Integrating Cytomics and Proteomics*

Tytus Bernas‡, Gérald Grégori§, Eli K. Asem¶, and J. Paul Robinson‡‖

Systems biology along with what is now classified as cytomics provides an excellent opportunity for cytometry to become integrated into studies where identification of functional proteins in complex cellular mixtures is desired. The combination of cell sorting with rapid protein-profiling platforms offers an automated and rapid technique for greater clarity, accuracy, and efficiency in identification of protein expression differences in mixed cell populations. The integration of cell sorting to purify cell populations opens up a new area for proteomic analysis. This article outlines an approach in which well defined cell analysis and separation tools are integrated into the proteomic programs within a core laboratory. In addition we introduce the concepts of flow cytometry sorting to demonstrate the importance of being able to use flow cytometry as a cell separation technology to identify and collect purified cell populations. Data demonstrating the speed and versatility of this combination of flow cytometry-based cell separation and protein separation and subsequent analysis, examples of protein maps from purified sorted cells, and an analysis of the overall procedure will be shown. It is clear that the power of cell sorting to separate heterogeneous populations of cells using specific phenotypic characteristics increases the power of rapid automated protein separation technologies. ***Molecular & Cellular Proteomics 5:2–13, 2006.***

A cell is the smallest living unit in any organism, and its structure and morphology often determine function. Under different conditions a cell may or may not express a functional marker. Although an identified gene might well be recognized for carrying the code for specific proteins, the regulated expression of that gene leads to the production of proteins associated with that phenotype. The translation from genome to proteome is necessary to understand the functional level of the cytome. *Cytomes* can be defined as cellular systems and subsystems and functional components of the organism. Cytomics is the study of the role of the cell within the context of genomic and proteomic discoveries. It is defined as the study of the heterogeneity of *cytomes* or more precisely the study of molecular single-cell phenotypes resulting from genotype and exposure in combination with exhaustive bioinformatic knowledge extraction (1).

Current approaches to understanding the functional diversity of an organism preferentially strive for a systems approach whereby first the phenotypic classification of a specific cytome is achieved prior to an attempt to perform proteomic analysis. The fundamental basis for this approach has been well established in studies of cellular systems over many years, and when combined with advanced proteomic approaches, this approach can achieve rapid and specific identification of a direct link between biomarkers and their functional roles in complex organisms.

Proteomics refers to the study of the proteome, which is the total protein complement of a genome. It is presently considered to be the key technology in the global analysis of protein expression and in the understanding of gene function and regulation in the postgenomic era. The scope of proteomics is broad; it encompasses identification and quantitation of proteins in cells, tissues, and biological fluids; analysis of changes in protein expression in normal *versus* diseased cells or between cells grown under different conditions; characterization of post-translational modifications; and studies of protein-protein interactions. The goals of proteomic research include clarification of molecular mechanisms that govern cellular processes, characterization of complex protein networks and their perturbations, discovery of biomarker proteins for detection and diagnosis of diseases, and identification of targets for the design of drug treatments (2, 3). Many different technologies have been and are still being developed to collect the information contained in the properties of proteins (Fig. 1).

The following three characteristics of these technologies are apparent. First, no single platform can accomplish the desired proteomic measurements. Second, the closer the measurement is to protein function, the less mature the technology. Third, "true" proteomic technology has yet to reach maturity.

In this article we propose the integration of a very mature cell identification and separation technology with a rapidly developing protein separation technology. The combination of these approaches provides a uniquely sensitive and accurate approach to protein profiling in biological systems. It is reproducible, relatively rapid, and robust, three valuable assets in any problem-solving modality.

