TWO-DIMENSIONAL DIFFERENTIAL IN GEL ELECTROPHORESIS (DIGE): A NOVEL METHOD FOR HIGH THROUGHPUT PROTEOMICS

Belinda Attard[1], John Aitken[1,2] and Mark Baker[2]
[1]Mothers and Babies Research Center, Hunter Medical Research Institute, and
[2]The ARC Centre of Excellence in Biotechnology and Development, School of Environmental and Life Science, University of Newcastle, NSW 2308

Introduction

With completion of the human genome project, the post-genomic era is upon us, and rapid advances are being made in understanding the functional diversity of protein interactions. Although identifying the complete proteome within tissue samples or organelles (1) offers a vast wealth of resources to researchers, this approach does not take into account what we would term "functional proteomics" - that is, those proteins that undergo modification, from one said state to another, in order to produce a phenotypic change.

Difference in Two-Dimensional Gel Electrophoresis (DIGE) is a powerful tool employed in monitoring the complex differences in proteomic profile between cells in different functional states. This technology allows the inter-comparison of up to three different samples within the one two dimensional polyacrylamide gel (2D-PAGE). When DIGE is combined with mass spectrometry protein identification, this system has the potential for expeditious separation and identification of proteins undergoing relevant changes, allowing comprehensive studies into protein expression in the context of development, pathology and experimental manipulation.

DIGE vs. Traditional Comparative 2D-PAGE

The most powerful technique currently available to resolve thousands of proteins in a single run is 2D-PAGE (2). With this in mind, researchers have sought to use such technology and compare 2D-PAGE of different samples in different gels. Samples are often separated in multiple gels and then overlapped in order to locate different spots or intensity of spot changes from one gel to another. Often discrepancies between gels in spatial resolution and spot intensities makes the overlaying of images and correct matching of proteins difficult, thus making it impossible to distinguish biological from experimental variation. DIGE builds on traditional 2D-PAGE by pre-labelling protein mixtures with cyanine dyes that guarantee co-migration of proteins, prior to electrophoresis (3).

Dye Development

The cyanine dyes (Cy2, Cy3 and Cy5) used for DIGE are all charge-matched and molecular mass-matched to prevent alterations of pI, and minimise dye-induced shifting of labelled proteins during electrophoresis. The dyes are N-hydroxy succinimidyl ester derivatives and have been designed to undergo a specific chemical reaction with the ε-amino group of lysine. The nucleophilic substitution of lysine removes the ε-amino group positive charge and replaces it with the positive charge of the dye itself, ensuring no change in protein electronegativity occurs (4,5). Lysine is targeted for cyanine labelling as it is found at high levels in a number of different proteomes studied (6). Due to the addition of ~0.5 kDa, the cyanine dyes add to the total molecular mass of a labelled protein such that the latter undergo a small but typical shift. However, binding of the dye to the protein appears to have no effect in protein identification through mass spectrometry techniques (4,7). This is due to the concept of minimal labelling, whereby the protein dye ratio is minimised ensuring that only one lysine residue per protein is labelled (4,8,9). This strategy of minimally labelling lysine residues in itself may present a problem, since the abundance of spots labelled with either Cy dye could be altered in relation to the number of lysine residues present, thereby representing a protein's abundance in a deceptive manner while proteins without a lysine residue would not be labelled at all (7). Nevertheless the approach of minimal labelling is ultimately a necessity since the hydrophobicity of the dyes is comparatively high and can cause protein precipitation at high dye:protein ratios (3).

A new strategy has recently emerged using new cyanine dyes that exhibit properties similar to the previously described dyes, however, instead of minimally labelling lysine residues they saturate cysteine residues. Again the cyanine dyes used are mass and charged matched; however, the saturation cyanine dyes are maleimide fluors that react with thiols, found on cysteine, with a nucleophilic addition to form a thioether. The reaction requires a pH range of 6.5-7.5, at which pH maleimides react faster with thiols than with amines, therefore they will preferentially label cysteine, however, there is a smaller chance of protein modification and pI shift (9). A number of advantages come with this new strategy of labelling. Photobleaching is lower using both minimal and saturation Cy3 and Cy5 compared to a gel stained with Sypro Ruby (8) over a four hour time period. Saturation cysteine dyes were found to possess superior sensitivity, efficient enough to detect 0.1 ng of albumin compared to 1 ng by minimal lysine labelling, and almost double the number of spots detected using saturation Cy5 compared with minimal Cy5 (9). As with the minimal lysine labelling cyanine dyes, if a cysteine amino acid is not present in a protein then the saturation dyes will not be able to label it, leaving the protein undectatable. Other experiments, however, have

