this new science will serve to improve the genomic data. Some of the present questions associated with genome annotations surround the fundamental question of how many genes there are, which, a priori, translates into how many functional proteins. Annotations of genomes depend somewhat on the algorithms applied to identify genes and these parameters include the minimum exon size set and the identification of bona fide promoters and of pseudogenes. Proteomics has a service to play here, as it certainly has in correction of miss-annotations and sequencing errors. Programmes exist such as Arabidopsis 2010 in the US, GarNet in the UK and various European wide consortia to take this further and confirm functional identity of all plant genes in the existing genomes. This will obviously involve proteomics and metabolomic inputs. Phytochemistry wishes to reflect these technological innovations and the resultant explosion in knowledge in a series of special issues. A Metabolomics issue edited by Dieter Strack and Richard Dixon has already appeared (volume 62(6) March 2003), the present issue, in two parts, deals with Proteomics and an issue on Evolution of Metabolic Diversity, a genomics issue dealing with secondary metabolism is planned for 2005 to be edited by Dieter Strack, Toni Kutchan and Thomas Hartmann.

The proteomes that have been appearing, already support the contention that the number of proteins will amplify, functionally, the number of genes. In humans, for instance, it is proposed that 30,000 genes translate to at least 90,000 proteins. This added diversity arises through alternative splicing and protein post-translational modification. The latter involves processing cleavages, protein splicing, glycosylation, phosphorylation, lipid modifications and at least a hundred other possible modifications. Although the general consequences of many of these modifications are known, their temporal and spatial manifestations and the signalling bringing them about requires proteomic analysis, using available and developing technologies many of which are highlighted in this issue.

Added complexity derived from proteome analysis can be seen at the tissue level, resulting from differential gene expression. This is spectacularly exemplified by Koller et al. (2002) who carried out ground breaking MudPIT/2D-gel analysis on rice and which was reviewed by one of us (Whitelegge, 2002). They analysed leaf, root and seed and identified 2528 unique proteins. Unexpectedly, only 189 of these were common to all three tissues, therefore revealing a large degree of specialization. In terms of phytochemistry, this probably means that in plants central metabolic pathways do not require large genetic commitment and there is high potential for intermediary and secondary pathways in the individual tissues. The limitations of genomic analysis have also been revealed in the proteomics of subcellular compartments. The two other editors had such an experience in their analysis of cell walls (Chivasa et al., 2002). Three hundred and twenty-eight spots were identified and analysed, giving 111 distinct proteins, representing 69 different gene products. However, 50% of these were not annotated in the genome and some of the known proteins had not been not described previously in walls. Such findings are further borne out by several of the papers within this special issue.

The proteome, unlike the genome, is dynamic and because of the need to measure changes in expression and post-translational modification in response to various over what can be short timescales, newer advances have addressed the need for accurate quantitation to allow meaningful comparisons. As a consequence of variability between measurements, current strategies in quantitation typically involve mixing of an experimental sample with a control, each coded with stable isotopes to allow for mass spectrometric discrimination between two samples in a single measurement. Labelling is achieved via growth in modified media or by chemical modification in vitro (Tao and Aebersold, 2003). The

0031-9422/$ - see front matter
doi:10.1016/j.phytochem.2004.06.001

paper by Whitelegge et al. in the first part of this issue introduces the use of minimal alteration of isotope ratio for isotope coding in proteomics. To complete the equation for proteome flux measurements it is necessary to measure protein turnover rate. Stable isotope strategies are again employed in pulse-chase protocols revealing, in the case of yeast, a wide variation in protein turnover rates (Pratt et al., 2002).

Another area where development of proteomic technologies are required is signal transduction. Protein kinase gene families have been annotated and the drive is now on to place individual members into their respective regulatory circuits. The transient expression system developed by Jen Sheen (Sheen, 2001) can address the biology in a spectacular breakthrough technology. However, the paper by Sawasaki et al. in this issue describes another emerging technology utilising in vitro transcription-translation and phosphorylation for further characterization of kinases and points to development of a protein chip. Although not represented in this issue, the area of phosphorylation needs development of high throughput technology for identification of target proteins and their respective phosphorylation sites. Despite claims of the manufacturers of the hardware for the ability of their equipment, this technology is not yet routine.

