FEBS Letters 369 (1995) 122–125 FEBS 15686

Minireview

Mapping cells and sub-cellular organelles on 2-D gels: 'new tricks for an old horse'

Lukas A. Huber*

Département de Biochimie, Université de Genève, 30, quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland

Received 25 April 1995

Abstract Nowadays, investigators in all fields are faced with the identification of unknown, up- or down-regulated, modified proteins that they are trying to identify. Two-dimensional (2-D) gel electrophoresis, with its ability to resolve several thousand proteins, is an extremely powerful technique. The current resolution and reproducibility of 2-D gel technology and the establishment of computer assisted 2-D gel protein databases have paved new ways for the identification of proteins.

Key words: 2-D gel electrophoresis; 2-D GTP ligand blotting; 2-D gel protein databases; Sub-cellular fractionation; Protein identification

1. Introduction

The analysis of complex patterns of protein expression has been tremendously improved since the development of two-dimensional (2-D) gel electrophoresis in which proteins are first separated according to their charge (pI) and then to their molecular weight [1]. For example, this sophisticated technique has enabled the separation of all the proteins of *E. coli* or a yeast cell. Mammalian cells, however, are likely to express more proteins than a standard 2-D gel can routinely resolve. Computer program systems have been designed that enable rigorous qualitative and quantitative analysis of protein expression on digitized 2-D gels. Using these programs one can match together and catalogue large numbers of 2-D gel patterns, thereby allowing the construction of comprehensive databases. The power of this approach is enhanced when combined with Western immunoblotting, transient expression of cDNAs, protein micro sequencing and amino acid analysis. Today 2-D gel protein databases are linked to each other as well as to international specialized data banks on nucleic acid and protein sequences, genome mapping, protein structures and NMR experimental data.

This minireview describes how the combination of 2-D gel electrophoresis with new satellite technologies can lead to the identification of a defined protein present in a complex pattern. Available 2-D gel methods, as well as established databases will be described, and strategies will be given for identifying a spot on a gel.

*Corresponding author. Fax: (41) (22) 702-6464.
E-mail: LHUBER@sc2a.unige.ch

2. The first hurdle: compatibility

The immense amount of information generated by 2-D gel analysis is permanently feeding cell- and tissue-catalogues in form of protein databases. The main 2-D gel systems used at present are the traditional carrier ampholine system (here cited as IEF) and more recent the immobilized pH gradients (IPG). The first one (IEF) is based on the original methodology described in the landmark publication by O'Farell in 1975 and the pH gradient in the first dimension IEF is established in tube gels by soluble ampholines [1,2]. In contrast, immobilized pH gradients are used for isoelectric focusing in the IPG system [3,4]. There gels are bonded to a plastic backing and run horizontally. Whereas in IEF the key to reproducibility of the system lies simply in the choice and mixture of ampholines, IPG-strips are commercially available and exist in a variety of gradients (linear vs. nonlinear, gradient intervals, etc.). This results in higher reproducibility, probably a very important attempt within the field to standardize the procedure. Furthermore, in order to resolve rather basic proteins (pI $\geq$ 7.5) anodic or cathodic sample application can be done with the same strip, which offers an interesting alternative to the technically demanding non-equilibrium gels (NEPHGE) in the ampholine-carrier system. The IPGs also have a higher loading capacity (in terms of protein amounts as well as of sample volume), which greatly facilitates micro sequencing and amino acid analysis. However, the IEF gels still represent the most commonly used system, and working with 2-D gels becomes a major advantage when many people are using them.

As a matter of fact 2-D gels focused in the first dimension with IEF- or IPG-gradients, from one and the same sample, are looking considerably different and make it difficult to directly exchange 2-D gel protein data between the two systems. An useful tool that has been developed to simplify matching of protein-patterns is 'spot transfer', where identified proteins are excised and eluted from a 2-D gel run on one gel system to transfer identification of the proteins to a different 2-D gel system by comigration with a comparable sample [5]. More fanciful, comprehensive software packages as MELANIE or QUEST II are capable of simulating different pH gradients, as well as different size separations in the second dimension (SDS-PAGE) on digitized gel patterns by so called 'warping modules'. With these software packages 2-D gels performed with IEF- or IPG-gradients, could be electronically overlaid and matched directly, at least to some extent.

The most important conclusion for the inexperienced user is the endeavor for direct compatibility with existing databases.

SSDI 0014-5793(95)00614-1

After choosing a specialized 2-D gel database by criteria as species or cell type, the provided protocols should be followed exactly in order to link your results later to the database [6–9]. Current 2-D protein databases will be listed and discussed briefly in the next section.

