CHAPTER 27

USE OF HIGH-RESOLUTION TWO-DIMENSIONAL ELECTROPHORESIS IN HUMAN HEALTH EFFECTS STUDIES

Norman G. Anderson and N. Leigh Anderson

ABSTRACT

High-resolution two-dimensional electrophoresis, coupled with methods for protein identification and compositional analysis, is being developed with the aim of detecting and measuring the effects of radiation and of environmental pollutants in man. The techniques described allow a large fraction of all tissue proteins to be detected and measured under conditions which allow about one third of amino acid substitutions to be seen. Samples required are in the submilligram range. Internal standards for charge (pI) and for molecular weight have been developed, together with systems for doing many analyses in parallel (the ISO-DALT system). The present capacity is up to 100 analyses per day. Scanners and computerized data reduction system for image analysis and protein identification are now on line and allow many patterns to be analyzed and intercompared. Present emphasis is on the analysis of human lymphocyte, fibroblast, muscle, plasma, urinary, and prostatic fluid proteins. The incidence of polymorphisms seen in cellular proteins is less than one tenth of that estimated from previous studies, suggesting that the system will be useful for detecting mutations in man. The development of the Human Protein Index, in support of this work, and of the associated analytical systems to measure and detect both genetic and somatic injury in human populations are discussed.

N.G. Anderson and N.L. Anderson are affiliated with the Molecular Anatomy Program, Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Il.

INTRODUCTION

Cost and human health effects are two central concerns in the evaluation of any and all energy sources. The greatest interest is in the effects of low exposures to radiation or to toxic agents to operating personnel and to surrounding populations. The effects of high exposures on man can be detected and studied by existing techniques. Low-level exposure effects have thus far been primarily evaluated in large controlled animal experiments or by epidemiological studies. One of the aims of present research is to develop more sensitive detection methods so that the effects of lower doses of both radiation and energy-related or environmental pollutants may be studied directly in both man and animals. In this presentation we discuss the present status of attempts to approach the theoretical limits of human health effects measurements by studying alterations in the structure, amount, or distribution of proteins using readily obtainable human samples. We describe first the types of measurements to be made, then the development and validation of suitable techniques. Lastly we discuss the work remaining to be done before the methods can be applied to human populations at risk, including the development of a Human Protein Index.

TYPES OF EFFECTS TO BE MEASURED

Cytotoxicity

A common denominator of many types of cytotoxicity is cell leakage. This may be a direct effect of the toxic agents on the cell membrane, or it may be one step along a pathway which may lead to cell death. In man, many cellular proteins are small enough to pass through the kidney and appear in the urine, and hence may be used as indicators of specific organ damage, quite independent of the specific molecular cause of the injury (Anderson, Anderson, and Tollaksen, 1979). (Note that the detection of enzymes from heart and liver in blood or urine have long been used clinically to detect injury to these organs.) Analysis of the spectrum of proteins found in human urine in response to toxic injury therefore offers great promise. Injury to tissues resulting in tissue leakage must, however, be distinguished from alterations in urine proteins due to damage to the kidney itself. In cadmium poisoning, for example, the kidney tubules may be injured, leading to failure to reabsorb and catabolize many low molecular weight proteins. These then appear in the urine and are frequently used as indicators of toxicity. The basic requirement for this work is a method for analyzing urine and plasma for a large number of different protein species at one time.

Carcinogenicity

In tumor cells some genes which are silent in the cell of origin may be activated, while others which are normally switched on are switched off. The effects of carcinogens on human populations might be detected if specific cancer-related proteins can be found, and if it can be shown that these are being produced in response to tumor promoters or carcinogens. The problem here is to be able to study alterations in gene expression in detail on very small human samples. Parenthetically, it is important to be able to distinguish clearly substances which may be toxic but do not alter gene expression from those which may or may not be very toxic, but which *do* alter gene expression.

Many tumors shed antigens which may appear in plasma or urine. It is important to be able to monitor these easily and efficiently in large populations both for health effects measurements, and for cancer monitoring generally. The techniques required again are those which allow large numbers of proteins to be "seen" in very small samples at one time.

