Fundamentals of Protein Separations: 50 Years of Nanotechnology, and Growing

David A. Egas and Mary J. Wirth

Department of Chemistry, University of Arizona, Tucson, Arizona 85721; email: egas@email.arizona.edu, mwirth@email.arizona.edu

Annu. Rev. Anal. Chem. 2008. 1:833–55

First published online as a Review in Advance on April 10, 2008

The *Annual Review of Analytical Chemistry* is online at anchem.annualreviews.org

This article's doi: 10.1146/annurev.anchem.1.031207.112912

1936-1327/08/0719-0833$20.00

Key Words

proteomics, microarrays, electrophoresis, isoelectric focusing, chromatography, biomarkers

Abstract

The separation of proteins in biology samples has long been recognized as an important and daunting endeavor that continues to have enormous impact on human health. Today's technology for protein separations has its origins in the early nanotechnology of the 1950s and 1960s, and the methods include immunoassays and other affinity extractions, electrophoresis, and chromatography. What is different today is the need to resolve and identify many low-abundance proteins within complex biological matrices. Multidimensional separations are the rule, high speed is needed, and the separations must be able to work with mass spectrometry for protein identification. Hybrid approaches that combine disparate separation tools (including recognition, electrophoresis, and chromatography) take advantage of the fact that no single class of separation can resolve the proteins in a biological matrix. Protein separations represent a developing area technologically, and understanding the principles of protein separations from a molecular and nanoscale viewpoint will enable today's researchers to invent tomorrow's technology.

1. INTRODUCTION

The sequencing of the human genome has greatly accelerated the field of protein analysis. Whereas the genome is the blueprint for the organism, the proteins carry out the functions of biological cells. The analysis of proteins has emerged as a vital area of medical research and clinical diagnostics. Every disease can potentially be described by its unique set of proteins and their concentrations. The magnitude of the analytical task is daunting: The human genome codes for over 20,000 proteins, and there are scores of physiologically relevant post-translational modifications. Consequently, protein analysis represents an enormous challenge to the separation scientist. We begin this review with illustrations of why protein analysis is such a vital and active field today, and we show the principles of the separations used in these analyses. We demonstrate that nanotechnology pervades protein separations, and the design of new separation technology rests on a fundamental understanding of basic principles.

2. EXTRACTIONS OF SINGLE PROTEINS

Disease is associated with a change in protein levels. This means that the diagnosis of disease, in principle, can be made through the detection of unusual protein levels. As a common example, diseases of the thyroid and pituitary glands are diagnosed in part by the analysis of the protein thyrotropin, which is a hormone secreted by the pituitary gland, stimulating the thyroid gland (1). Clinical labs use a commercial kit that tests for this protein in blood samples by immunoassay, which is the most common technique for clinically sensing protein levels in body fluids. A protein, or any other species, whose concentration is indicative of a disease is called a biomarker. In this case, multiple biomarkers are used to distinguish between thyroid and pituitary diseases. Other common clinical immunoassays include biomarkers for heart attack, prostate cancer, infectious disease, and pregnancy. Immunoassays are among the most powerful separations today, selectively extracting one type of protein at nanomolar concentration levels out of a mixture of tens of thousands of proteins. Immunoassays are not normally discussed in the analytical chemistry literature as separation tools, but they are extractions, and the fundamentals are common to all extractions: Selectivity and fast mass transport are the primary technological issues.

Yalow & Berson (2) originated the concept of the immunoassay in 1959, earning Yalow the Nobel Prize in 1977 for this invention, which revolutionized clinical analysis. The immunoassay is based on nature's nanotechnology, in which an antibody specifically binds to the protein through a spatial arrangement of functional groups that fit like a lock and key with the functional groups of the protein or other analyte of interest (3). **Figure 1** illustrates the commonly used sandwich immunoassay, in which two different antibodies recognize different parts of a protein, which is the analyte. The capture antibody is bound to the surface, and the protein of interest is selectively removed from the complex mixture, such as blood serum, by virtue of the high specificity and strength of the binding. The immunoassay is essentially a highly selective solid-phase extraction. The orientation of the capture antibody cannot be controlled so nicely in practice, and the analysis relies on a sufficient number of capture antibodies that have their recognition sites accessible to the protein of interest.

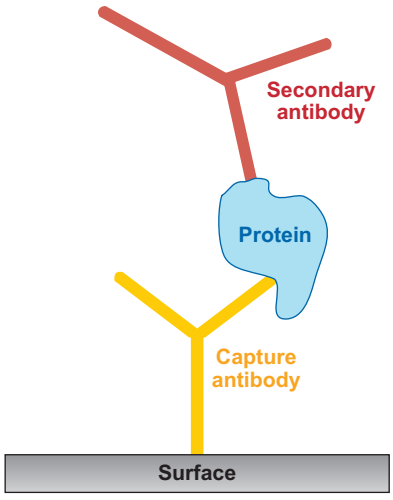

Figure 1

Depiction of antibodies in an immunoassay of a protein.

The binding constant, K_b, is defined in Equation 1 for the antibody, antigen, and complex concentrations ([Ab], [Ag] and [AbAg], respectively):

$$K_b = \frac{[\mathrm{AbAg}]}{[\mathrm{Ab}]\,[\mathrm{Ag}]}, \tag{1}$$

where K_b ranges from 10^5 to 10^9. Binding equilibria allow the analysis of concentrations approximately two orders of magnitude lower than $1/K_b$, which is the nanomolar to picomolar scale for immunoassays. High binding constants correlate with the slow kinetics of dissociation of the complex because K_b is the ratio of rate constants for binding, k_b, and dissociation, k_d:

$$K_b = \frac{k_b}{k_d}. \tag{2}$$

Because the rate constant for dissociation is so small, unwanted species can be rinsed from the substrate with minimal loss of the targeted protein. To detect the protein that has been extracted, one uses a secondary antibody. This antibody is typically labeled to use fluorescence or chemiluminescence for signal transduction. In developing new immunoassays, the selectivity (or its opposite, the cross-reactivity) must be characterized because proteins sometimes have similar sequences or even a subunit in common with another protein.

The speed of immunoassays can be undesirably slow. The full expression for the rate of formation of the antibody-antigen complex, AbAg, has three terms, which result from the rate of diffusion of antigen to the antibody, the rate of binding of antigen to antibody, and the rate of dissociation of the complex (4):

$$\frac{d\,[\mathrm{AbAg}]}{dt} = D \cdot \nabla^2\,[\mathrm{Ag}] + k_b \Gamma_{\mathrm{Ab}}\,[\mathrm{Ag}] - k_d\,[\mathrm{AbAg}]\,, \tag{3}$$

where D is the diffusion coefficient of the antigen, and Γ is the surface concentration of the accessible antibody. As an illustration, a study of the binding kinetics of the HIV-1 capsid protein p24 to a monoclonal antibody revealed that $k_b = 2.25 \times 10^5\ M^{-1}\ s^{-1}$ and $k_d = 0.0062\ s^{-1}$ (5). One can see from this very low rate constant for dissociation that rinsing removes the loosely bound species without significantly reducing the selectively bound species. One can also see that low antigen concentrations give low rates of binding despite the high rate constant. For example, if the concentration of antigen were 1 nM, then the time constant, $(k_b[\mathrm{Ag}])^{-1}$, would be 1300 s. The antibody can also deplete the surrounding solution of antigen, in which case diffusion is required to supply antigen, and this can be the slow step. Modeling Equation 3 using these parameters, or even a 100-fold lower value of k_b (4), shows that diffusion is typically the slow step. Therefore, researchers who want to decrease analysis times assist the mass transport with flow, electromigration, or the use of beads as substrates, which can be shaken in the solution of analyte.

Unmet needs for immunoassays include higher sensitivity, smaller quantities of reagents (particularly the antibodies), faster binding times, less nonspecific adsorption, and the development of inexpensive synthetic antibodies. Aptamers are an attempt to address this last need (6).

By the time a protein analysis has made its way into the clinical lab as an immunoassay, the research to discover the protein biomarkers, and the research to understand why the disease is correlated with each biomarker, has already been done. Protein separations that can analyze large numbers of proteins are needed for this research, and the first such separation we discuss below is the protein microarray.