THE CYTOMIC APPROACH TO PROTEOMICS

The driving force for integrating cytomics into proteomics is the need to generate pure populations of cells from hetero-

From the ‡Cytometry Laboratories and ¶School of Veterinary Medicine, Purdue University, West Lafayette, Indiana 47907-2064 and §Laboratoire de Microbiologie, Géochimie et Ecologíe Marines, CNRS UMR 6117, 163 Avenue de Luminy, Case 901, Batîment TPR1, 13288 Marseille cedex 9, France

Received, August 1, 2005, and in revised form, October 27, 2005
Published, MCP Papers in Press, October 28, 2005, DOI 10.1074/mcp.R500014-MCP200

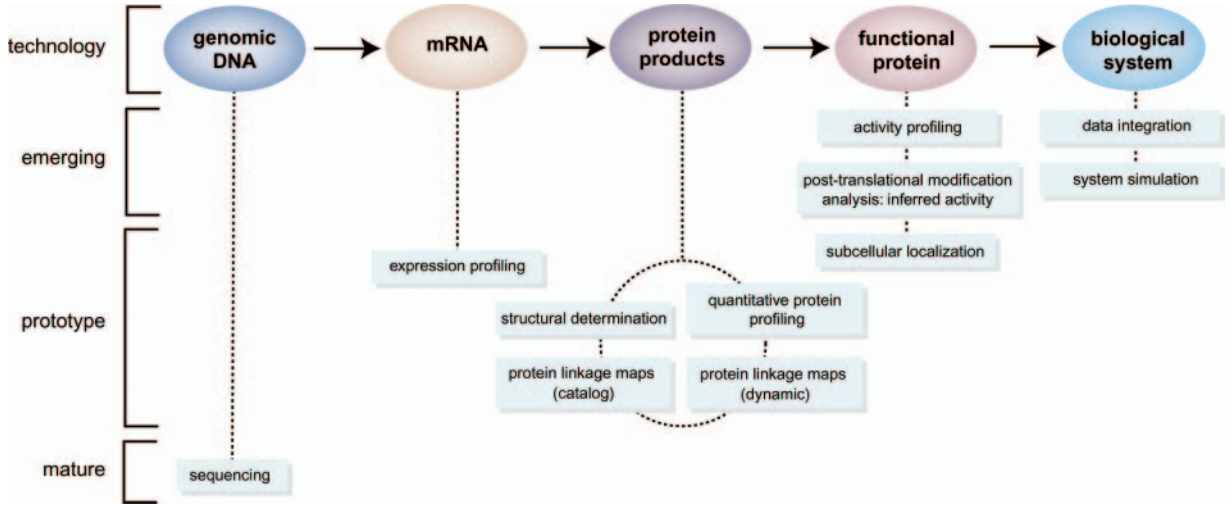

FIG. 1. **The current status of proteomic technologies.** The different data typically collected in proteomic research and the available technologies are listed. The relative maturity of the proteomic technologies and other key discovery science tools is apparent from the position of the respective technology on the graph. Reprinted with permission from Patterson and Aebersold (3) (Nature Publishing Group, www.nature.com).

geneous and highly integrated mixtures as are found in a majority of biological environments. The first component of this system is the use of very specific tools for single-cell analysis, specifically flow cytometry. Flow cytometry has been a mature technology for over 30 years, having been developed in the late 1960s and early 1970s.

Flow cytometry is a technology that has made a significant impact in cell biology and clinical medicine as well as in several other associated fields (*e.g.* microbiology and bioengineering). The principle of flow cytometry is that single particles suspended within a stream of liquid are illuminated within a short time (2–10 μs typically) while they pass through a light source focused within a very small region (also called the interrogation volume). Particles will then scatter light and may emit natural or induced fluorescence if they contain natural fluorochromes or are stained with fluorescent dyes. The optical signals emitted from the particles are then collected as spectral bands of light mostly within the visible spectrum (400 up to 900 nm). These signals represent the detection of various chemical or biological components principally based on fluorescence. Because flow cytometers can analyze single particles/cells, it is possible to separate particles/cells into clusters based upon a statistical difference of any of 5–20 variables that can be measured on the current flow cytometers for each particle/cell. It is possible to electronically separate these populations based on statistical analyses with subsequent identification using multivariate analytical techniques.