Mutagenicity

Proposed human mutation rate measurements depend on finding protein variants in offspring which are not found in either parent. The number of trios which must be examined is inversely proportional to the number of individual proteins examined in each analysis (Neel, Tiffany, and Anderson, 1973). Hence it is again important to develop very high resolution methods which can see as many proteins in each sample as possible.

ANALYTICAL REQUIREMENTS

It is estimated that the human genome contains 30,000–50,000 structural genes, and that possibly 10 percent or less of these are expressed in any one cell type. Some proteins are post-translationally modified, and most would be found to have variant forms, hence the analytical requirements (ability to resolve well over 50,000 proteins) are formidable indeed. It is very unlikely that any one-dimensional separation will be capable of dealing with this number of entities.

For human studies there are additional requirements. Only very small samples are routinely available, and these include urinary proteins, plasma, red cells, white cells, hair follicles, skin biopsies, and for special purposes, a variety of other small-sample biopsies. It is important to be able to do all analyses on less than one milligram of protein. Ultimately it must be possible to identify each of the proteins resolved, and to assign some function to it, i.e., to produce a Human Protein Index. Note that an estimated 500–1,000 human proteins have been well enough characterized to have some analytical number associated with each. This is only 1 to 2% percent of the estimated total. Hence it appears that most human proteins are not only unknown, but have functions which are unsuspected. This does not detract from the value of their study, however. If one finds a variant protein of unknown function whose presence correlates with disease or with disease susceptibility, then one has a test for that disease or for susceptibility, and some clues as to function.

THE ISO-DALT SYSTEM

High-resolution two-dimensional electrophoresis (Klose, 1975; O'Farrell, 1975; Scheele, 1975; Iborra and Buhler, 1976; N.G. Anderson and N.L. Anderson, 1978; N.L. Anderson and N.G. Anderson, 1978; N.G. Anderson and N.L. Anderson, 1979) utilizes two analytical methods, each dependent on a different and unrelated property of proteins (Stegemann, 1970), and each capable of resolving separately mixtures of over 100 proteins. When these methods are combined two-dimensionally, the resolution is the product of the resolution of each separately, i.e., theoretically 10,000. In practice over 2,000 proteins can be resolved using extracts from one cell type (Willard and Anderson, 1980). This technique uses denaturing conditions, and hence sees monomeric proteins or protein subunits. At Argonne a semi-automated version of this method (the ISO-DALT SYSTEM) has been developed which allows up to 100 analyses to be run per day. Protein detection is accomplished either by staining, autoradiography, or by fluorography. With autoradiographs or fluorographs the developed X-ray films are scanned with a drum scanner, and the image processed using the TYCHO and KEPLER software systems (Taylor et al., 1980). These systems perform background subtraction, spot detection, and spot modeling to resolve each spot into two-dimensional Gaussian peaks; normalization (stretching) of gels to match a standard gel; and recording of the spot list in terms of spot coordinates and integrated absorbance. For stained gels, scanners based on rasters generated by a rotating mirror and a laser beam and by high-resolution CRT tubes are under development. While the limits of resolution of two-dimensional electrophoresis have not been reached, it is apparent that a number of other methods must be used in conjunction with it to achieve the analytical resolution ultimately required. These methods include cell separation, subcellular fractionation, and a variety of group separations largely based on affinity chromatography. Specific antibodies against each protein will be required for identification, cross-checking of results between laboratories, and the development of specific assays for individual proteins of diagnostic interest.

THE HUMAN PROTEIN INDEX

As this work has progressed, it has become evident that the proteins seen on two-dimensional maps should be cataloged to form an index (Anderson, 1979), and that some considerations must be given to how the index is constructed. Each entry (protein) in the Human Protein Index will be listed with the following descriptors:

1. Identifying number (rules for numbering will be discussed elsewhere).
2. Spot location on standard 2-D gel with reference to internal charge and molecular weight standards.
3. Cell type(s) in which protein is found.
4. Amount of protein in spot in gel analyzed.
5. Subcellular location and how determined.
6. Chromosomal location of the gene for the protein.
7. Amino acid composition.
8. Coregulational set to which a protein belongs, including a list of the other members of the set.
9. How protein varies under experimental conditions (i.e., after exposure to toxic agents, radiation, drugs, hormones, or physical variables.
10. Genetic polymorphisms, and the frequency of their occurrence.
11. Biophysical properties including thermal denaturation inflection point in the presence and absence of cofactors and substrates (Nance, Hickman, and Anderson, 1980).
12. Availability of identifying monospecific antibodies.
13. Relationship to disease states.
14. Tertiary structural data including the identity of other proteins with which it may be associated.