3. PROTEIN MICROARRAYS

Investigators can study large numbers of proteins using microarrays of capture antibodies, which are now commercially available from a variety of sources. A protein microarray is an organized high-density assortment of microscopic spots of proteins immobilized onto a support. This can be a solid planar support (7–10), microspherical beads (11), and polymers or gels (12). Ink-jet printing is commonly used for the impregnation of proteins onto the supports (13). Other procedures include microcontact printing (8) and self-assembling protein microarrays (14). All methods of printing give pronounced deformities in spot shapes, often giving rings rather than filled spots (10), which can be a challenge in quantitative interpretation. Additives prevent the protein from denaturing when the microarray is dried after printing (15). In the use of protein microarrays (such as arrays of antibodies), the microscopic spots of printed protein bind the analytes from solution, and the same considerations discussed above for immunoassays (such as enhancing mass transport and rinsing weakly bound species) also apply for microarrays (16). Protein microarrays represent a rapidly developing technology.

As an example of one type of protein microarray, cytokine arrays are promising in medicine (17). Cytokines are small proteins involved in the inflammatory response, which is activated by disease. Because cytokine biology is complex, typically 12 cytokines need to be measured simultaneously to characterize the inflammatory process,

hence the attractiveness of microarrays. One interesting application is in tailoring HIV therapy to the patient. HIV destroys a certain type of white blood cell called a CD4 T cell, and current treatment to kill the virus, if effective, restores the levels of these white blood cells. Some HIV patients continue to have low CD4 levels despite the therapy lowering the viral load. Sachdeva et al. (18) used a cytokine array to study the differences between the two populations of patients having low viral loads but differing levels of CD4 cells. The array recognized 12 different cytokines. The authors found that two cytokines had significantly different levels in the two groups, and the identities of these two proteins point to greater immune damage in the case of the patients whose CD4 cell levels had not recovered. With further research, insight into the biology of HIV could lead to improved therapy for the group whose CD4 cells did not recover. Most applications of protein arrays are in the research stage because it is a new technology, and as with DNA microarrays, the technology will likely result in new clinical tests and drugs tailored to the patient.

Unlike DNA, proteins have no amplification reaction that is analogous to polymerase chain reaction (PCR); therefore, sensitivity is a problem that has slowed the development of protein microarrays. Consequently, the flat glass substrates widely used for DNA arrays are used less for proteins arrays. The substrates for protein arrays are usually thin, hydrophilic polymer films, which allow for more moles per area to be printed to increase sensitivity (12). The polymeric substrates for protein arrays are compatible with the same microarray printers used for the well-established technology of DNA microarrays. The printed spots spread more on the polymers, but the density requirements for protein arrays are not as great as they are for DNA arrays. Polymer substrates bind antibodies and other proteins well, and they leave the proteins in their native conformation to retain binding specificity. The binding is not understood: Printed proteins appear to physisorb (15). As with DNA arrays, researchers must use a blocking solution after printing to deactivate the surface so that it does not bind the analytes. New substrate materials for protein microarrays would be valuable because polymers suffer from low optical transmission and fluorescence background.

Silica colloidal crystals, which are porous crystals of silica nanoparticles, have recently been introduced as a substrate for protein arrays (19). They offer a high binding capacity, optical transparency, no fluorescence, and fast mass transport. Silica colloidal crystals occur in nature as opals, which makes them another example of nature's nanotechnology. **Figure 2** compares a natural opal of gem quality with the synthetic silica colloidal crystal shown in **Figure 2*b***. Scanning electron micrographs, for both the natural and synthetic opals, show that the silica nanoparticles are arranged in a face-centered cubic lattice structure. Because the sizes of the silica particles are on the order of the wavelength of light, their crystalline order gives Bragg diffraction at visible wavelengths (20). Scientists can make a crystal of monodisperse nanoparticles that has just one color, whereas nature creates a patient mixing of crystallites of different sized nanoparticles and orientations. For protein analysis, the size of the nanoparticle determines the pore size, and simple geometry reveals the limiting pore diameter to be 15% of the particle diameter. One would probably want a pore diameter of at least 40 nm to accommodate the sandwiches of three proteins for an antibody

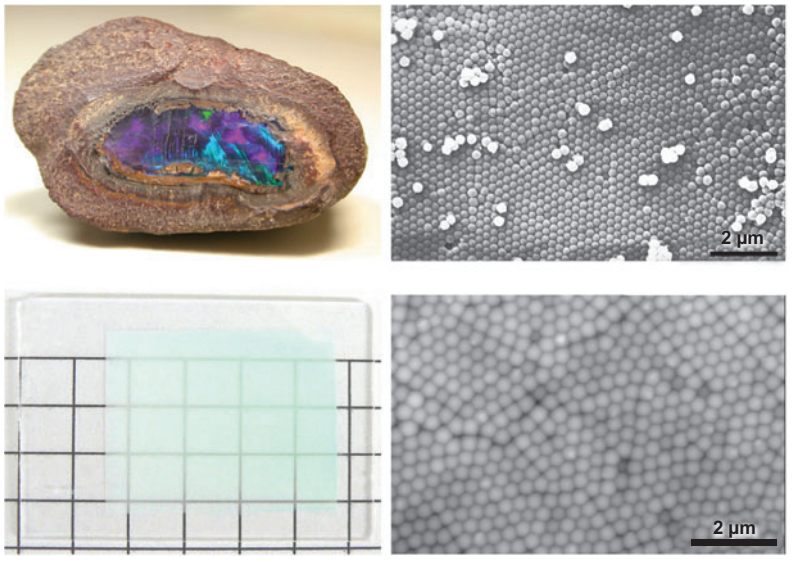

Figure 2

(*a*) Gem-quality opal (*left panel*) (courtesy of George R. Rossman, California Institute of Technology) and scanning electron micrograph (SEM) image (*right panel*) (courtesy of Chi Ma, California Institute of Technology). (*b*) Synthetic opal (*left panel*) and SEM image of the synthetic opal (*right panel*).

array. This would require nanoparticles on the order of 250 nm in diameter, which are readily synthesized and crystallized. Researchers compared a silica colloidal crystal and a flat surface for the capture of the protein streptavidin, using biotin bound to the surfaces of both types of substrates. The signal from labeled streptavidin was increased by a factor of 80 relative to that of the flat surface for a colloidal crystal 11 μm in thickness, made with nanoparticles 250 nm in diameter (19). This high enhancement agreed with theory, and the increase in fluorescence background was negligible. These materials are being evaluated currently as substrates for antibody arrays.

For both immunoassays and antibody arrays, it is necessary to know the identities of the proteins before the measurements can be made. Determining which proteins are involved in a disease, and discovering new biomarkers for disease, requires the separation and identification of unknown proteins from the complex sample. These separations must resolve enormous numbers of proteins, including the post-translationally modified forms of proteins. In addition, the separation has to be compatible with mass spectrometry to enable identification of the proteins. There is no one type of separation with adequate power to separate protein mixtures as complex as human serum or cell lysates. To address the challenge, researchers use combinations of powerful separations, each based on a complementary property of the protein. There is no one approach that works for all applications, and the field continues to develop. The next sections discuss the types of separations used.

4. GEL ELECTROPHORESIS: INTRODUCTION

Complex protein analyses use at least two orthogonal separations, combined with mass spectrometry. A given separation can be described by its peak capacity, which is

the number of peaks that can be fit into the separation window. A high peak capacity is 100. In two-dimensional (2D) separations, if one dimension has a peak capacity of n proteins and the other dimension has a peak capacity of m proteins, then the total peak capacity is the product, $n \times m$. This assumes that the mechanisms of the separations are independent of one another (i.e., they are orthogonal). If the mechanisms are perfectly correlated, the peak capacity remains on the order of n or m. Aside from this condition of orthogonality, the random overlap of peaks along one dimension reduces n and m each to 18% of their ideal values, based on statistics alone (21). The limited peak capacities, relative to the large number of proteins in serum samples, make multidimensional separations a requirement. Subsequent identification by mass spectrometry provides some additional resolution of overlapped peaks.