This special issue has been compiled in several thematic sections arranged over two parts (issues 65/11 and 65/12). The first part begins with a number of general contributions. Russell Newton et al. thoroughly review modern biological mass spectroscopy and its application to proteomics and plant biochemistry. Hisashi Hirano et al. review the technical aspects of functional genomics in plants. They consider all methodology from 2D electrophoresis, multi-dimensional protein identification technology (MudPIT), through quantitative techniques using fluorescent dyes (2D-DIGE) and isotope labelled affinity tags (ICAT) to the latest technologies using top-down protein analyses using Fourier transform ion cyclotron resonance. The review delves further into strategies for examining protein–protein interactions and future technical demands. The reference list is truly spectacular and will serve the community for some time to come. Francis Chevalier et al. examine current stains for polyacrylamide gels providing the first comparisons of some of the newest technology available. Whitelegge et al. explore the use of minor alterations to ^{13}C:^{12}C ratios for isotope coding in proteomics with the goal of providing an inexpensive alternative to current strategies requiring full isotope exchange and full compatibility with available protein identification technologies. Doubling the abundance of ^{13}C provides isotope coding without compromising protein identification by tandem mass spectrometry.

Following these general technological contributions are a number on functional genomics. There are two reviews by Millar et al. and Gottlieb et al. The paper from the Millar group reviews the techniques that are employed by plant proteomics researchers to link proteins to others on interaction maps. The problem of redundancy arising from gene duplication is emphasized and is a recurring theme throughout the work. Gottlieb et al. present a multivariate approach to the handling of data in classical proteomics and other data generating protein analysis methods like mass spectrometry and near infrared spectroscopy. This has the advantage that protein patterns can be identified at an early stage and consequently the proteins selected for sequencing can be selected intelligently and reveal the hidden structures present in these data. Tatsuya Sawasaki et al. report on a high efficiency in vitro cell free translational system from wheat germ which they have modified so that it works efficiently and gives high yield. This is a bonus for those looking at gene function as it has allowed them to investigate 439 cDNAs encoding protein kinases from Arabidopsis and demonstrate autophosphorylation and Ca^{2+}-dependent kinase activity. Norman Lewis's group present the utility of expression systems to study the detailed kinetics of individual isoforms within multigene families, i.e. the kinetic characterization of the four isoforms of the Arabidopsis phenylalanine ammonia lyase gene family. A contribution from Hans-Peter Mock's Group describes use of 2D-electrophoresis, MALDI TOF, post-source decay and western blotting to look at germins in two ecotypes of Arabidopsis. This is a multigene family which posess at least one site for glycosylation and one of the isoforms has oxalate oxidase activity which could be involved in cell wall cross-linking. Zhang et al. report on the growing use of mass spectroscopy to allow detailed structural characterization of complex glycoproteins and improve on our knowledge of how glycans influence stability and folding of proteins. Eric Sarnighausen et al. describe the first extensive study on proteomics in *Physcomitrella patens*. This is a plant, which has a high frequency of homologous recombination allowing for targetted gene deletions. Over 300 proteins have been identified and this will be a landmark paper in the area.

The application of proteomics in eco-devo research is well represented also. The group of Catherine Damerval has produced a 2D map of maize endosperm in an extensive study, which will be a resource of fundamental importance for all working in cereals. Christine Finnie et al. have looked at barley seed proteins with particular emphasis on effects of growth conditions, transgenes and pleiotropic effects. Bob Buchanan's group provide an informative description and critical assessment of their approach to identify thioredoxin protein targets in wheat seeds. Using monobromobimane fluorescence and affinity column chromatography, 68 putative protein thioredoxin targets were identified in wheat endosperm and flour. Significantly, 40 potential new targets were

identified for thioredoxin. The contribution from Wolfram Weckworth's group reports on a proteomic analysis of a single cell type, the *Arabidopsis* trichome. The authors collected thousands of individual trichomes and identified 69 proteins whose functions indicate high levels of sulphur metabolism and detoxification. Eliane Dumas-Gaudot et al. report the effect of xenobiotics on the interactions between the model plant *Medicago truncatula* and the arbuscular mycorrhizal fungus *Glomus mosseae* or the rhizobial bacteria *Sinorhizobium meliloti*. Symbiosis-related proteins were detected and identified by two-dimensional electrophoresis and matrix-assisted laser desorption ionization mass spectrometry, and image analysis was used to determine the effects on the *M. truncatula* symbiotic proteome.