Most of the classification in flow cytometry is based on the use of fluorescent molecules to specifically label the particles/cells of interest. If the particle is a cell such as a white blood cell for example, the fluorescent probe might be membrane-bound or cytoplasmically bound, or it could be attached to nuclear material. Monoclonal or polyclonal antibodies that recognize specific receptors on cells are used as the primary differential markers because of their specificity. Both the location and number of these conjugated antibodies can be monitored by conjugating fluorescent molecules to them. Particles or cells of almost any nature can be evaluated by flow cytometry as long as they are within a reasonable size range. They can be very small, even below the resolution limits of visible light (*i.e.* viruses) because they can still be detected by their fluorescent signatures. Similarly, by modifying the nature of the fluidic system of the instrument it is possible to study and isolate particles as large as several thousand micrometers.

The real advantage of flow cytometry is that a very large number of particles can be evaluated in a very short time period. Some systems for example can analyze particles at rates up to 100,000 particles/s while collecting 10–20 simultaneous parameters from each particle. Finally, the principle of cell sorting in flow cytometry allows this technology to physically separate single particles/cells from mixed populations. Thus single particles can be physically deposited under sterile conditions onto a defined location for further analysis or culture. If the cell is a rare event (1:100,000 or more) it is possible to identify it and physically isolate this rare population, although it may be difficult to physically collect an adequate number of such rare cells in a reasonable time. Thus, it is clear that the technology of flow cytometry has a unique advantage in any system where heterogeneity exists because the collection of a high number of variables allows discrimination of populations with minute differences followed by the subsequent purification of these cells.

Over 20 years ago, Shapiro (4) noted that multiparameter flow cytometry was now a reality in the field because of the availability of commercial instruments. The field has since expanded well beyond anything that was considered possible

at that time. Today's instruments have the capacity to measure 10–20 spectral bands simultaneously together with a variety of scatter signals. The newest next generation instrument will be capable of full spectral analysis and classification (5). Advanced computational analysis makes it is possible to perform complex multiparametric analyses virtually instantaneously with adequate time to make sorting decisions after measurements are made. These types of analysis have been frequently used in the field of cytometry (6), and current computational capability is advancing the field rapidly. The result of this technology is that it is now possible to rapidly generate clinical diagnostic information from complex heterogeneous mixtures of samples such as human blood (7) and perform this in real time (6).

A Brief History of Flow Cytometry—The basic principles of flow cytometry are based on some very old ideas generated early in the 20th century using the principles of laminar flow defined previously by Reynolds (39) in the late 19th century. Some 50 years later, Moldavan (8) published the design of an instrument that could have identified single cells using a microscope and a photodetector, and this article is accepted as one of the earliest attempts at single-cell analysis. In the 1940s Papanicolaou and Traut (9) demonstrated that they could identify cancerous cells from cervical cancer by observing the staining patterns obtained by staining tissues with specifically designed stains. At this time, there were very few dyes that were established for bright-field microscopy and were used for identifying cellular abnormalities. There were no specific fluorescence dyes prior to 1941 when Coons *et al.* (10) developed the technique for directly labeling antibody with fluorescein. It subsequently proved quite difficult to create analytical technology based on the poor capability of computers and imaging technology, thus resulting in a movement toward single-cell analysis using flow cytometry as opposed to image processing and recognition technologies.

In the 1960s, Louis Kamentsky designed and built a single-cell analyzer. Kamentsky's work at IBM's Watson Laboratories involved using optical character recognition techniques to identify cancer cells. Because of the lack of computation power at the time, it was clear to Kamentsky[1] that his efforts were better focused on single-cell analysis and the design of a cytometer that measured absorption and scatter rather than an image-based technology (11). Shortly thereafter he added the ability to sort cells using fluidic switching (12). During the same time period, Fulwyler was trying to solve a problem generated by the study of red blood cells using a single-cell analysis system at Los Alamos. Because of the use of impedance measurement using the newly developed Coulter volume, an observation had been made that there was a bimodal distribution of red blood cells. This suggested to some that two different types of red blood cells were being detected. Fulwyler recognized that physically separating these "different" cells was necessary. He had previously seen Richard Sweet's developments of high speed chart recorders using electrostatic drop generation (13) and pursued this technology at Los Alamos under the mentorship of Marvin Van Dilla. Fulwyler (14) visited Richard Sweet's laboratory and modified his newly invented ink jet recording technology to design and build a cell sorter to separate red blood cells. Once the instrument had been designed and tested, it was rapidly concluded that the bimodal distribution was related to spatial orientation rather than inherent red blood cell variability.[2] Amazingly this finding of great significance was never formally published, and it was immediately obvious that sorting of white blood cells was an opportunity to be investigated. The complete history of the development of cell sorting is covered by Shapiro (15, 16). Thus, the development of the cell sorter became a reality in 1965.