The peak capacity in one dimension is limited by the similarity of the properties of the components (22). For example, if the components are similar in size, a separation based on size inherently gives a small peak capacity. For multidimensional separations, Giddings (23) defined the term sample dimensionality as the number of independent types of properties that can be used to achieve a separation. For example, a polydisperse sample of polymer, such as polystyrene, has one dimension because only the molecular weight of polystyrene varies, not its composition. Proteins have multiple properties that vary because of the variety of amino acids. The most notable variables are molecular weight, charge, and hydrophobicity. A 3D separation of proteins is thus expected to be possible for complex protein samples.

In a typical 2D gel electrophoresis separation, one axis is the isoelectric point (i.e., charge) and the other axis is the molecular weight. These are quite orthogonal to one another. The third dimension of hydrophobicity is not used in conjunction with 2D gels because the proteins are denatured by sodium dodecyl sulfate (SDS) in the size separation, which masks the hydrophobicity of the protein. After separation by the 2D gel, the separated spots of denatured proteins are individually removed and digested by trypsin into peptides. These peptides are then separated by reversed-phase liquid chromatography (RPLC). The unique combination of peptides is used to identify the protein. Some overlap of protein spots in the gel can be tolerated because the mass spectrometer can resolve multiple peptides within overlapping peaks, and the algorithms for protein identification consider multiple proteins contributing to the collection of peptides from a given gel spot. One problem is that a high-abundance protein can hide a low-abundance protein. Nonetheless, owing to the orthogonality of its two dimensions and the power of mass spectrometers, the 2D polyacrylamide gel in combination with mass spectrometry is the workhorse of proteomics.

Investigators have used 2D gel separations in the search for biomarkers. Using a 2D gel technique called differential in-gel electrophoresis (DIGE), one can search for potential biomarkers by labeling proteins from healthy and diseased cells with different color dyes (24–26). The chief difficulty is that there are 12 high-abundance proteins in serum that compose 85% of the total serum protein, and these proteins are approximately six orders of magnitude more abundant than the known biomarkers (27). These give broad bands in the gel owing to precipitation, and they swamp out the signals of lower-abundance proteins. To alleviate this problem, researchers remove the high-abundance proteins by passing the serum through an affinity column containing

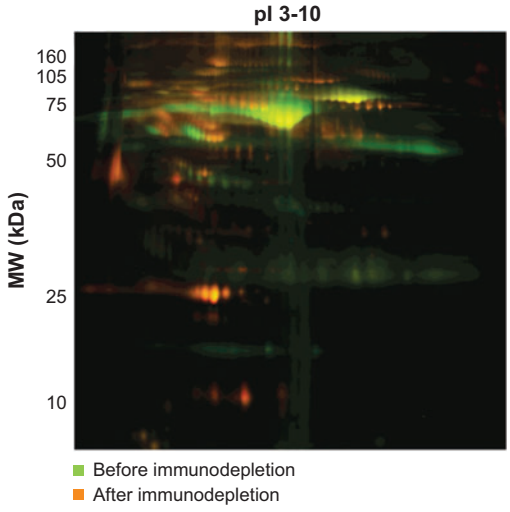

Figure 3

Differential in-gel electrophoresis image before and after the immunodepletion of abundant serum proteins. Figure taken with permission from Yu et al. (29).

antibodies that bind the high-abundance proteins (28). **Figure 3** illustrates an example of a separation by 2D DIGE, comparing the total serum protein with serum that has been depleted of the 12 most abundant proteins (29). One high-abundance protein, human serum albumin, is responsible for the bright peaks at approximately 66 kDa. The loss of the high-abundance proteins allows one to scale the intensity to reveal many proteins that were not visible in the whole serum sample.

In one study of biomarkers in pancreatic cancer, DIGE revealed apolipoprotein E as a potential biomarker because it is 7.7-fold higher in cancer patients than in healthy patients (29). This protein, however, was not validated as a biomarker after the study of a larger number of serum samples. Aside from the low number of serum samples that is reasonable for studies with DIGE, there may be a more fundamental reason why this method has not fulfilled its promise in biomarker discovery. Clinically successful biomarkers are very-low-abundance proteins (e.g., nanomolar concentration levels), and even with serum depletion, DIGE is not sensitive to such low levels (27). The potential biomarker mentioned above, apolipoprotein E, is 100-fold higher in abundance than prostate specific antigen, which is a widely used clinical biomarker whose concentration is on the nanomolar scale when indicative of cancer. More separation power to achieve a higher dynamic range is needed. In addition, much higher speeds are needed because it takes 1 day just to separate with DIGE. Below we examine the principles of electrophoresis to understand the factors that limit its resolution and speed.

5. GEL ELECTROPHORESIS: MECHANISMS

Isoelectric focusing takes advantage of the fact that the velocity, v, of the protein depends on the charge, z, of the protein:

$$v = zFE/f = \mu_e E, \tag{4}$$

where E is the applied electric field, f is the friction coefficient for the protein, and F is the Faraday constant. The terms z, F, and f are usually lumped together as the electrophoretic mobility, μ_e. The charge of the protein is dependent on pH, and every protein has a pH at which its charge is zero (i.e., its isoelectric pH). Isoelectric focusing uses a medium that has a pH gradient spatially imposed, and the protein electromigrates until it reaches the region of its isoelectric point, and then it stops. As it diffuses in either direction, it picks up a charge that causes it to electromigrate back to the position of its isoelectric pH, hence the term focusing. Resolution depends on the steepness of the pH gradient (*dpH/dx*), the change in protein charge with pH (*dz/dpH*), and the temperature (30):

$$R_s = \frac{\Delta pI}{4}\left(\frac{FE}{RT}\frac{-dz/d\mathrm{pH}}{d\mathrm{pH}/dx}\right)^{1/2}. \quad (5)$$

Extraordinary resolution is possible: $\Delta\mathrm{pH} = 0.001$ can be resolved for a shallow pH gradient and a large change in charge with pH.

Isoelectric focusing is typically performed in gel strips for proteomics, and this separation is allowed to run overnight. In principle, one could apply an enormous electric field and separate the proteins quickly. High fields are used for isoelectric focusing in capillaries, giving focusing in 7 min (31) and more recently in 30 s (32). The high initial currents are a problem for the larger cross-sectional areas of gels because the Joule heat cannot be dissipated quickly. Capillaries are thinner, giving less current for a given voltage and faster heat transfer. Any approach to speeding up 2D electrophoresis will likely involve miniaturizing the volume to minimize the effects of Joule heating.

The second dimension of the gel electrophoresis separation is sieving. This separates proteins based on their sizes, using the porous nature of the gel. Referring to Equation 4, the friction coefficient of the protein is related by Stokes' law to its radius, r:

$$f = 6\pi\eta r, \quad (6)$$

where η is the viscosity of the medium. Because the migration is electrically driven, the charges of the proteins must be the same for the separation to depend only on size. This is accomplished by denaturing the protein with the surfactant SDS, which binds to the protein in a ratio of 1.4 g of SDS per 1 g of protein (33). Because the charge per gram is constant, the electrophoretic mobility is constant. Separating the SDS-denatured proteins based on size now requires a sieving medium.

Investigators use nanotechnology to sieve proteins because of their nanoscale sizes. To achieve sieving, the pores through which the protein passes must be on the order of the size of the proteins. Researchers developed sieving media well before the use of SDS, beginning in the 1950s with starch gels (34). We can consider these starch gels to be the advent of nanotechnology in separations. A few years later, in 1959, polyacrylamide gels were introduced for protein separations (35), and agarose gels were introduced in 1966 (36). Investigators demonstrated the use of cross-linked polyacrylamide gel to sieve SDS-denatured proteins as early as 1967 (37). The polyacrylamide gels and SDS denaturation of the proteins remain the dominant technology in proteomics.

