Submitted to the Symposium on Application of Biochemical and Molecular Biology Techniques to Proteins of Parasite and Vector Identification to be held in Geneva, Switzerland, November 8-10, 1982.

HIGH RESOLUTION TWO-DIMENSIONAL ELECTROPHORESIS AS A METHOD OF IDENTIFYING ORGANISMS AND CELL TYPES

N. Leigh Anderson

Molecular Anatomy Program

Division of Biological and Medical Research

Argonne National Laboratory

Argonne, Illinois 60439

Keywords: 2-D electrophoresis
cell type identification
organism identification
protein taxonomy

HIGH-RESOLUTION TWO-DIMENSIONAL ELECTROPHORESIS AS A METHOD OF IDENTIFYING ORGANISMS AND CELL TYPES

N. Leigh Anderson
Molecular Anatomy Program
Division of Biological and Medical Research
Argonne National Laboratory, Argonne, IL 60439

INTRODUCTION

The set of proteins synthesized in a particular cell effectively characterizes that cell in terms of both function and species. Within one organism, each type of cell produces enzymes, structural proteins, etc., appropriate to its function, while evolution has resulted in sequence differences between analogous proteins in different organisms. Any method capable of resolving cellular proteins and distinguishing between evolutionary analogues should thus be able to characterize cells as to type and species. This paper attempts to give a brief overview of the usefulness of protein mapping in areas related to parasite and host identification and study.

High Resolution Protein Separation

The highest resolution method for protein separation currently available is two-dimensional gel electrophoresis (1). The technique consists of isoelectric focusing in 9 M urea in the first dimension, 2% NP-40 followed by sodium dodecyl sulfate (SDS) electrophoresis in a perpendicular direction, producing a pattern of spots, each of which is a distinct protein component. The ISO-DALT system of 2-D electrophoresis allows resolution of 1000-2000 proteins in sets of 20-40 samples simultaneously (2,3). Figure 1 shows a representative 2-D pattern of proteins from a human fibroblast cell line. Since the separation is sensitive to molecular charge in the first dimension, amino acid changes (such as those resulting from mutation) resulting in a charge change can be detected as a shift in protein spot position. The introduction of computer-based data reduction systems (4,5) now makes it possible to quantitate more than 1,000 proteins per gel and to compare results from large numbers of analyses. Typical detection limits for an average protein are 50 ng using Coomassie Blue staining, 0.5 ng using the silver stain, and about 1 dpm for radiolabeled proteins using fluorography.

Characterization of Cell Types

Within one organism, there is a hierarchy of differences among cell types reflecting their relative origins in ontogeny. Mammalian fibroblasts and lymphocytes share perhaps 50-70% of the proteins expressed in either, while lymphocytes and monocytes (which derive from a common bone marrow precursor) are much more alike and differ by a set of about 50 qualitative marker proteins (6). T- and B-lymphocytes are yet more alike (showing about 30 mainly quantitative differences), while suppressor and helper T-cells show almost indistinguishable two-dimensional protein patterns (K. E. Willard-Gallo, personal communication). These results are as yet only fragmentary, but they indicate that by examining large numbers of proteins as possible markers, it should be possible to define the positions of cells on an ontogenetic tree in

