Introductory Review

99. Separation-dependent approaches for protein expression profiling

Michael J. Dunn and Stephen R. Pennington
Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland

REVISED PAGE PROOFS

1. Introduction

Intense efforts over the last few years have resulted in the availability, at the time of writing (February 2005), of complete genome sequences for 256 organisms (21 archael, 203 bacterial, 32 eukaryotic), including man. This wealth of information is an invaluable resource that will allow comprehensive studies of gene expression that will in turn lead to new insights into cellular functions that determine biologically relevant phenotypes in health and disease. The understanding that one gene can encode more than a single protein has led to a realization that the functional complexity of an organism far exceeds that indicated by its genome sequence alone. While powerful techniques such as DNA microarrays and serial analysis of gene expression (SAGE) make it possible to undertake rapid, global transcriptomic profiling of mRNA expression, processes including alternative mRNA splicing, RNA editing, and co- and posttranslational protein modification make it essential to undertake expression studies at the protein level. The concept of mapping the human complement of protein expression was first proposed more than 25 years ago (Anderson and Anderson, 1982), with the development of a technique in which large numbers of proteins could be separated simultaneously by two-dimensional polyacrylamide gel electrophoresis (2-DE) (O'Farrell, 1975). The term "proteome" was not established until the mid-1990s (Wasinger *et al*., 1995) when it was proposed to define the protein complement of a genome. This article will give an introduction to expression profiling in which the individual proteins in a complex sample are separated prior to semiquantitative analysis, and then identified usually using techniques of mass spectrometry. For an introduction to alternative strategies for expression profiling of unresolved complex protein samples, see article 99, **Separation-independent approaches for protein expression profiling**, Volume 0.

g302110

2. Two-dimensional gel electrophoresis (2-DE)

The technique of two-dimensional gel electrophoresis (2-DE) in which proteins are separated in the first dimension according to their charge properties (isoelectric point, p*I*) under denaturing conditions, followed by their separation in the second dimension according to their relative molecular mass (M_r) by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was developed more than 25 years ago (O'Farrell, 1975). Nevertheless, it remains the core technology of choice for the majority of applied proteomic projects (Görg *et al*., 2004; *see* article 99, **Two-dimensional gel electrophoresis (2-DE)**, Volume 0, article 99, **Two-dimensional gel electrophoresis**, Volume 0) due to its ability to separate simultaneously thousands of proteins and to indicate posttranslational modifications that result in alterations in protein p*I* and/or M_r. Large-format (24 × 21 cm) 2-D gels can routinely separate around 2000 protein spots. Moreover, recent developments including the use of narrow range "zoom" gels (*see* article 99, **Two-dimensional gel electrophoresis (2-DE)**, Volume 0, article 99, **Two-dimensional gel electrophoresis**, Volume 0) and fluorescent dyes that facilitate the multiplex analysis of samples (*see* article 99, **2D DIGE**, Volume 0, article 99, **2-D Difference Gel Electrophoresis – an accurate quantitative method for protein analysis**, Volume 0) make it possible to achieve greater proteomic coverage combined with more accurate differential expression analysis. Additional advantages of 2-DE are the high-sensitivity visualization of the resulting 2-D separations (*see* article 99, **Detecting protein posttranslational modifications using small molecule probes and multiwavelength imaging devices**, Volume 0), compatibility with quantitative computer analysis to detect differentially regulated proteins (Dowsey *et al*., 2003; *see* article 99, **Image analysis**, Volume 0), and the relative ease with which proteins from 2-D gels can be identified and characterized by mass spectrometry (*see* article 99, **MS-based methods for identification of 2-DE-resolved proteins**, Volume 0).

g302406 g302207 g302406 g302207 g302312 g302404 g302313 g302305 g302409

3. Alternatives to 2-DE

Despite the many advantages of 2-DE, there are alternative protein separation strategies. Perhaps the simplest alternative to 2-DE is the use of one-dimensional SDS-PAGE to separate proteins in the sample, on the basis of their M_r followed by protein identification by tandem mass spectrometry (MS/MS), such that several proteins comigrating in a single band can be identified. This method is limited by the complexity of the protein mixture that can be analyzed but is well suited for the analysis of membrane proteins, and has also been successfully applied to the study of protein complexes (Figeys *et al*., 2001). Other approaches avoid the use of gels altogether by combining liquid chromatography (LC) and MS. In these so-called shotgun approaches, a tryptic digest of the sample is separated by one or more dimensions (typically ion-exchange combined with reverse-phase) of LC to reduce the complexity of peptide fractions. These are subsequently introduced (either on- or off-line) into a tandem mass spectrometer for sequence-based identification. For example, the so-called MudPIT approach (Wolters *et al*., 2001) identified around