The principles of sieving electrophoresis explain why polyacrylamide gels have stood the test of time, and how a new material might be developed to improve the speed. Giddings (30) described the partition coefficient, K, of a protein, as it distributes between nanoscale pores and free solution by relating K to the fraction of accessible volume in the pore:

$$K = \frac{\text{accessible volume}}{\text{true volume}} = \frac{\pi\left(R^2 - r^2\right) \cdot L}{\pi R^2 \cdot L}, \tag{7}$$

where r and R are the radii of the protein and pore, respectively; and L is the pore length. This partition coefficient describes the mechanism of size-exclusion chromatography (SEC). Rodbard & Chrambach (38) showed that the electrophoretic mobility, μ_e, relative to that in free solution, μ_e^0, is equal to the same K as in SEC (i.e., both are described by the ratio of accessible to true volume of the pore):

$$K = \frac{\text{accessible volume}}{\text{true volume}} = \frac{\mu_e}{\mu_e^0}. \tag{8}$$

The ability of protein to electromigrate in a gel is thus dictated by the accessible volume in the gel pores. An intuitive way to understand this is that the velocity of the protein has to be zero when it is contact with a wall, and the accessible volume is proportional to the probability that the protein is not in contact with a wall. Entropy thus controls sieving separations.

The pores of polyacrylamide gels are not monodisperse. Ogston (39) first modeled the random structure of the gel fibers. Using this structural description, we arrive at the Ogston-Morris-Rodbard-Chrambach model for gel electrophoresis, which describes the electrophoretic mobility relative to that in free solution:

$$\frac{\mu_e}{\mu_e^0} = \exp\left[-\frac{\pi}{4}\left(\frac{d+r}{R}\right)^2\right], \tag{9}$$

where d is the radius of the gel fibers; and r and R are again defined as the protein and pore radii, respectively. This model agrees well with experiment (40). Other models give comparable agreement with experiment (40), so one should not take Equation 9 as the law of the land. Even using Equation 7 (which would ignore the distribution of pores), one can obtain a similar dependence of mobility on protein size.

The essential issue for any sieving model is that the radius of the pore must be on the same order as the radius of the protein to give a sieving separation. Experimentally, the pore size of the gel is controlled by the percentage of both monomer and cross-linker in the gel (41). To illustrate the role of pore size, **Figure 4** shows a series of plots of computed mobility versus protein radius (using Equation 9) for three different pore radii. In each case, the fiber radius was assumed to be equal to the pore radius for simplicity. The 10-nm pore radius is shown to be well suited to the molecular weight range of 10 to 200 kDa that is typically studied in proteomics. **Figure 4** also shows the same relation for the monodisperse pores, which require a somewhat larger pore radius to span this range of molecular weights. The graph illustrates how similar the behaviors are for the gel and the monodisperse pores. The flatter slope indicates that the monodisperse pores give somewhat greater selectivity than the distribution

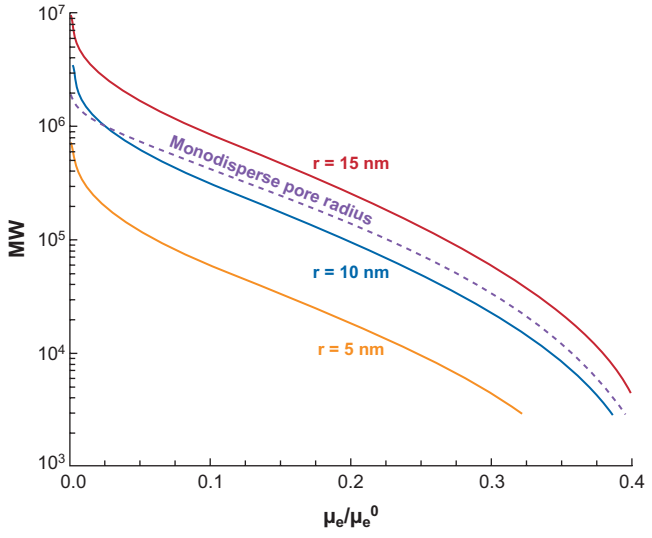

Figure 4

Calculated dependence of electrophoretic mobility on molecular weight (MW) for gel sieving with three different pore radii (*solid lines*). The dotted line is the same curve, but for monodisperse pore radius.

of pores. Another difference is the abrupt cutoff at $r = R$ for monodisperse pores, which is smoothed in the case of the gel because the distribution of pores provides a path for proteins larger than the average pore size. Overall, the differences between gels and monodisperse pores are rather small, and both illustrate that nanoscale pores enable a size-based separation of proteins.

Researchers have used atomic force microscopy (AFM) to image the nanoscale structure of gels used in electrophoresis, using tapping mode, which imparts minimal force to the soft material (42). **Figure 5*a*** shows an AFM image of a 3% agarose gel, revealing a web structure. This image gives us a visual way of thinking about gels. Image analysis revealed the distribution of pore diameters in **Figure 5*b***, which agrees with the Gaussian distribution in Ogston's (39) structural model. The AFM results further showed that the pore-size distribution narrowed with higher agarose concentration, decreasing by a factor of two for a 5% agarose solution. The pore diameters in all cases are much larger than the protein diameter. AFM images of polyacrylamide gels show the expected small pore sizes (43). The wide pores of the agarose introduce an important question: Why can such a wide-pore medium sieve small proteins?

The answer is interesting, and it has consequences for speeding up sieving separations. Even lower concentrations of agarose give size-based separations of small proteins (44). In general, large hydrophilic polymers in homogeneous solution impart a size-based separation without forming permanent pores. The explanation is that agarose is a soluble polymer, and at sufficiently high concentrations, the polymer chains become entangled to form transient pores (45, 46). The larger pores obvious in **Figure 5*a*** are not the same pores responsible for the sieving of much smaller proteins. Many hydrophilic polymers of high molecular weight are now available for sieving of proteins and DNA fragments, including polyethylene glycol, cellulose derivatives and linear polyacrylamide, and copolymers (47). These polymer solutions have been called replaceable gel media, non-cross-linked polymers, and entangled polymer

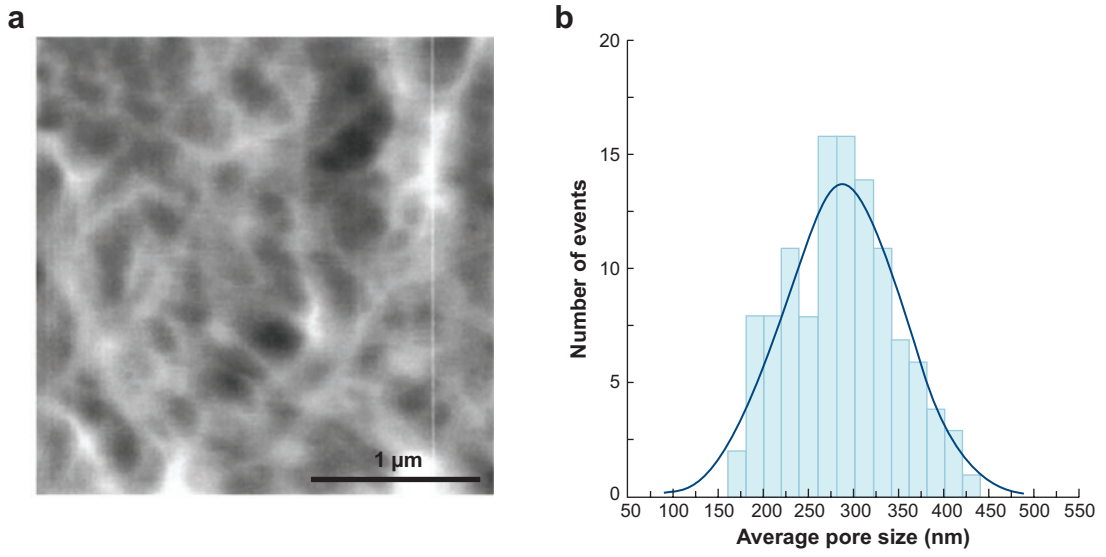

Figure 5

(*a*) Atomic force microscopy (AFM) image of agarose gel. (*b*) Pore size distribution determined from image analysis of the AFM image. Figure taken with permission from Pernodet et al. (42).

solutions. These media have enabled the routine use of gels in capillaries for fast and automated DNA sequencing (45, 48–50). This advance led to the completion of the human genome project ahead of schedule (51). Entangled polymer solutions are now used commercially for fast microchip-based protein sieving (52). In the latter case, a set of protein molecular weight standards was easily resolved in 40 s using an electric field of 340 V cm^{-1}. This is a tenfold higher electric field than a slab gel can tolerate. The microchip can accommodate the higher field strength because it is only 13 μm deep, giving lower current and faster heat dissipation. The ability to avoid preparing a gel, which must be thick to be reasonably uniform, is thus possible by using entangled polymer solutions in thin capillaries and thin channels of microchips. The drawback is that these media cannot be cast readily in a 2D format to allow 2D electrophoresis. The polymer solution is also difficult to couple with other separations or with mass spectrometry.