The second part begins with a section on organelles. Application of proteomics at the subcellular level is also well represented through six contributions. Starting with the nucleus, Md. Khan et al. review the state of play with respect to proteomics of rice tissues and organelles with particular emphasis on the annotation of the nuclear localized proteins. More than 190 proteins are now identified, the highest subset of which, naturally are involved in signal transduction and gene regulation. The contribution from Hans-Peter Braun's demonstrates the utility of blue-native gels as a first dimension for proteomic analyses of supercomplexes from solubilized thylakoid membranes from *Arabidopsis*. This separation system sets the bar for all other experimental approaches aimed at subfractionation of cellular membrane systems. Sabine Brugerie et al. have increased by 40% the number of identifiable proteins now known to be present in *Arabidopsis* mitochondrial membranes. Although many more remain to be identified, this study must represent one of the most complete for an individual membrane compartment from plant cells. Bonnie Watson et al. have used proteomics to investigate the cell wall of *M. sativa* (alfalfa). Differential extraction of the cell wall has been achieved using different salt solutions and the high LiCl treatment has provided a separate fraction for analysis. Over 100 proteins have been identified and their subcellular location by this method forms the basis for further studies. The related *M. truncatula* has become the leguminous model plant of choice for which abundant EST and genomic data is accumulating. Beniot Valot et al. have made one of the first major proteomic studies to aid annotation and report 96 microsomal proteins some of which could be important in establishing symbiotic relationships. Brian Mooney and Jay Thelen demonstrate the utility of the Unigene expressed tag database for protein identification in soybean.

Proteomics has also been applied successfully to understanding signalling components. A review from Reddy and Reddy shows that the Ca^{2+}-signaling circuit consists of three major "nodes" – generation of a Ca^{2+}-signature in response to a signal, recognition of signature by Ca^{2+} sensors and transduction of signature message to targets that participate in producing signal-specific responses. Molecular genetic and protein–protein interaction approaches together with bioinformatic analysis of the Arabidopsis genome have resulted in identification of a large number of proteins at each "node" – ~80 at Ca^{2+} signature, ~400 sensors and ~200 targets – that form a myriad of Ca^{2+} signalling networks in a "mix and match" fashion. They discuss the current status of Ca^{2+} signalling components, their known functions and potential of emerging high-throughput genomic and proteomic technologies in unravelling complex Ca^{2+} circuitry. Although the kinase families, CDPKs, MAPKs, etc. are well annotated, there is a vast lack of knowledge of individual targets for each kinase. Birgit Kirsten's group have used such high-throughput technology and generated protein microarrays that were incubated with an expression library-derived barley CK2α in the presence of [γ-^{33}P]ATP, and signals were detected by X-ray film or phosphor imager. They were able to demonstrate the power of the protein microarray technology by identification of 21 potential targets out of 768 proteins including such well-known substrates of CK2α as high mobility group proteins and calreticulin. A contribution from Gerhard Link's group explores this phylogenetically conserved kinase (CK2) and shows, in mustard, it is both phosphorylated and subject to thiol regulation for hypothetical transcriptional control in the chloroplast.