The purpose of the KEPLER software system is to organize this information and to search it for interesting correlations. For example, one could search for protein variants whose presence in a population correlates with susceptibility to injury by a specific toxic agents.

PRESENT STATUS

The purpose of this program is to lay the groundwork for human health effects detection and measurement at the molecular level, and to develop and validate the necessary analytical systems for studies on human populations thought to be at risk. This requires that we work out methods for obtaining and analyzing several different types of samples, identify as many constituents in each type of sample as possible, and then analyze a series of normal samples and samples from individuals with known disease or exposures to radiation or deleterious chemicals.

The prototype versions of the ISO-DALT system have thus far been used to run over 35,000 high-resolution two-dimensional gels. Systems for automatic processing of negatives of stained gels, autoradiographs, and prints have been installed and calibrated. Internal standards for charge (pI) (Anderson and Hickman, 1979; Hickman et al., 1980) and molecular weight (Giometti et al., 1980) have been developed, together with methods for immunochemical identification of spots. Patterns have been published for human plasma (Anderson

and Anderson, 1977a,b; N.L. Anderson and N.G. Anderson, 1979), saliva (Giometti and Anderson, 1980), peripheral lymphocytes (N.G. Anderson and N.L. Anderson, 1979), leukemic cells (N.L. Anderson, in preparation), normal and diseased muscle (Giometti, Anderson, and Anderson, 1979; Giometti et al., 1980), hair follicles (N.G. Anderson and N.L. Anderson, 1979), human urinary proteins (Anderson et al., 1979; Anderson, Anderson, and Tollaksen, 1979; Tollaksen and Anderson, 1980), and red cell proteins (Edwards et al., 1979; Edwards, Hahn, and Anderson, 1980). Over 80 spots have been resolved in the light chain fraction of human and mouse immunoglobulins (Anderson and Anderson, forthcoming). Over 500 human urine samples have been analyzed in the development of human urinary protein concentration and identification techniques. Marked and characteristic differences are seen between normal urinary proteins and proteins from patients with glomerular disease, tubular disease, and multiple myeloma (in addition to Bence-Jones protein), and interesting and suggestive differences are seen between normal urinary proteins and proteins from cancer patients and from individuals with viral infections. Protein sets altered by tumor promoters, Epstein-Barr virus infection, and interferon in human leukemic cells have been identified (N.L. Anderson, unpublished results).

A simple optical comparator system (Anderson and Anderson, 1977b; Anderson, Anderson, and Tollaksen, 1979) has been used to compare two patterns, which is useful when only small numbers of samples are to be analyzed. Three-dimensional analysis using electrophoretic transfer to cellulose nitrate is now routine. We have assisted in the organization of laboratories and training of staff in the United States, England, and Norway.

DISCUSSION

In studies involving the effects of a wide range of toxic substances on human peripheral blood lymphocytes, clear indication of the division of toxic agents into two groups was obtained. With the first group no change in gene expression was seen up to exposure levels which killed the cells. The second group of compounds, which was a very small one, produced characteristic changes in cell proteins, including both the suppression of synthesis of proteins normally present, and the de novo synthesis of proteins not seen in untreated cells. This strongly suggests that toxic agents and new drugs should be tested using this technique. Note that many of the control circuits for gene expression which are now examining are absent from cells which have been maintained in tissue culture, suggesting that such studies should be carried out with fresh cells.