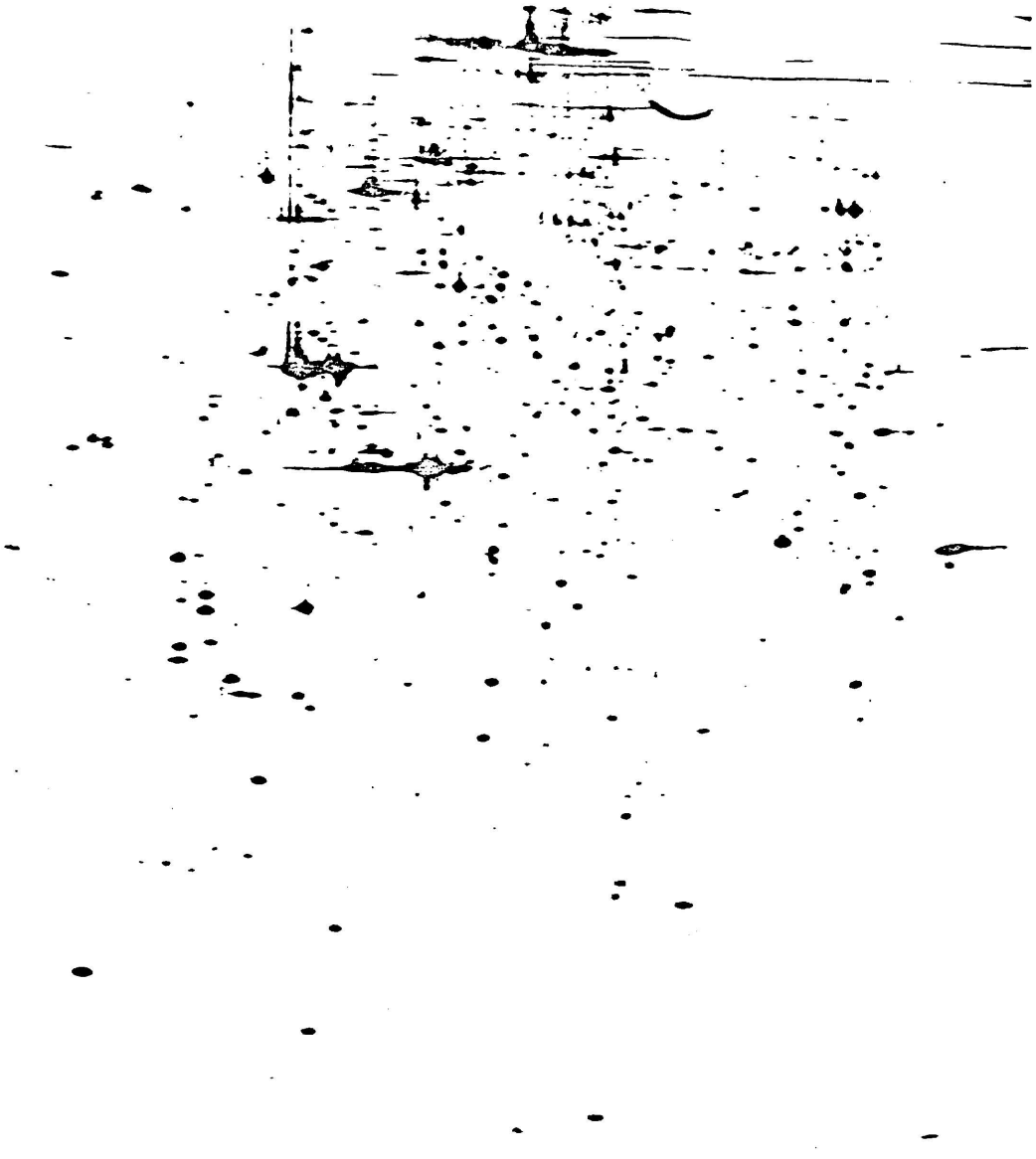

Fig. 1. Two-dimensional electrophoretic patterns of proteins from a human fibroblast line (GM3040a). Acid isoelectric points are to the left (basic to the right) and high SDS-molecular weight proteins are at the top (lower MW's at the bottom). The major pair of spots at the center of the patterns are non-muscle actins.

precise terms. This would in turn provide a method of investigating the effects of parasitic infections on differentiation of particular host cell types.

Identification of Organisms

On an evolutionary time scale, all of the genes of an organism can be mutated and these mutations fixed in the population by selection or random processes. Different proteins "evolve" at different rates, so that by examining large numbers of polypeptides, one has access to molecular clocks running

at widely varying speeds. Since two-dimensional electrophoresis is sensitive to amino acid substitutions that result in a charge change (about 1/3 of single base charge mutations), it is possible to use the protein patterns as tools for measuring the relatedness of species. Previous investigations have shown that some proteins (β and γ actin and some mitochondrial proteins, for instance) are charge-invariant throughout most of the mammals, while other proteins vary in charge between closely-related species (such as the principal heat-shock protein; 7). Thus protein patterns from goat and bovine fibroblast look quite similar (though easily distinguishable) and the two are more alike than either is to dolphin or human fibroblasts (NLA, unpublished observations). Variation between protein patterns of cells from different individuals of the same species appears limited to about 1-2% of the proteins. Differences between species are thus of a much greater magnitude than between individuals of the same species, at least for mammals. A determinative classification of species based on high-resolution protein maps is thus a feasible undertaking.