Due to the massive changes in global gene expression that occur, it is not surprising that a proteomics approach to understanding biotic stress has been particularly fruitful. Christina Walz et al. have applied proteomics to look at cucurbit phloem exudates so revealing a defensive network. The system is of particular interest as the proteins come from phloem, which has no endogenous transcriptional or translational machinery. The majority of the proteins identified were involved in responses to stress and in defence reactions. An interaction, which involves parasitizing of the plant vascular system, is described by the group of Jesus Jorrin. They compared resistant and susceptible accessions of pea to broomrape and have achieved positive identification of a number of proteins of carbohydrate and nitrogen metabolism in addition to those that are defence related that show differential responses. Murray Grant and John Mansfield discuss responses to microbial pathogens and report alterations in the proteome of *Arabidopsis thaliana* leaves during early responses to challenge by *Pseudomonas syringae* pv. *tomato* DC3000 (DC3000). Protein changes characteristic of the establishment of basal resistance and *R*-gene mediated resistance were examined by comparing responses to DC3000, a *hrp* mutant and DC3000 expressing *avrRpm1*, respectively. The abundance of selected transcripts was also analysed

in GeneChip experiments. They present comparative data from the soluble fraction of leaf protein, highlighting changes in two antioxidant enzyme groups; the glutathione *S*-transferases and peroxiredoxins and stressing the importance of oxidative events in innate defence (basal resistance) and gene for gene interactions. A feature of the latter is the hypersensitive response although there is debate whether this involves true programmed cell death (PCD).

Programmed cell death also occurs in response to abiotic stresses. Lee Sweetlove and colleagues report a proteomics approach to this in general which will help define the exact parameters of this phenomenon in plants. They identified 11 proteins that increased in abundance relative to total protein in response to heat- and senescence-induced PCD despite extensive degradation of other proteins in cultured cells of *Arabidopsis*. The increased abundance of several antioxidant proteins as well as a measured increase in free Fe^{2+} content of the cells indicates an oxidative stress in this system. Several mitochondrial proteins were identified, confirming the importance of this organelle during PCD. They also identified an extracellular glycoprotein that may function in the transmission of a 'death signal' from cell to cell. Kristensen et al. devised an elegant protocol for proteomic analysis of protein carbonyl formation in rice leaf mitochondrial matrix following exposure to mild oxidative stress in vitro. Protein oxidation in vivo is an aspect of proteomics that is only just being addressed. Bart Devreese's group also study the heat shock response. In this study, a proteomic analysis (2D-PAGE, MS) was used to detect the effects of heat shock on the protein pattern of barley cultivars tolerant and susceptible to abiotic stresses. They show the power of an automated nano-HPLC system coupled to an electrospray ionization-quadrupole linear ion trap (Q-TRAP) instrument which allowed the identification of six isoforms of a 16.9 kDa small heat shock proteins (sHSP) in one single spot. Other differentially expressed proteins were *S*-adenosylmethionine synthetase (SAM-S) and distinct isoforms of sHSP.

In conclusion, we hope we have gathered together a group of contributions that manifest the growing impact of proteomics on modern plant science. Clearly still in its infancy, proteomics will yield significant insights on an ever broadening basis as we seek to understand the relationship between the genome and the living plant system. Ultimately, this understanding will yield acceptable mathematical models of plant metabolism and development enabling us to predict the results of targeted engineering prior to experimentation. The scope for innovative improvements in proteomics performance over the next 10 years is huge; young researchers should not underestimate the contribution they might make.

References

Chivasa, S., Ndimba, B.K., Simon, W.J., Robertson, D., Yu, X.L., Knox, J.P., Bolwell, P., Slabas, A.R., 2002. Proteomic analysis of the *Arabidopsis thaliana* cell wall. Electrophoresis 23, 1754–1765.

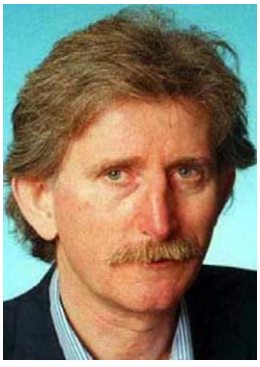
Koller, A., Washburn, M.P., Lange, B.M., Andon, N.L., Deciu, C., Haynes, P.A., Hays, L., Schieltz, D., Ulaszek, R., Wei, J., Wolters, D., Yates 3rd, J.R., 2002. Proteomic survey of metabolic pathways in rice. Proceedings of the National Academy of Sciences of the United States of America 99, 11969–11974.