It is quite evident that much further development and standardization must be done before the system can come into wide use clinically. Orienting studies now in progress, however, suggest many clinical applications. We believe that if methods for measuring low-level effects in many are successfully developed, they will become part of the routine of the clinical chemistry laboratory, and will have many uses in addition to those for which they were developed. Monitoring human populations at risk for the effects of mutagens, carcinogens, and toxic agents generally requires that both normal populations and groups having known exposures to radiation or to toxic agents be extensively examined. The technology to begin to do this now exists.

REFERENCES

Anderson, N.G. 1979. "Making molecular catalogues." *Nature 279:* 122-123.

Anderson, N.G., and N.L. Anderson. 1978. "Analytical techniques for cell fractions. XXI. Two-dimensional analysis of serum and tissue proteins: Multiple isoelectric focusing." *Anal Biochem 85:* 331-340.

Anderson, N.G., and N.L. Anderson. 1979. "Molecular anatomy." *Behring Inst Mitt 63:* 169-210.

Anderson, N.G., N.L. Anderson, and S.L. Tollaksen. 1979. "Proteins of human urine. I. Concentration and analysis by two-dimensional electrophoresis." *Clin Chem 25:* 1199-1210.

Anderson, N.G., et al. 1979. "Analytical techniques for cell fractions. XXV. Concentration and two-dimensional electrophoretic analysis of human urinary proteins." *Anal Biochem 95:* 48-61.

Anderson, N.L., and N.G. Anderso. 1977a. "High-resolution two-dimensional electrophoresis of human plasma proteins." *Proc Natl Acad Sci USA 74:* 5421-5425.

Anderson, N.L., and N.G. Anderson. 1977b. "Two-dimensional separation of serum proteins, use of antibodies for identification, and optical methods for gel intercomparison." *Fed Proc 36:* 1063.

Anderson, N.L., and N.G. Anderson. 1978. "Analytical techniques for cell fractions. XXII. Two-dimensional analysis of serum and tissue proteins: Multiple gradient-slab gel electrophoresis." *Anal Biochem 85:* 341-354.

Anderson, N.L., and N.G. Anderson. 1979. "Microheterogeneity of serum transferrin, haptoglobin and alpha-2 HS glycoprotein examined by high-resolution two-dimensional electrophoresis." *Biochem Biophys Res Commun 88:* 258-265.

Anderson, N.L., and N.G. Anderson. "The potential of high-resolution protein mapping as a method of monitoring the human immune system." IN *Proceeding of Conference on the Biological Relevance of Immune Suppression as Induced by Genetic, Therapeutic, and Environmental Factors.* Van Nostrand Reinhold Company, New York. (forthcoming)

Anderson, N.L., and B.J. Hickman. 1979. "Analytical techniques for cell fractions. XIV. Isoelectric point standards for two-dimensional electrophoresis." *anal Biochem 93:* 312-320.

Edwards, J.J., H.J. Hahn, and N.G. Anderson. 1980. "Two-dimensional electrophoretic mapping of human erythrocyte proteins." pp. 383-393. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berlin. 858 pp.

Edwards, J. J., et al. 1979. "Red cell proteins. I. Two-dimensional mapping of human erythrocyte lysate proteins." *Blood 53:* 1121-1132.

Giometti, C.S., and N.G. Anderson. 1980. "Two-dimensional electrophoresis of human saliva." pp. 395-404. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berline. 858 pp.

Giometti, C.S., N.G. Anderson, and N.L. Anderson. 1979. "Muscle protein analysis. I. High-resolution two-dimensional electrophoresis of skeletal muscle proteins for analysis of small biopsy samples." *Clin Chem 25:* 1877-1884.

Giometti, C.S., et al. 1980. "Analytical techniques for cell fractions. XXVII. Use of heart proteins as reference standards in two-dimensional electrophoresis." *Anal Biochem 102:* 47-58.

Giometti, C.S., et al. 1980. "Muscle protein analysis. II. Two-dimensional electrophoresis of normal and diseased human skeletal muscle." *Clin Chem 26:* 1152-1155.

Hickman, B.J., et al. 1980. "Internal charge standardization for two-dimensional electrophoresis." pp. 341-350. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berlin. 858 pp.

Iborra, G., and J.M. Buhler. 1976. "Protein subunit mapping. A sensitive high resolution method." *Anal Biochem 74:* 503-511.