Silica colloidal crystals have been introduced as a thin separation medium that can be cast in a 2D format (53, 54). These materials are described above in the context of proteins arrays. The particle size can be tailored to give optimal pore sizes for sieving. Investigators demonstrated the separation of proteins by sieving through the pores of colloidal crystals in a microchip: The channel was filled with a colloidal crystal of 160-nm silica particles, and proteins from 20 to 205 kDa were separated using a field strength of 31 V cm^{-1} (54). Much higher fields can be used, and the authors did not observe any deleterious heating effects in electrochromatography up to 1000 V cm^{-1} (55). The surfaces of silica nanoparticles can be made to be as high in quality as the best chromatographic silica gel (56); in fact, silica nanoparticles are constituents of

the highest-quality silica gel particles (57). The use of these materials in separations will advance as better and faster means are being developed for making these crystals, such as calcining to enable crack-free material crystals (58) and spin coating to enable the deposition of crystals in less than 1 min (59). These materials have the potential for 2D protein separations because they are inherently 2D in format, and they have the potential for interfacing with mass spectrometry, either through matrix-assisted laser desorption/ionization or protein extraction. This could increase the sensitivity of DIGE because the colloidal crystals are intrinsically low in fluorescence (19).

There are many unmet needs for protein electrophoresis and 2D gel electrophoresis. Obviously the speed and resolution of the 2D separation need to be improved. Another unmet need is low cost. DIGE analyses in particular are expensive, with one 2D gel separation typically costing approximately $2000 without the mass spectrometry. Devices for fast 2D separations on the benchtop in researchers' labs that allow detection of lower-abundance proteins would be valuable.

6. CHROMATOGRAPHY OF PROTEINS

Gel electrophoresis utilizes size and charge as the two properties on which to base a protein separation. In high-performance liquid chromatography (HPLC), it is common to use the third distinguishing property of proteins, which is hydrophobicity. This is achieved by RPLC. These separations use a hydrophobic stationary phase, typically a monolayer of short hydrocarbon chains covalently bonded to the silica surface through a siloxane linkage. This monolayer is on the order of 1 nm in thickness, which allows fast mass transport to the mobile phase. These monolayer bonded phases represent another example of the early use of nanotechnology in separations (60). Separations of proteins using stationary phases of hydrophobic monolayers are best achieved by gradient elution RPLC, for which the gradient begins with a mobile-phase composition that is mostly water and ends with a composition that is mostly organic, typically acetonitrile. For peptides, the elution occurs in order of increasing hydrophobicity (61). This necessarily would correlate with size if two peptides or proteins had similar amino acid compositions, but overall the correlation with size is not high. For example, for bovine serum, albumin (66 kDa) elutes earlier than the much smaller myoglobin (16.7 kDa) (62). RPLC thus contributes in a complementary way to selectivity in protein separations by separating on the basis of hydrophobicity.

Chromatography can also separate on the basis of size. SEC is based on the same fundamental principles as sieving electrophoresis, as detailed above. SEC offers one potential advantage over sieving electrophoresis: There is no need for denaturing with SDS. The need for SDS denaturation in electrophoresis arises because the native proteins have different charges, and so much SDS is adsorbed that it masks the differences in protein charges. Because there is no electric field in SEC, the charges of the proteins do not contribute to the retention times; therefore, SDS is not needed in SEC. The problem with using SDS in electrophoresis is that it prevents the use of mass spectrometry on whole proteins because there is no effective way of removing SDS from the gel spots. Size exclusion, by avoiding this denaturing, would allow the proteins to be directly analyzed by mass spectrometry.

There are two reasons why size exclusion is not used nearly as extensively as sieving for protein separations. First, chromatography columns do not lend themselves readily to 2D separations. Second, the peak capacity of SEC is low. We can overcome the first issue to some extent by using sequential columns to achieve 2D separations, in which one column is eluted slowly and the other column analyzes each peak quickly (63). Coupling two 1D columns to achieve a 2D separation requires more planning and design than separations that have a spatially 2D format, such as a slab gel or a thin layer chromatography plate. As one illustration, Moore & Jorgenson (64) performed a 3D separation of peptides from a tryptic digest of ovalbumin using SEC as the first dimension, followed by reversed-phase HPLC, and finally by capillary electrophoresis. These are the three orthogonal properties pointed out above for proteins, and the same idea applies to peptides. **Figure 6*a*** shows the SEC separation of the peptides, illustrating that this technique does not separate the rather limited number of peptides well. (We discuss the reasons below.) **Figure 6*b*** shows the final

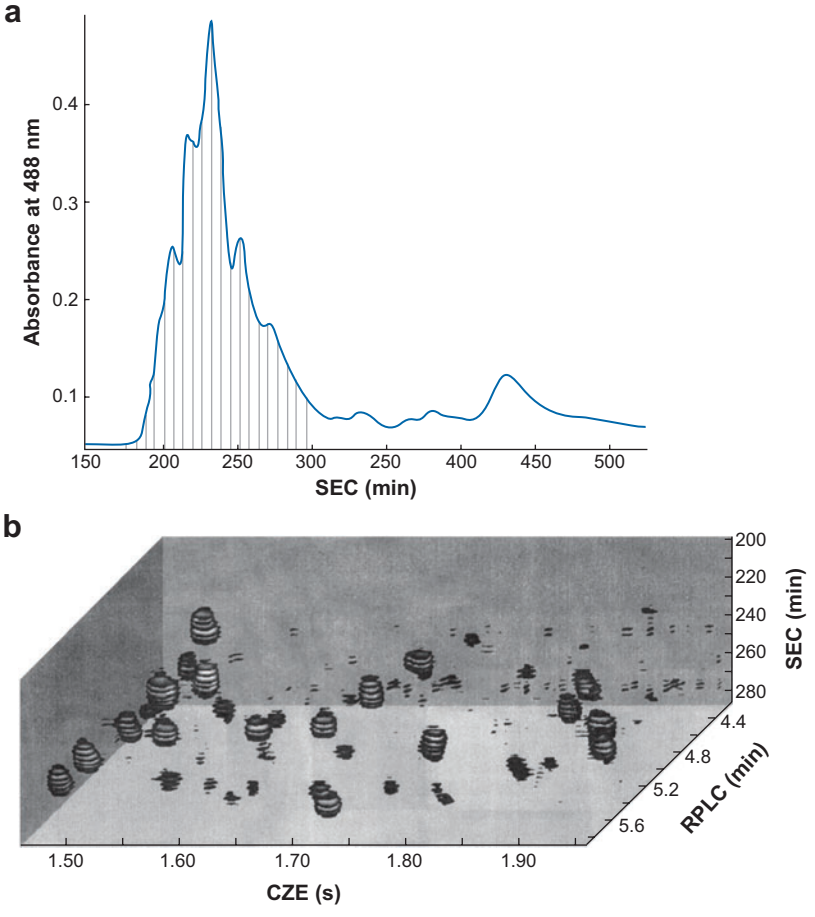

Figure 6

(*a*) Size-exclusion chromatogram of peptides. (*b*) Display of three-dimensional separation using size-exclusion chromatography (SEC), reversed-phase liquid chromatography (RPLC), and capillary zone electrophoresis (CZE) for the three dimensions. Figure taken with permission from Moore & Jorgenson (64).