REVISED PAGE PROOFS

1500 yeast proteins in a single analysis (Washburn *et al*., 2001). An alternative to this approach that is more robust than multidimensional chromatography, while still allowing complex samples to be analyzed, has been termed *GeLC-MS/MS* (Schirle *et al*., 2003). Here, tryptic digests of protein bands excised from the SDS-PAGE gel are separated by one-dimensional RP-HPLC prior to on- or off-line MS/MS analysis. However, a major limitation of such approaches is that unless combined with some form of stable isotope labeling or "mass tagging", they provide no information on semiquantitative abundance of proteins and are very limited in their ability to detect posttranslational modifications.

The former problem is currently being addressed by the development of a range of MS-based techniques in which stable isotopes are used to differentiate between two or more populations of proteins (Han *et al*., 2001). In general, this approach consists of four steps: (1) differential isotopic labeling of the two (or more) protein mixtures, (2) digestion of the combined labeled samples with a protease such as trypsin or Lys-C, (3) separation of the peptides by multidimensional LC, and (4) semiquantitative analysis and identification of the peptides by MS/MS. Currently, the most widely used method is the isotope-coded affinity tag (ICAT) (Han *et al*., 2001; *see* article 99, **ICAT and other labeling strategies for semiquantitative LC-based expression profiling**, Volume 0), but there are a variety of other approaches involving labeling with stable isotopes at the whole cell, intact protein, or tryptic peptide level (Julka and Regnier, 2004; Ross *et al*., 2004; *see* article 99, **ICAT and other labeling strategies for semiquantitative LC-based expression profiling**, Volume 0). Although these approaches are promising, there are caveats: (1) their quantitative reproducibility needs to be established, (2) the dynamic range of the these techniques may be little better than 2-DE, and (3) there is evidence that they can be complementary to a 2-DE approach in identifying a different subset of proteins from a given sample (Kubota *et al*., 2003).

g302201

g302201

4. Conclusion

The current array of proteomic techniques makes it possible to characterize global alterations in protein expression associated with the progression of many different biological processes, including human disease. However, there is still no one method that is suitable for the analysis of all samples, and for many projects it is likely that a combination of proteomic platforms, both gel and nongel based, will have to be applied to provide the required depth of proteomic coverage.

Acknowledgments

MJD is the recipient of a Science Foundation Ireland Research Professorship, and is grateful to SFI for the generous support of research in his laboratory. The Proteome Research Centre established by SRP has been supported by grants from the Higher Education Authority in Ireland.

REVISED PAGE PROOFS

References

Anderson NG and Anderson L (1982) The human protein index. *Clinical Chemistry*, **28**, 739–748.

Dowsey AW, Dunn MJ and Yang GZ (2003) The role of bioinformatics in two-dimensional gel electrophoresis. *Proteomics*, **3**, 1567–1596.

Figeys D, McBroom LD and Moran MF (2001) Mass spectrometry for the study of protein-protein interactions. *Methods*, **24**, 230–239.

Görg A, Weiss W and Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics*, **4**, 3665–3685.

Han DK, Eng J, Zhou H and Aebersold R (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nature Biotechnology*, **19**, 946–951.

Julka S and Regnier F (2004) Quantification in proteomics through stable isotope coding: a review. *Journal of Proteome Research*, **3**, 350–363.

Kubota K, Wakabayashi K and Matsuoka T (2003) Proteome analysis of secreted proteins during osteoclast differentiation using two different methods: two-dimensional electrophoresis and isotope-coded affinity tags analysis with two-dimensional chromatography. *Proteomics*, **3**, 616–626.

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *The Journal of Biological Chemistry*, **250**, 4007–4021.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, *et al*. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics*, **3**, 1154–1169.

Schirle M, Heurtier MA and Kuster B (2003) Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Molecular & Cellular Proteomics*, **2**, 1297–1305.

Washburn MP, Wolters D and Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology*, **19**, 242–247.

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL and Humphery-Smith I (1995) Progress with gene-product mapping of the mollicutes: mycoplasma genitalium. *Electrophoresis*, **16**, 1090–1094.

Wolters DA, Washburn MP and Yates JR (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry*, **73**, 5683–5690.

REVISED PAGE PROOFS