revealed that the saturation cyanine dyes will detect small concentrations (5-10 ng) of myoglobin, which contains no cysteine residues (9). It is possible that in the absence of a cysteine residue the saturation cyanine dyes will label another amino acid, possibly lysine, in its stead. The potential of labelling without the desired amino acid present opens the possibility of identifying an entire stratum of proteins that would have previously escaped detection. Although they have many advantages, saturation cyanine dyes do come with an increased propensity for artefactual protein changes and a decrease in protein solubility at higher molecular masses.

The power in using cyanine dyes comes from both the high fluorescent extinction coefficient and the fact that each possesses unique absorption and emission spectra, thus, allowing three differentially labelled samples to be analysed within the one gel. Once appropriately labelled, the mixture of samples are isoelectrically focused into an IPG strip and run out on a gel, as with any normal 2D gel. Each dye is then scanned using a different emission filter and the band pass during scanning is kept to a minimum in order to prevent leaking between fluorescence labels (5), or imaged using a charge-coupled device camera (3,10). After the gels have been imaged and compiled they are analysed with DeCyder Differential In-Gel Analysis Software (Amersham Biosciences). This software allows accurate protein alignment and quantification between scanned images. Spots may be directly picked through an automated system off a previously scanned gel, or preferably, a preparative gel run and stained with either colloidal Coomassie, silver staining (4) or Sypro Ruby (8). These stains should be run in conjunction with Cy dye labelling, since when used alone they exhibit higher variability between gels and less sensitivity than Cy dyes (4,8).

Although the cyanine dyes are most commonly used for protein labelling in the DIGE method, there have been reports of the use of Alexa dyes. Von Eggeling and associates (11) have used the fluorescent dyes, Alexa 488 and 532, in DIGE analysis of proteins extracted from a colorectal adenocarcinoma cell line, Colo 320, spiked with BSA. Alexa dyes are similar to the cyanine dyes, as they are a succinimidyl ester derivative that covalently binds with the amines of lysine; however, the Alexa dyes used in the Von Eggeling experiment carry a negative (Alexa 488) or neutral charge (Alexa 532). Due to the discrepancy in charge carried by the Alexa dyes, the pI of proteins labelled will change according to the dye they are labelled with and as a consequence will migrate different distances on the IPG strip during the first dimension. Although Alexa dyes and radionuclotide labelling (12) have been employed in the past, these methods have proven to be inferior or flawed and should not be used as a substitution for cyanine dye labelling.

Refining Resolution and Quantification of DIGE

DIGE can be used in conjunction with other techniques to improve identification of proteins that would normally go without detection. Purification and pre-fractionation of complex protein samples would be advantageous, as less abundant proteins can be masked behind those present in higher concentration. This can be achieved through a number of methods that can purify the sample. Reverse Phase High Performance Liquid Chromatography (RP-HPLC) has been used to improve the resolving power of normal 2D electrophoresis (13) in the past. RP-HPLC has also been used in conjunction with DIGE to determine protein changes in the striated cortex of adult (14), and 30 day old kittens (10). Previous experiments had identified 28 protein differences between the adult and kitten samples without fractionation (15), however, with RP-HPLC pre-fractionation, a further 15 spots were resolved that had previously been concealed (10). Reducing the complexity of samples through RP-HPLC is a useful procedure to use in conjunction with DIGE as it improves the resolving power of DIGE without compromising its quantitative capabilities (10).