3. 2-D gel databases

The human keratinocyte 2-D gel database at Aarhus, Denmark [6], has been set up to study cell proliferation, differentiation and disease within a single cell type. Currently more than 3000 cellular proteins (2168 IEF, 919 NEPHGE) are listed and almost 1000 polypeptides have been identified. The database is updated and published yearly in Electrophoresis (VCH, Weinheim, Germany), an international journal for electrophoresis research. The protein entries are listed both in alphabetical order and with increasing sample spot number (SSP number), together with their molecular weight, pI, sub-cellular localization and annotations regarding cellular differentiation patterns, response to various activation pathways and links to skin diseases as psoriasis or cancer. Efforts are currently underway to make this rich source of information available through international networks.

Other human reference protein maps have been established for plasma, red blood cells, liver, cerebrospinal fluid and other human samples [7,8,10–14].

An *E. coli* gene-protein database links each of the protein spots from 2-D gels of different strains grown under different conditions to the gene that encodes the protein [15,16].

A theoretical yeast 2-D gel database has been generated based on the genome sequence information of yeast (*Saccharomyces cerevisiae*) and the literature on yeast proteins. This database contains information such as electrophoretic coordinates, calculated codon bias and known modifications of yeast proteins [17] and can be linked to an experimental one, regularly updated, an approach being used for protein databases of other organisms [17,18].

Many other groups have started to generate 2-D gel protein databases with the intention to linking these to protein and DNA sequence databases. The major problem so far has been that these databases were not linked properly to each other and no uniform way to access the large volumes of data was available. Whereas IPG databases are more recent and, thus, constrain less information, the SWISS-2-D PAGE (see Table 1) computerized database was an important step towards standardizing and sharing database information over international networks [9]. In this electronic database scanned images of 2-D gels (IPGs) are linked to information on the identity of the protein spots. Information on mapping procedures, physiological and pathological data and bibliographical references are also provided, as well as dynamic links to SWISS-PROT protein sequence and other related databases such as the European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database, Protein Sites and Patterns (PROSITE) and On-line Mendelian Inheritance in Man (OMIM). A molecular biology server called ExPasy, located in Geneva, Switzerland, gives public access to the SWISS-2D PAGE database and the related protein maps (see Table 1). Information on the protein spots, including cross references, can be simply obtained by clicking on a spot of one of the supplied 'master 2-D gel' images (presently human liver, human serum and several human cell lines) [7–9]. Alternatively the theoretical location of any protein in the SWISS-PROT database can be mapped onto any of the 'master'-images. Most importantly all protocols developed and established by the SWISS-2-D gel team (Hochstrasser, D., et al.) are accessible by a simple mouse click and public courses are offered via internet for introducing the new IPG system.

Another internet accessible database server is the REF52 global gel navigator, housed at Cold Spring Harbor Laboratories (Cold Spring Harbor, NY, USA) [18]. This server is also directly linked to the ExPasy-molecular biology server. By servers like ExPasy or the REF52 (see Table 1) information from various fields could be grouped, catalogued, and offered to the scientific community as an uniform information retrieval model as well as a communication platform which can be publicly accessed.

Table 1
2-D gel databases available via Internet

Name of server	Database description	Electronic address
ExPasy Server	human liver, plasma etc. SWISS-PROT SWISS-2DPAGE SWISS-3DImage BIOSCI Melanie software	http://expasy.hcuge.ch/
CSH QUEST Protein Database Center	rat, mouse embryo, yeast 2D REF52 Quest software	http://siva.cshl.org/
NIMH-NCI Protein Disease Database (PDD)	common body fluids GELLAB software	http:/www-pdd.ncifcrf.gov/
Argonne Protein Mapping Group Server	mouse liver, human breast etc. hyperthermophilic microorganisms	http://www.anl.gov/CMB/pmg_welcome.html
E. coli Gene-Protein Database Project	ECO2DBASE	ftp://ncbi.nlm.nih.gov/repository
Cambridge 2DPAGE	rat neuronal database	http://sunspot.bioc.cam.ac.uk/
Human Myocardial Two-Dimensional Electrophoresis Protein Database	human heart proteins	http://www.chemie.fu-berlin.de/user/pleiss/dhzb.html

4. From a spot to a name

Most of us want to be able to identify their proteins of interest directly from 2-D gels, or in case of unknown proteins, to obtain sufficient protein-sequence information from a spot for cloning the gene. Proteins can usually be identified in a sample by a number of techniques: these include micro sequencing, peptide mapping, amino-acid composition, 2-D Western blotting, immunoprecipitations and transient over expression of cDNAs. However, alternative techniques could serve as a useful first pass method in an integrated approach to protein identification.

Obviously the first step is taking full advantage of the existing pool of information by matching your gels against protein patterns in databases. Proteins that have been mapped already in one of the existing 2-D databases could simply be identified by comigration experiments with so-called 'gold-standard samples' from the database of choice. Even though this approach is limited to experiments performed with the same standardized gel systems, immense information could be achieved already within a single comigration experiment.