Flow Cytometry Principles—Flow cytometers operate only with particles in liquid suspensions. Cells must be in a single cell form rather than aggregated or in tissue before they are analyzed one at a time. Most flow cytometers operate by using a sheath flow system. Cells are suspended in one medium, whereas another medium is used as sheath liquid. Both can be identical fluids such as saline. If cells are to be sorted on an electrostatic sorter the sheath fluid must indeed be composed of an ionized liquid such as saline to conduct electrical charges as described above. There are other sorting techniques based on a mechanical system that do not operate this way, but they are rare and mostly very slow (less than 300 cells/s).

Once the cells are in suspension, they are injected into the flow cytometer through a nozzle. The sample is either driven into the instrument by air pressure applied on top of the sample or injected into the instrument by a syringe (Hamilton type) controlled by a precision motor. Particles are then driven to the flow cell, which is the keystone of the flow cytometer (Fig. 2).

The sheath fluid fills the flow cell and progressively accelerates as it goes through the conic tip of the flow cell. The particles are injected within the flow cell at the very base of this conic area. As soon as a particle emerges from the injection nozzle it is immediately accelerated and centered within the surrounding fluid sheath. This phenomenon, also called hydrodynamic focusing, is responsible for the separation, alignment, and centering of the particles within the jet made by the sheath fluid surrounding the sample core. Every single particle is then intercepted as it flows by a fixed light source, usually a laser beam (Fig. 3).

In most flow cytometers two types of scatters are recorded: forward-angle and right-angle scatter. Forward-angle light scatter is the amount of light scattered at small angle ($<10°$) in the forward direction with respect to the direction of the laser beam. Forward-angle light scatter is known to be related to the size of the particle.

[1] L. Kamentsky, personal communication.

[2] M. J. Fulwyler, personal communication.