Standard 2D gels separate over a typical pH range of 3-10 in the first dimension, giving a broad survey of all proteins present, however, this approach often does not detect low abundance proteins. By concentrating on a smaller pI range over the same length of IPG strip, it is possible to detect proteins that were previously not visualised; these are known as zoom gels (16,17). Zoom gels have been successfully used in both 2D (16,17) and DIGE (15) analysis to improve resolution of unseen proteins or to further separate out multiple proteins identified by mass spectrometry from a single spot (17). Although zoom gels and pre-fractionation improve resolution, the accurate quantification of protein variation between samples is imperative for determining biological significance.

Standard DIGE encompasses a simple strategy involving direct comparison of two different samples. Although this technique has been effectively used to investigate a number of proteomes (3,4,7,18,19) it is a restrictive methodology and does not allow the freedom necessary for a multi-sample study, without requiring multiple gels and the variation associated therein. In order to efficiently expand beyond a simple picture of direct two-protein-sample comparisons into multiple sample analysis, the introduction of a third dye, Cy2, was of great significance. An equal quantity of each sample to be analysed is labeled with Cy2 and this allows for an internal control to be added to each gel run (**Fig. 1**). The internal control is a pooled mixture and represents the average of all samples and proteins present. This allows for normalisation of protein abundance measurements across multiple gels.

By labelling a pooled internal control with Cy2 and the experimental and control samples with Cy3 and Cy5, it was found that the addition of the Cy2-pooled internal control improved the accuracy of protein quantification and inter-gel spot matching. Experiments performed using *Escherichia coli* lysate, spiked with 8 known increasing concentrations of BSA have clearly demonstrated the power of using the pooled internal standard when analysing spot variations between two different gels.

When known concentrations of spiked BSA spot volumes were assessed using the internal standard, the coefficient of variation ranged from 2.29 to 6.45 with an average of 4.0. By contrast, when these same experiments were analysed without the internal control, the coefficient of variation ranged from 8.51 to 120.93 with an average of 43.86. Similar results were seen with three other known proteins examined including conalbumin, GAPDH and trypsin inhibitor (5). These results show that the incorporation of a pooled internal control is imperative for accurate protein quantification. Although pooled internal controls have been used in a two dye system (20) the level of analysis does not go beyond direct comparison of two samples. It is through the addition of a third dye that the sophistication of an experimental design is refined with the ability to compare between a greater number of samples and observe minute changes in protein expression with precision.

Summary

Through the use of the minimal lysine labelling 2D-DIGE system, proteomics takes a step into the future. Employing innovative chemistry, this powerful tool enables direct comparisons of protein samples in a manner that was never previously possible. 2D-DIGE has already assisted in a better understanding of protein markers associated with infiltrating ductal carcinoma of the breast (21) the development of the cat brain (10,15), the reaction of the Saccharomyces cerevisiae proteome to metal stress (18) and protein changes in mitochondrial proteins from mouse hearts deficient in creatine kinase (19), to name a few. Although DIGE is a relatively new technique it is continually developing into a better one, with the introduction of saturation cysteine dyes and development of internal standard controls. Armed with a technique of such potential, the possibilities for comparative proteomics expand beyond anything that was previously conceivable.