In the following, two examples (from the biased view of a cell biologist) will be given to illustrate the usefulness of such an approach:

(i) In a recent study we used a 2-D gel approach and the currently available human keratinocyte 2-D gel protein database to characterize the composition of phagosomes/phagolysosomes formed in human and rodent macrophages [19]. This way, we were able to resolve over 200 polypeptides in the human preparations, identified more than 30 proteins associated specifically with phagosomes, and yielded new information on yet unknown proteins of this organelle [19]. Reference maps of different sub-cellular organelles will be available on database servers in the near future and the generation of specific 'organelle fingerprints' will be an important step towards our understanding and definition of sub-cellular membrane boundaries. An example of IPG-mapped mammalian early endosomes (Huber, L.A in collaboration with the SWISS-2D PAGE-team, Geneva Switzerland) is available on the ExPasy Molecular Biology server in the SWISS-2D PAGE gel museum (http://expasy.hcuge.ch/ch2d/cours2d-gels1.html).

(ii) Relatively little is known about the protein composition of sub-cellular organelles. Only few organelle-specific marker proteins are known and it is difficult to distinguish between organelle specific proteins and contaminants. Small GTPases of the Ras superfamily can serve as specific protein markers for sub-cellular membrane boundaries. They behave as molecular switches and regulate a diverse spectrum of intracellular processes, including cellular proliferation and differentiation, vesicular trafficking and cytoskeletal control. To study the distribution of small GTPases in different cells and their organelles we have developed a mapping technique based on [α^{32}P]GTP ligand blotting of high resolution 2-D gels. By using carrier ampholines we have mapped recently more than 30 small GTP-binding proteins relative to their isoelectric points and according to their molecular weights [20,21]. With this technique we in collaboration with many other groups [20,22–27] have established a catalogue of small GTP-binding proteins from different sub-cellular fractions (Golgi, plasma membrane, endosomes, post-Golgi vesicles, clathrin-coated vesicles, synaptosomes, etc.). Combination of the 2-D gel GTP-binding reference map with the GTP-binding patterns from sub-cellular fractions offers an elegant and straight forward approach to identify sub-cellular organelle fractions simply by their specific GTP-binding protein profile. This example demonstrates that sub-cellular fractionation prior to especially adapted 2-D gel analysis should make it possible to identify more regulatory proteins of relatively low abundance in individual cell types and to construct databases for all sub-cellular organelles.

Protein sequence analysis (N-terminal or internal) has been the bedrock for identification of unknown proteins from gels. Since this technique is both time consuming and expensive, the continued use of micro sequencing as first stage in protein analysis from 2-D gels represents in many cases a waste of sequencer time, when the protein turns out to be known and characterized already. This has pushed forward the development of a complementary technique, that is rapid and less expensive. Amino acid analysis and experimentally obtained data (pI, molecular weight) on the protein of interest together with available information in sequence databases can be used in algorithms to produce computerized 'match' tables with scores as indication of the likelihood of the match being correct.

The recent developments of standardized and commercially available first dimensions (IPGs) further facilitate now the possibility of directly comparing 2-D maps of various cell types generated in different laboratories. Furthermore, network serves with 2-D gel databases can now be accessed from any computer connected to internet network (see Table 1), and electrophoretic coordinates of known proteins or theoretical positions of yet unknown proteins can be displayed on your computer screen [9,28].

5. Perspectives

With a little imagination a scenario can be envisaged for rapidly scanning an entire 2-D gel (keyword: laser scanning by MALDI-MS) and providing information that can be compared or shared over international networks. As the genome sequencing projects from different species proceed apace, 2-D protein databases will provide tremendous information on the expression of sequenced genes, including the sub-cellular localization, cell and tissue distribution and changes in production of their products under various conditions.

Acknowledgements: I am especially grateful to Drs. G. van der Goot, J. Gruenberg, T. Harder and M. Rojo, all at the Department of Biochemistry, Sciences II, University of Geneva for helpful discussions and carefully reading the manuscript.