Klose, J. 1975. "Protein mapping by combined isoelectric focusing and electrophoresis in mouse tissues. A novel approach to testing for induced point mutation in mammals." *Humangenetik 26:* 231-243.

Nance, S.L., B.J. Hickman, and N.L. Anderson. 1980. "A method for studying proteins in 2-D gels using thermal denaturation analysis." pp. 351-360. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berlin. 858 pp.

Neel, J.V., T.O. Tiffany, and N.G. Anderson. 1973. "Approaches to monitoring human populations for mutation rates and genetic disease." pp. 105-150. IN Hollaender, A. (ed.), *Chemical Mutagens, Volume 3.* Plenum Press, New York.

O'Farrell, P.H. 1975. "High resolution two-dimensional electrophoresis of proteins." *J Biol Chem 250:* 4007-4021.

Scheele, G.A. 1975. "Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins." *J Biol Chem 250:* 5375-5385.

Stegemann, H. 1970. "Proteinfraktionierungen in polyacrylamid und die Anwendung auf die genetische Analyse bei Pflanzen." *Angew Chem 82:* 640.

Taylor, J., et al. 1980. "Estimation of two-dimensional electrophoretic spot intensities and positions by modeling." pp. 329-339. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berlin. 858 pp.

Tollaksen, S.L., and N.G. Anderson. 1980. "Two-dimensional electrophoresis of human urinary proteins in health and disease." pp. 405-414. IN Radola, B. (ed.), *Electrophoresis '79.* Walter de Gruyter & Company, Berlin. 858 pp.

DISCUSSION

Paul Ts'o, The Johns Hopkins University: I would like to add a comment first that this is really very exciting. It's the beginning of the description of cells at gene expression level and/or disease in human beings. Our laboratory and several others are starting on this type of program at the RNA and DNA level, but of course it is very difficult surveying without the advantages you have. On the other hand, as we progress closer to the mechanism for gene control with which you and I are both concerned, we do come up with one very important challenge. As we are using the molecular probe to better define molecular mechanisms, the situation of the cell biology has to be more and more defined because we find normal cell 1 is not equivalent to normal cell 2 is not equivalent to normal cell 3. Therefore, the comparison of individual normal cells and tumor cells means that as you express it very well, you don't know which difference really counts. That's one of the most important challenges. At the present, we have to pursue that matter through clonal cell transformation: isolate a cell and get its progeny and follow the transformation of its progeny. We cannot use populations any more.

Anderson: We find much more day-to-day variation in tissue culture cells than we do in lymphocytes taken straight from people.

Richard Albertini, University of Vermont: Following up on that, you are if you are talking about phenotypes of populations. Do you have subpopulations in your lymphocyte work? Are you restricting it to peripheral blood lymphocytes or to any subpopulation?

Anderson: We have done cell sorter work and separated out tissue and blood cells, and we'll continue with the subfractionation of these. But our initial question was really how variable are different people, looking at the whole population of peripheral lymphocytes. The answer is that the level of polymorphism is about one tenth of what we expected, and other people have found the same thing. It takes a little bit of study to find out what the differences are, but the identification can be done. There are not so many differences that we can't find our way around in these patterns. Also, the levels of posttranslational modification is much less inside a cell than outside, fortunately.

Paul Selby, Oak Ridge National Laboratory: I think you are wise to be cautious in trying to figure out what these differences really mean. What is the chance that many of the differences between people would be of consequence as far as health risk and not be reflected by any differences in the type of system that you are looking at?

Anderson: This is probably the most sensitive system we've got. With Jim Neil, we're going through all the mutations that he determines by classical techniques to see whether they can all be screened by this method, and the answer so far is that they can. Now, there are many mutations that are silent by this technique. But there are many others, of course, where protein just plain disappears. We're adding now a third dimension to this system which will solve some of these problems.

Robert Painter, University of California, San Francisco: In your estimate, how many silent mutations occur compared to the ones that would move one of these patterns?

Anderson: Two thirds of them should be silent by this technique, because, on the average, those are point mutations. Two thirds of single amino-acid substitutions would not produce a charge shift that would be visible. Those you would not see.