3D separation. The vertical discs result from the fact that the broad, slowly eluting SEC peaks are sampled discretely by the reversed-phase column. The peptides are well resolved in three dimensions. **Figure 6*b*** illustrates that the three dimensions are orthogonal to one another. For any of the three axes correlated with one another, the effect would be that the data points lie within a plane rather than in 3D space. The figure shows that the points fill 3D space rather uniformly, demonstrating that these three separation modes are orthogonal. **Figure 6*b*** also illustrates that the peak capacity is smallest for the SEC dimension.

Coupling three 1D chromatographic separations together is an engineering feat that must be carefully planned. For the 3D peptide separation, each successive dimension has a progressively faster timescale: The SEC scale spans 80 min, the reversed-phase axis spans 1.5 min, and the capillary zone electrophoresis axis spans 27 s. This contrasts with the 2D gel, in which sampling and timing are not issues. Also, the mobile phases must be compatible in successive dimensions. To our knowledge, a 3D separation has not been used for proteins. A nice example of coupling two 1D separations of proteins uses two capillaries; in this study, Hu et al. (65) coupled micellar electrokinetic chromatography (based on hydrophobicity) and capillary zone electrophoresis (based on charge) for fast 2D separations in proteomics studies of single cells. In this example, the authors achieved very high sensitivity and low loss to enable single cells to be analyzed. New and well-grounded ideas, such as the ones presented here, to achieve multidimensional protein separations will ultimately enable the resolution of low-abundance proteins for mechanistic studies of diseases and biomarker discovery.

The limited peak capacity of SEC is an interesting problem because the same physical phenomenon gives rise to sieving separations in gels, yet sieving gels have peak capacities on the order of 50, whereas SEC can give a peak capacity as high as 10. The fundamentals of SEC explain why this is so. First, SEC has a smaller window in which the peaks must fit, and second, the peaks are broad compared to the width of this window. This can be explained from first principles. Recalling our above discussion of sieving, we give the equation for the partition coefficient in size exclusion here for the retention volume, V_R; the mobile-phase volume, V_0; and the total volume (mobile phase plus pores), V_T:

$$K = \frac{\text{accessible volume}}{\text{true volume}} = \frac{\pi\left(R^2 - r^2\right) \cdot L}{\pi R^2 \cdot L} = \frac{V_R - V_0}{V_T - V_0}. \qquad (10)$$

All the peaks in the chromatogram fit within the window bounded by V_0 and V_T. The largest proteins, which do not fit into the pores, elute at V_0, and the lowest–molecular weight species enter these pores without any size exclusion and elute at V_T. To estimate how many peaks fit between these two extremes, we consider a typical set of column parameters. If a column is randomly packed with spheres as the stationary phase, $V_0 = 40\%\ V_{sphere}$ (66). The maximum possible pore volume equals the sphere volume, which gives $V_T = V_0 + V_{sphere}$. In this limiting case, the range of retention volumes is V_0 to $V_0 \times 2.5$. More typically, the range is V_0 to $V_0 \times 1.5$ (67). Retention volumes larger than the theoretical maximum can be observed if adsorption occurs, which is irrelevant to this retention mechanism. The peak capacity, n, neglecting statistical

overlap, is the ratio of this range to the average peak width, 4σ. Converting retention volume to retention time makes for a more convenient expression:

$$n = \frac{1.5 \times V_0 - V_0}{4\sigma_V} = \frac{0.5 \times t_0}{4\sigma_t}. \quad (11)$$

For pure size exclusion, let us suppose that $L = 10$ cm and $V_0 = 0.3$ mL. If the plate height is estimated to be 10 μm (which is about as good as one can do with SEC columns), we can calculate the peak width through the relation $H = \sigma^2/L$, where the width is 4σ. These conditions give a peak capacity of 5. This illustrates the problem with SEC: The window is so narrow that even for this high efficiency of $H = 10$ μm, there are few resolved peaks. One way to increase the number of resolved peaks would be to increase the column length to 25 cm, but the number of resolved peaks would go up only as $L^{1/2}$, which is a factor of 1.6. Another way would be to use smaller particles, 1.8 μm in diameter, if these were available, which would give a factor of 3 at the most by reducing H to an impressive 1 μm. There are not any anticipated factors at this time that would increase the peak capacity enough to catch up with sieving. Fundamentally, because the window is narrow relative to the peak width, the peak capacities of SEC are not likely to approach those of sieving electrophoresis or those of other HPLC techniques.

7. SUMMARY OF PROTEIN SEPARATION TECHNIQUES

The various techniques described above, among others, are shown in **Table 1**. The properties used by each technique are also shown. The additional techniques

Table 1 Protein separation techniques and their properties

Technique	Property			
	Size	Charge	Hydrophobicity	Affinity
PAGE	√			
IEF		√		
2D-PAGE[a]	√	√		
CZE	√[b]	√[b]		
RPLC			√	
SCX		√		
MudPIT[c]		√	√	
SEC	√			
IMAC				√
Microarrays				√

[a]Sequential combination of IEF and PAGE in a two-dimensional gel.
[b]Separation based on ratio of charge to size.
[c]Sequential combination of SCX and RPLC.
Abbreviations: 2D, two-dimensional; CZE, capillary zone electrophoresis; IEF, isoelectric focusing; IMAC, immobilized metal affinity chromatography; MudPIT, multidimensional protein identification technology; PAGE, polyacrylamide gel electrophoresis; RPLC; reversed-phase liquid chromatography; SCX; strong cation exchange; SEC, size exclusion chromatography.

presented include immobilized metal ion affinity chromatography (IMAC), strong cation exchange (SCX), and multidimensional protein identification technology (MudPIT). In IMAC the specific binding of histidine residues to metal ions is utilized as the separation principle, and proteins are typically genetically engineered to have a hexahistine tag for purification by IMAC. In SCX the proteins adsorb to cationic sites in a resin and are differentially eluted with step increases in ionic strength of the mobile phase.

8. COMBINATION OF SEPARATIONS

In proteomics, investigators often use 2D electrophoresis or, alternatively, 2D HPLC. Combining the strengths of diverse separation tools is a promising approach in proteomics. One illustration is the proteomics of glycoproteins. These species are interesting because altered glycosylation is the hallmark of cancer (68). Glycosylation is a common type of post-translational modification that attaches a carbohydrate to a protein. A novel approach for biomarker discovery uses this idea to analyze only the glycoproteins of serum samples, in which the salient concept is that the biomarker does not need to be a protein whose level has changed owing to the disease; instead, it is the structure of the carbohydrate attached to the protein that has changed (69). To search for proteins of altered carbohydrate structure, Zhao et al. (69) applied a combination of affinity extraction, RPLC, and protein microarrays to serum samples of patients with pancreatic cancer. They first used affinity extractions, removing the high-abundance serum proteins with a multiple antibody column and then used a lectin column to extract the glycoproteins. Lectins are proteins that selectively bind glycoproteins, and a mixture of lectins in the affinity column ensured that a variety of glycoproteins were extracted. The study of only the protein bound to the lectin column further reduces the amount of high-abundance proteins and simplifies the mixture to the relevant proteins.

Subsequent gradient elution RPLC of the lectin-extracted protein sample gave protein peaks that were rather well resolved. The authors used a column of nonporous 1.8-μm silica particles for this separation. The smaller particle diameter is needed to achieve a reasonable surface area for separation. This is because in chromatography the resolution, R_s, is proportional to surface area, A, in the limit of low surface area:

$$R_s = \frac{\sqrt{L/H}}{4}(\alpha - 1)\frac{K \cdot A/V_m}{1 + K \cdot A/V_m}. \quad (12)$$

This equation strictly holds for isocratic elutions, but we can think of the gradient elution as having a progressively changing K for each protein as mobile-phase composition is varied. The nonporous particles provide a lower plate height by eliminating broadening due to intraparticle diffusion. The other terms are the selectivity, α, and the volume of the mobile phase, which are comparable to conventional gradient elution RPLC. Their minimizing H while maintaining sufficient surface area enabled approximately 24 protein peaks to be observed in RPLC.