Gene Expression Systems

At the other end of the protein difference spectrum from species differences are the changes caused in a single cell type by some treatment (infection by an intracellular parasite or virus, treatment with chemicals, etc.). Such changes can, if they are inducible in vitro, be observed directly in comparisons of protein maps of treated cells with appropriate controls. In human lymphoid cells, specific proteins have been identified whose synthesis is strongly affected by heat shock (7), anti-mitochondrial agents (8), 5-azacytidine (NLA and M. A. Gemmell, manuscript in preparation), interferon and a variety of other substances (9), or by Epstein-Barr virus infection (10). In addition, differences due to clonal variations can be observed, as in the cases of light and heavy immunoglobulin chains (11), or variant-specific surface antigens in trypanosomes (12).

CONCLUSIONS

Two-dimensional electrophoresis of proteins is a general tool for observing gene expression, and for obtaining information about the gene products expressed. Since proteins are the principal working parts of cells, the inventory of proteins expressed by a given cell type can indicate what the cell is and what organism it comes from. It follows that an important long-term goal in any field devoted to the analysis of a particular class of organisms should be the construction of a complete catalog of the proteins encoded by the genome(s) in question (13,14). This has already been accomplished for some viruses, and is between 25 and 50% complete for E. coli. However, it appears that a more intense and systematic effort will be required to achieve the goal of a complete catalog for major eukaryotes, since major organisms of interest (man, for instance) may have between 30- and 50,000 structural genes, compared to about 1,500 in E. coli.

Protein mapping as an analytical tool is not divorced from the technologies of DNA cloning and sequencing. Sufficient protein can be obtained from 2-D gels to allow n-terminal sequencing of 20-40 amino acid residues; this in turn is sufficient to allow synthesis of a DNA probe capable of identifying the associated gene in a clone library. Protein mapping may thus be regarded

as an effective method of identifying interesting proteins for further investigation. Unlike present nucleic acid techniques, it is ideally suited to an overall examination of large numbers of gene products simultaneously, with the concomitant ability to detect quantitative changes in each due to the influence of experimental variables. As such, it can be a useful tool in the identification of parasites, hosts, and vectors, and in the study of the interactions between cells of these classes of organism.

ACKNOWLEDGMENT

This work is supported by the U.S. Department of Energy under contract No. W-31-109-ENG-39.

REFERENCES

1. O'Farrell, P. H. (1975) J. Biol. Chem. 250, 4007-4021.

2. Anderson, N. G. and Anderson, N. L. (1978) Anal. Biochem. 85, 331-340.

3. Anderson, N. L. and Anderson, N. G. (1978) Anal. Biochem. 85, 341-354.

4. Anderson, N. L., Taylor, J., Scandora, A. E., Coulter, B. P., and Anderson, N. G. (1981) Clin. Chem. 27, 1807-1820.

5. Taylor, J., Anderson, N. L., Scandora, A. E., Jr., Willard, K. E., and Anderson, N. G. (1982) Clin. Chem. 28, 861-866.

6. Gemmell, M. A., and Anderson, N. L. (1982) Clin. Chem. 28, 1062-1066.

7. Anderson, N. L., Giometti, C. S., Gemmell, M. A., Nance, S. L., and Anderson, N. G. (1982) Clin. Chem. 28, 1084-1092.

8. Anderson, L. (1981) Proc. Nat. Acad. Sci. 78, 2407-2411.

9. Anderson, N. L. (1981) Electrophoresis '81, Allen, Arnaud, Eds., W. deGruyter, pp 309-316.

10. Willard, K. E. (1982) Clin. Chem. 28, 1031-1035.

11. Anderson, N. L. (1981) Immunology Letters 2, 195-199.

12. Pearson, Terry, and Anderson, Leigh (1980) Anal. Biochem. 101, 377-386.

13. Anderson, N. L. (1982) Trends in Analytical Chemistry (TRAC) 1, 131-135.

14. Anderson, N. G. and Anderson, N. L. (1979) Behring Inst. Mitt. 63, 169-210.