References

[1] O'Farell, P.H. (1975) J. Biol. Chem. 250, 4007–4021.
[2] Bravo, R. (1984) In: Two-dimensional Gel Electrophoresis of Proteins (Celis, J.E. and Bravo, R., Eds.), pp. 3–36. Academic Press, Orlando, FL.
[3] Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, J.C. and Hochstrasser, D. (1993) Electrophoresis 14, 1357–1365.
[4] Corbett, J.M., Dunn, M.J., Posch, A. and Gorg, A. (1994) Electrophoresis 15, 1205–1211.
[5] Dean, D.P., Cronan, M.T., Merenda, J.M., Gardner, J.P., Connelly, P.A. and Celis, J.E. (1994) Electrophoresis 15, 540–543.
[6] Celis, J.E., Rasmussen, H.H., Olsen, E., Madsen, P., Leffers, H., Honore, B., Dejgaard, K., Gromov, P., Vorum, H., Vassilev, A., Baskin, Y., Liu, X.D., Celis, A., Basse, B., Lauridsen, J.B., Ratz, G.P., Andersen, A.H., Walbum, E., Kjaergaard, I., Andersen, I.,

Puype, M., Vandamme, J. and Vandekerckhove, J. (1994) Electrophoresis 15, 1349–1458.
[7] Hughes, G.J., Frutiger, S., Paquet, N., Pasquali, C., Sanchez, J.-C., Tissot, J.-D., Bairoch, A., Appel, R.D. and Hochstrasser, D.F. (1993) Electrophoresis 11, 1216–1222.
[8] Golaz, O., Hughes, G.J., Frutiger, S., Paquet, N., Pasquali, C., Sanchez, J.-C., Tissot, J.-D., Appel, R.D., Walzer, C., Balant, L. and Hochstrasser, D.F. (1993) Electrophoresis 11, 1223–1231.
[9] Appel, R.D., Sanchez, J.-C., Bairoch, A., Golaz, O., Min, M., Vargas, J.R. and Hochstrasser, D.F. (1993) Electrophoresis 14, 1232–1238.
[10] De Curtis, I., Howell, K.E. and Simons, K. (1988) Exp. Cell Res. 175, 248–265.
[11] Anderson, N.L. and Anderson, N.G. (1991) Electrophoresis 12, 883–906.
[12] Edwards, J.J., Anderson, N.G., Nace, S.L. and Anderson, N.L. (1979) Blood 53, 1121–1132.
[13] Taylor, J., Anderson, N.L., Scandora, A.E.J., Willard, K.E. and Anderson, N.G. (1982) Clin. Chem. 28, 1062–1066.
[14] Gemell, M.A. and Anderson, N.L. (1982) Clin. Chem. 28, 955–961.
[15] Neidhardt, F.C., Appleby, D.B., Sankar, P., Hutton, M.E. and Philips, T.A. (1989) Electrophoresis 10, 116–122.
[16] VanBogelen, R.A., Sankar, P., Clark, R.L., Bogan, J.A. and Neidhardt, F.C. (1992) Electrophoresis 13, 1014–1054.
[17] Garrels, J.I., Futcher, B., Kobayashi, R., Latter, G.I., Schwender, B., Volpe, T., Warner, J.R. and Mclaughlin, C.S. (1994) Electrophoresis 15, 1466–1486.
[18] Boutell, T., Garrels, J.I., Franza, B.R., Monardo, P.J. and Latter, G.I. (1994) Electrophoresis 15, 1487–1490.
[19] Desjardins, M., Celis, J.E., Vanmeer, G., Dieplinger, H., Jahraus, A., Griffiths, G. and Huber, L.A. (1994) J. Biol. Chem. 269, 32194–32200.
[20] Huber, L.A. and Peter, M.E. (1994) Electrophoresis 15, 283–288.
[21] Huber, L.A., Ullrich, O., Takai, Y., Lutcke, A., Dupree, P., Olkkonen, V., Virta, H., Dehoop, M.J., Alexandrov, K., Peter, M., Zerial, M. and Simons, K. (1994) Proc. Natl. Acad. Sci. USA 91, 7874–7878.
[22] Huber, L.A., Pimplikar, S., Parton, R.G., Virta, H., Zerial, M. and Simons, K. (1993) J. Cell Biol. 123, 35–45.
[23] Urbé, S., Huber, L.A., Zerial, M., Tooze, S.A. and Parton, R.G. (1993) FEBS Lett. 334, 175–182.
[24] Desjardins, M., Huber, L.A., Parton, R.G. and Griffiths, G. (1994) J. Cell Biol. 124, 677–688.
[25] Nickel, W., Huber, L.A., Kahn, R.A., Kipper, N., Barthel, A., Fasshauer, D. and Soling, H.D. (1994) J. Cell Biol. 125, 721–732.
[26] Dehoop, M.J., Huber, L.A., Stenmark, H., Williamson, E., Zerial, M., Parton, R.G. and Dotti, C.G. (1994) Neuron 13, 11–22.
[27] Deretic, D., Huber, L.A., Ransom, N., Mancini, M., Simons, K. and Papermaster, D.S. (1995) J. Cell Sci. 108, 215–224.
[28] Celis, J.E., Honore, B., Bauw, G. and Vandekerckhove, J. (1990) Bioessays 12, 93–97.