The researchers collected fractions from the RPLC separation to isolate the glycoproteins, which were then spotted onto a substrate to make a glycoprotein array. It

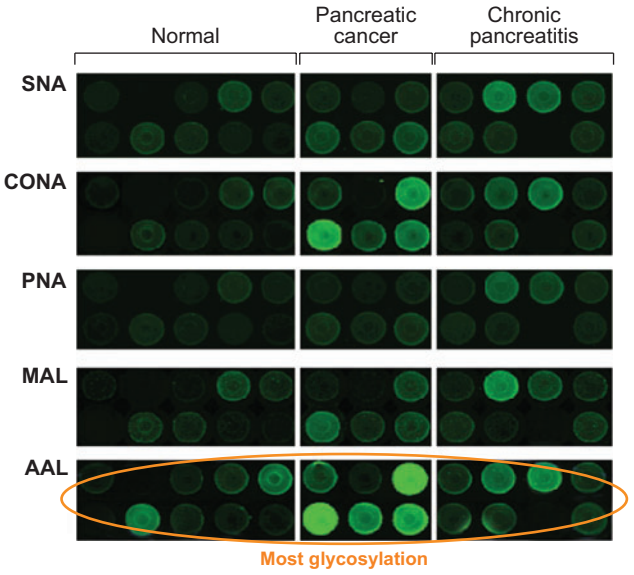

Figure 7

Protein array using five different lectins (SNA, CONA, PNA, MAL, and AAL) to detect glycoproteins in reversed-phase liquid chromatography (RPLC) fractions of the same retention time for 24 different patient serum samples. AAL detects the most glycosylation, and the samples from cancer patients are the most extensively glycosylated. Figure taken with permission from Zhao et al. (69).

is called a reverse protein array when the analytes, rather than the capture agents, are spotted onto the substrate. The protein array was probed by labeled lectins, with each type of lectin chosen to recognize a different type of carbohydrate. **Figure 7** shows a section of the protein array corresponding to one protein fraction from the serum samples of 24 patients, probed by five different lectins. One can see that the six samples from cancer patients typically had more glycosylation than either the ten samples from healthy patients or the eight samples from patients with pancreatitis. For this protein fraction, the most glycosylation was observed for the lectin AAL (*Aleuria aurentia*), which selects for fucosylation. The microarray revealed multiple proteins whose fucosylation was altered. Mass spectrometry of the corresponding RPLC protein fractions enabled identification of the proteins, showing that biomarkers originate from the liver rather than the pancreas, apparently owing to the inflammatory reaction associated with pancreatic cancer. Higher-sensitivity detection might enable the identification of pancreatic proteins. This study illustrates the use of a strategic combination of a wider variety of separations tools, and this approach could have a significant impact on biomarker discovery and mechanistic studies of disease if even higher sensitivity can be achieved.

9. CONCLUSIONS

The era of proteomics has placed large demands on the analysis of highly complex protein samples, requiring multidimensional separations coupled with mass spectrometry. The modern evolution of separations toward both higher speed and higher resolution through miniaturization has improved proteomics mainly in the case of chromatographic separations, in which particles smaller than 2 μm provide sharper peaks and faster elution. Immunodepletion columns and protein microarrays are new

manifestations of antibody-antigen interactions, and these provide valuable new tools for proteomics. Gel electrophoresis has evolved little, and this is one area in which breakthrough might have the most impact. Higher resolution in all types of protein separations is a general need for proteomics.

DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by NIH R01 GM65980. We thank Professor George R. Rossman of the California Institute of Technology for kindly providing the images for **Figure 2**.

LITERATURE CITED

1. Lepage R, Albert C. 2006. Fifty years of development in the endocrinology laboratory. *Clin. Biochem.* 39:542–57
2. Yalow RS, Berson SA. 1959. Assay of plasma insulin in human subjects by immunological methods. *Nature* 184:1648–49
3. Woof JM, Burton DR. 2004. Human antibody–Fc receptor interactions illuminated by crystal structures. *Nat. Rev. Immunol.* 4:89–99
4. Hu GQ, Gao YL, Li DQ. 2007. Modeling micropatterned antigen-antibody binding kinetics in a microfluidic chip. *Biosens. Bioelectron.* 22:1403–9
5. Glaser RW, Hausdorf G. 1996. Binding kinetics of an antibody against HIV p24 core protein measured with real-time biomolecular interaction analysis suggest a slow conformational change in antigen p24. *J. Immunol. Methods* 189:1–14
6. Tombelli S, Minunni M, Mascini M. 2007. Aptamers-based assays for diagnostics, environmental and food analysis. *Biomol. Eng.* 24:191–200
7. Zhu H, Snyder M. 2001. Protein arrays and microarrays. *Curr. Opin. Chem. Biol.* 5:40–45
8. MacBeath G, Schreiber SL. 2000. Printing proteins as microarrays for high-throughput function determination. *Science* 289:1760–63
9. Zhou FX, Bonin J, Predki PF. 2004. Development of functional protein microarrays for drug discovery: progress and challenges. *Comb. Chem. High Throughput Screen.* 7:539–46
10. Liu YS, Li CM, Yu L, Chen P. 2007. Optimization of printing buffer for protein microarrays based on aldehyde-modified glass slides. *Front. Biosci.* 12:3768–73
11. Melton L. 2004. Protein arrays: proteomics in multiplex. *Nature* 429:101–7
12. Angenendt P, Glokler J, Murphy D, Lehrach H, Cahill DJ. 2002. Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal. Biochem.* 309:253–60

13. Lee BH, Nagamune T. 2004. Protein microarrays and their applications. *Biotechnol. Bioprocess Eng.* 9:69–75
14. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, et al. 2004. Self-assembling protein microarrays. *Science* 305:86–90
15. Wu P, Grainger DW. 2006. Comparison of hydroxylated print additives on antibody microarray performance. *J. Proteome Res.* 5:2956–65
16. Kusnezow W, Syagailo YV, Ruffer S, Baudenstiel N, Gauer C, et al. 2006. Optimal design of microarray immunoassays to compensate for kinetic limitations: theory and experiment. *Mol. Cell. Proteomics* 5:1681–96
17. Huang RP. 2007. An array of possibilities in cancer research using cytokine antibody arrays. *Expert Rev. Proteomics* 4:299–308
18. Sachdeva N, Yoon HS, Oshima K, Garcia D, Goodkin K, Asthana D. 2007. Biochip array-based analysis of plasma cytokines in HIV patients with immunological and virological discordance. *Scand. J. Immunol.* 65:549–54
19. Zheng SP, Zhang H, Ross E, Van Le T, Wirth MJ. 2007. Silica colloidal crystals for enhanced fluorescence detection in microarrays. *Anal. Chem.* 79:3867–72
20. Jiang P, Bertone JF, Hwang KS, Colvin VL. 1999. Single-crystal colloidal multilayers of controlled thickness. *Chem. Mater.* 11:2132–40
21. Davis JM, Giddings JC. 1983. Statistical theory of component overlap in multicomponent chromatograms. *Anal. Chem.* 55:418–24
22. Davis JM, Giddings JC. 1985. Statistical method for estimation of number of components from single complex chromatograms: application to experimental chromatograms. *Anal. Chem.* 57:2178–82
23. Giddings JC. 1995. Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation. *J. Chromatogr. A* 703:3–15
24. Morita A, Miyagi E, Yasumitsu H, Kawasaki H, Hirano H, Hirahara F. 2006. Proteomic search for potential diagnostic markers and therapeutic targets for ovarian clear cell adenocarcinoma. *Proteomics* 6:5880–90
25. Queloz PA, Crettaz D, Thadikkaran L, Sapin V, Tissot JD. 2006. Two-dimensional gel electrophoresis based technologies for potential biomarkers identification in amniotic fluid: a simple model. *Protein Pept. Lett.* 13:959–63
26. Huang HL, Stasyk T, Morandell S, Dieplinger H, Falkensammer G, et al. 2006. Biomarker discovery in breast cancer serum using 2-D differential gel electrophoresis/MALDI-TOF/TOF and data validation by routine clinical assays. *Electrophoresis* 27:1641–50
27. Hoffman SA, Joo WA, Echan LA, Speicher DW. 2007. Higher dimensional (Hi-D) separation strategies dramatically improve the potential for cancer biomarker detection in serum and plasma. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 849:43–52
28. Huang L, Harvie G, Feitelson JS, Gramatikoff K, Herold DA, et al. 2005. Immunoaffinity separation of plasma proteins by IgY microbeads: meeting the needs of proteomic sample preparation and analysis. *Proteomics* 5:3314–28
29. Yu KH, Rustgi AK, Blair IA. 2005. Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry. *J. Proteome Res.* 4:1742–51