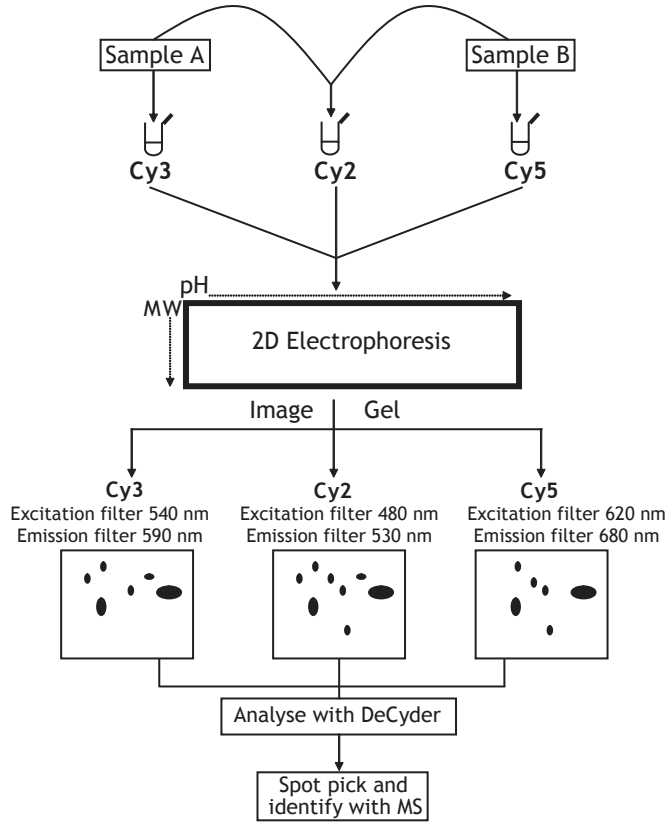

Fig. 1. Multiple sample analysis DIGE using Cy2

References

1. Huber, L.A., Pfaller, K., and Vietor, I. (2003) *Circ. Res.* **92**, 962-968
2. O'Farrell, P.H. (1975) *J. Biol. Chem.* **250**, 4007-4021
3. Unlu, M., Morgan, M.E., and Minden, J.S. (1997) *Electrophoresis* **18**, 2071-2077
4. Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., and Davison, M. (2001) *Proteomics* **1**, 377-396
5. Alban, A., David, S.O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S., and Currie, I. (2003) *Proteomics* **3**, 36-44
6. Schwartz, R., Ting, C.S., and King, J. (2001) *Genome Res.* **11**, 703-709
7. Zhou, G., Li, H., DeCamp, D., Chen, S., Shu, H., Gong, Y., Flaig, M., Gillespie, J.W., Hu, N., Taylor, P.R., Emmert-Buck, M.R., Liotta, L.A., Petricoin, E.F., and Zhao, Y. (2002) *Mol. Cell Proteomics* **1**, 117-124
8. Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M.J., Cramer, R., Waterfield, M.D., and Timms, J.F. (2002) *Mol. Cell Proteomics* **1**, 91-98
9. Shaw, J., Rowlinson, R., Nickson, J., Stone, T., Sweet, A., Williams, K., and Tonge, R. (2003) *Proteomics* **3**, 1181-1195
10. Van den Bergh, G., Clerens, S., Vandesande, F., and Arckens, L. (2003) *Electrophoresis* **24**, 1471-1481
11. Von Eggeling, F., Gawriljuk, A., Fiedler, W., Ernst, G., Claussen, U., Klose, J., and Romer, I. (2001) *Int. J. Mol. Med.* **8**, 373-377
12. Goldman, R.C., Trus, B.L., and Leive, L. (1983) *Eur. J. Biochem.* **131**, 473-480
13. Badock, V., Steinhusen, U., Bommert, K., and Otto, A. (2001) *Electrophoresis* **22**, 2856-2864
14. Van den Bergh, G., and Arckens, L. (2004) *Curr. Opin. Biotechnol.* **15**, 38-43
15. Van den Bergh, G., Clerens, S., Cnops, L., Vandesande, F., and Arckens, L. (2003) *J. Neurochem.* **85**, 193-205
16. Wildgruber, R., Harder, A., Obermaier, C., Boguth, G., Weiss, W., Fey, S.J., Larsen, P.M., and Gorg, A. (2000) *Electrophoresis* **21**, 2610-2616
17. Hoving, S., Voshol, H., and van Oostrum, J. (2000) *Electrophoresis* **21**, 2617-2621
18. Hu, Y., Wang, G., Chen, G.Y., Fu, X., and Yao, S.Q. (2003) *Electrophoresis* **24**, 1458-1470
19. Kernec, F., Unlu, M., Labeikovsky, W., Minden, J.S., and Koretsky, A.P. (2001) *Physiol. Genomics* **6**, 117-128
20. Yan, J.X., Devenish, A.T., Wait, R., Stone, T., Lewis, S., and Fowler, S. (2002) *Proteomics* **2**, 1682-1698
21. Somiari, R.I., Sullivan, A., Russell, S., Somiari, S., Hu, H., Jordan, R., George, A., Katenhusen, R., Buchowiecka, A., Arciero, C., Brzeski, H., Hooke, J., and Shriver, C. (2003) *Proteomics* **3**, 1863-1873