Pratt, J.M., Petty, J., Riba-Garcia, I., Robertson, D.H., Gaskell, S.J., Oliver, S.G., Beynon, R.J., 2002. Dynamics of protein turnover, missing dimension in proteomics. Molecular and Cellular Proteomics 8, 579–591.

Sheen, J., 2001. Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. Plant Physiology 127, 1466–1475.

Tao, W.A., Aebersold, R., 2003. Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. Current Opinion in Biotechnology 14, 110–118.

Whitelegge, J.P., 2002. Plant proteomics: BLASTing out of a MudPIT. Proceedings of the National Academy of Sciences of the United States of America 99, 11564–11566.



Paul Bolwell is the UK-based editor of Phytochemistry. He moved to Cambridge University for a year after graduating in Botany from Oxford University and in 1971 returned to Oxford to complete a D. Phil in Plant Biochemistry in 1974. Subsequent research posts at Leeds and Cambridge University were followed by an industrially sponsored research fellowship at Royal Holloway, University of London. From 1986 to 1989 he was a senior lecturer in Molecular Biology at the City of London Polytechnic and returned to Royal Holloway where he was awarded a personal chair in Plant Biochemistry in 1998. In 1998–1999 he was a visiting Professor at Harvard University. He joined the Board of Phytochemistry in 1995 and became UK-based editor in 1999. Over the past five years, he has worked with the other editors to maintain the position of Phytochemistry as a forum for reporting all aspects of pure and applied plant biochemistry leading to a deeper understanding of the factors underlying the growth, development and metabolism of plants and the chemistry of plant constituents. This has involved encouraging the applications of post-genomic technologies, exemplified in the present special issue.

Toni Slabas graduated from Queen Mary College, London and obtained his D. Phil from Oxford University in 1975. Following research at Sheffield University and University College London he joined Unilever Research, Colworth House. He established the Protein Chemistry Section and was section manager of Cell Sciences. He returned to academia as Professor of Plant Biochemistry in Durham University in 1988. He has broad research interests. In the area of lipid metabolism he is interested in the biochemical pathway of lipid metabolism and its regulation. This metabolic pathway involves de novo fatty acid biosynthesis and incorporation of fatty acids into triacyglycerols (TAGs) using acyltransferases. The molecular structure of several of these components has been determined and

these studies will be extended to elucidate protein–protein interactions between components of lipid metabolism. Using anti-sense plants and other approaches, he is investigating regulation of lipid metabolism. A major programme has also been initiated in proteomics. This includes analysis of the endoplasmic reticulum (ER) and its role in triacylglycerol biosynthesis, use of Arabidopsis suspension cultures to identify components in biotic and abiotic stress responses, cyanobacterial adaptation to stress and recently androgen signaling in prostate cancer cell lines.

Julian P. Whitelegge was trained in the biological sciences, studying marine and freshwater biology at Westfield College prior to its merger with Queen Mary, and plant physiology at Imperial College in the group of James Barber FRSC, Ernst Chain Professor of Biochemistry. After a postdoctoral position at Royal Holloway, London, he moved to Los Angeles to continue studies of covalent modification of the integral membrane proteins that constitute the reaction center of Photosystem II. Joining The Pasarow Mass Spectrometry Laboratory in 1995, Julian was well placed to exploit the growing field of biological mass spectrometry and expanded his interests into the neurosciences and other medically important areas. Over the last few years, he has continued to pioneer the application of mass spectrometry to integral membrane proteins and tackle the problem of applying proteomics to the bilayer domain. Julian has organized the 'Membrane Protein' session of the annual meeting of the American Society for Mass Spectrometry since its inception in 2002.

G. Paul Bolwell
Division of Biochemistry
School of Biological Sciences
University of London Division of Biochemistry
329b Bourne Building, Egham
Surrey TW20 0EX, UK
E-mail addres: phytochemistry@rhul.ac.uk

Antoni R. Slabas
Department of Biological Sciences
Durham University
Durham DH1 3LE, UK
E-mail address: a.r.slabas@durham.ac.uk

Julian P. Whitelegge
Department of Psychiatry and Biobehavioral Sciences
David Geffen School of Medicine
University of California
Los Angeles, CA, USA
E-mail address: jpw@chem.ucla.edu