30. Giddings JJ. 1991. *Unified Separation Science*. New York: Wiley & Sons
31. Liu Z, Lemma T, Pawliszyn J. 2006. Capillary isoelectric focusing coupled with dynamic imaging detection: a one-dimensional separation for two-dimensional protein characterization. *J. Proteome Res.* 5:1246–51
32. Yao B, Yang HH, Liang QL, Luo G, Wang LD, et al. 2006. High-speed, whole-column fluorescence imaging detection for isoelectric focusing on a microchip using an organic light emitting diode as light source. *Anal. Chem.* 78:5845–50
33. Reynolds JA, Tanford C. 1970. Binding of dodecyl sulfate to proteins at high binding ratios: possible implications for state of proteins in biological membranes. *Proc. Natl. Acad. Sci. USA* 66:1002–7
34. Smithies O. 1955. Grouped variations in the occurrence of new protein components in normal human serum. *Nature* 175:307–8
35. Raymond S, Weintraub L. 1959. Acrylamide gel as a supporting medium for zone electrophoresis. *Science* 130:711–11
36. Laurell CB. 1966. Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. *Anal. Biochem.* 15:45–52
37. Shapiro AL, Vinuela E, Maizel JV. 1967. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem. Biophys. Res. Commun.* 28:815–20
38. Rodbard D, Chrambach A. 1970. Unified theory for gel electrophoresis and gel filtration. *Proc. Natl. Acad. Sci. USA* 65:970–77
39. Ogston AG. 1958. The spaces in a uniform random suspension of fibres. *Trans. Faraday Soc.* 54:1754–57
40. Kopecka K, Drouin G, Slater GW. 2004. Capillary electrophoresis sequencing of small ssDNA molecules versus the Ogston regime: fitting data and interpreting parameters. *Electrophoresis* 25:2177–85
41. Holmes DL, Stellwagen NC. 1991. Estimation of polyacrylamide-gel pore size from Ferguson plots of linear DNA fragments. 2. Comparison of gels with different cross-linker concentrations, added agarose and added linear polyacrylamide. *Electrophoresis* 12:612–19
42. Pernodet N, Maaloum M, Tinland B. 1997. Pore size of agarose gels by atomic force microscopy. *Electrophoresis* 18:55–58
43. Suzuki A, Yamazaki M, Kobiki Y. 1996. Direct observation of polymer gel surfaces by atomic force microscopy. *J. Chem. Phys.* 104:1751–57
44. Acevedo F, Marin V, Wasserman M. 1995. Electrophoretic size separation of proteins treated with sodium dodecyl sulfate in 1% agarose gels. *Electrophoresis* 16:1394–400
45. Tietz D, Gottlieb MH, Fawcett JS, Chrambach A. 1986. Electrophoresis on uncrosslinked polyacrylamide: molecular sieving and its potential applications. *Electrophoresis* 7:217–20
46. Zhou D, Wang YM. 2006. Mechanisms of DNA separation by capillary electrophoresis in nongel sieving matrices. *Prog. Chem.* 18:987–94
47. Chu BJ, Liang DH. 2002. Copolymer solutions as separation media for DNA capillary electrophoresis. *J. Chromatogr. A* 966:1–13

48. Best N, Arriaga E, Chen DY, Dovichi NJ. 1994. Separation of fragments up to 570 bases in length by use of 6′5 T-non-cross-linked polyacrylamide for DNA sequencing in capillary electrophoresis. *Anal. Chem.* 66:4063–67
49. Ganzler K, Greve KS, Cohen AS, Karger BL, Guttman A, Cooke NC. 1992. High-performance capillary electrophoresis of SDS protein complexes using UV-transparent polymer networks. *Anal. Chem.* 64:2665–71
50. Wu D, Regnier FE. 1992. Sodium dodecyl sulfate capillary gel electrophoresis of proteins using non-cross-linked polyacrylamide. *J. Chromatogr.* 608:349–56
51. Dovichi NJ, Zhang JZ. 2000. How capillary electrophoresis sequenced the human genome. *Angew. Chem. Int. Ed. Engl.* 39:4463–68
52. Bousse L, Mouradian S, Minalla A, Yee H, Williams K, Dubrow R. 2001. Protein sizing on a microchip. *Anal. Chem.* 73:1207–12
53. Zhang H, Wirth MJ. 2005. Electromigration of single molecules of DNA in a crystalline array of 300-nm silica colloids. *Anal. Chem.* 77:1237–42
54. Zeng Y, Harrison DJ. 2007. Self-assembled colloidal arrays as three-dimensional nanofluidic sieves for separation of biomolecules on microchips. *Anal. Chem.* 79:2289–95
55. Zheng SP, Ross E, Legg MA, Wirth MJ. 2006. High-speed electroseparations inside silica colloidal crystals. *J. Am. Chem. Soc.* 128:9016–17
56. Le TV, Ross EE, Velarde TRC, Legg MA, Wirth MJ. 2007. Sintered silica colloidal crystals with fully hydroxylated surfaces. *Langmuir* 23:8554–59
57. Kirkland JJ, Kohler J. 1989. *U.S. Patent No. 4,874,518*
58. Chabanov AA, Jun Y, Norris DJ. 2004. Avoiding cracks in self-assembled photonic band-gap crystals. *Appl. Phys. Lett.* 84:3573–75
59. Mihi A, Ocana M, Miguez H. 2006. Oriented colloidal-crystal thin films by spin-coating microspheres dispersed in volatile media. *Adv. Mater.* 18:2244–49
60. Kirkland JJ. 1971. High speed liquid-partition chromatography with chemically bonded organic stationary phases. *J. Chromatogr. Sci.* 9:206–14
61. Krokhin OV, Craig R, Spicer V, Ens W, Standing KG, et al. 2004. An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol. Cell. Proteomics* 3:908–19
62. Koyama J, Nomura J, Shiojima Y, Ohtsu Y, Horii I. 1992. Effect of column length and elution mechanism on the separation of proteins by reversed-phase high-performance liquid chromatography. *J. Chromatogr.* 625:217–22
63. Bushey MM, Jorgenson JW. 1990. Automated instrumentation for comprehensive two-dimensional high-performance liquid-chromatography capillary zone electrophoresis. *Anal. Chem.* 62:978–84
64. Moore AW, Jorgenson JW. 1995. Comprehensive three-dimensional separation of peptides using size-exclusion chromatography reversed-phase liquid-chromatography optically gated capillary zone electrophoresis. *Anal. Chem.* 67:3456–63
65. Hu S, Michels DA, Fazal MA, Ratisoontorn C, Cunningham ML, Dovichi NJ. 2004. Capillary sieving electrophoresis/micellar electrokinetic capillary chromatography for two-dimensional protein fingerprinting of single mammalian cells. *Anal. Chem.* 76:4044–49

66. Schure MR, Maier RS. 2006. How does column packing microstructure affect column efficiency in liquid chromatography? *J. Chromatogr. A* 1126:58–69
67. Yau WW, Ginnard CR, Kirkland JJ. 1978. Broad-range linear calibration in high-performance size-exclusion chromatography using column packings with bimodal pores. *J. Chromatogr.* 149:465–87
68. Gerber-Lemaire S, Juillerat-Jeanneret L. 2006. Glycosylation pathways as drug targets for cancer: glycosidase inhibitors. *Mini-Rev. Med. Chem.* 6:1043–52
69. Zhao J, Patwa TH, Qiu WL, Shedden K, Hinderer R, et al. 2007. Glycoprotein microarrays with multi-lectin detection: unique lectin binding patterns as a tool for classifying normal, chronic pancreatitis and pancreatic cancer sera. *J. Proteome Res.* 6:1864–74