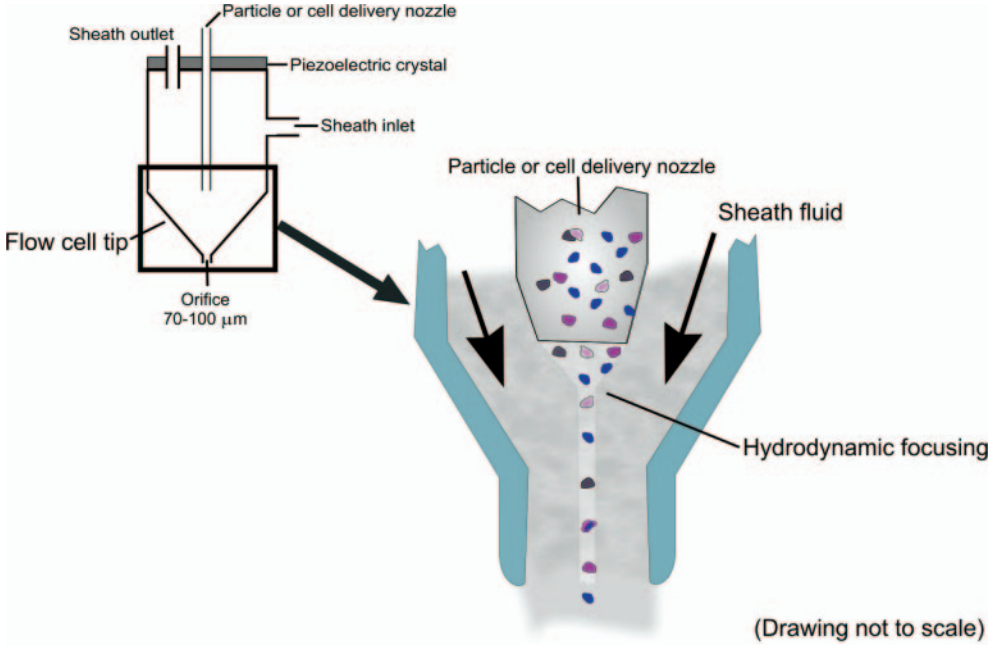

FIG. 2. **Scheme of the flow cell and the principle of hydrodynamic focusing whereby cells are aligned in single file for subsequent analysis by laser-based scatter and fluorescence detection.**

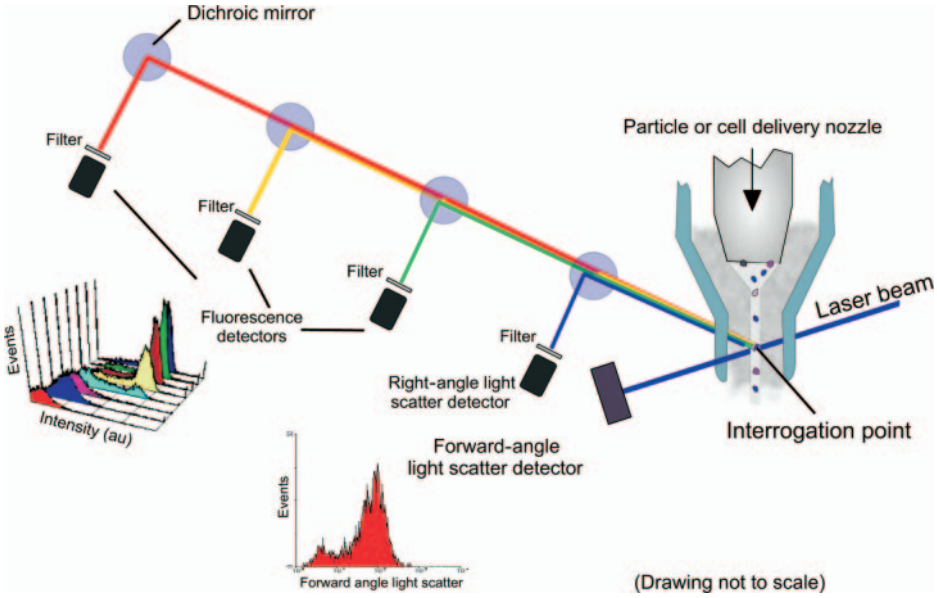

FIG. 3. **As single particles pass through an excitation source (laser beam), emission is detected by a series of photodetectors.** Shown is a general scheme of flow cytometer optics. Scattered light and specific fluorescence emission bands are separated at individual detectors by emission filters.

The photons scattered at 90° (or right-angle light scatter) with respect to the incident laser beam are also recorded. The right-angle light scatter intensity depends mostly upon the cell structure and granularity.

The cell fluorescence emitted after excitation by the laser beam is also recorded. This fluorescence can be natural (autofluorescence) from molecules such as photosynthetic pigments or can be induced by a fluorescent dye added to the sample. This dye emits a fluorescence signal when excited at the proper wavelength and thus helps to identify cells of interest by identifying some phenotypic characteristic.

Scatter and fluorescence signals are separated from each other and sent to specific photodetectors by a set of dichroic mirrors and filters that split the optical signals in spectral wavelength ranges (Fig. 3). For each measured parameter, data are then acquired on a personal computer. Thus flow cytometry provides a multiparametric analysis at the single-cell level.

Data can be displayed in single dimension (histogram) or in two dimensions (cytogram). A histogram displays a frequency distribution of the particle for any particular parameter, whereas a cytogram displays the data in two dimensions.

Histograms and cytograms will represent the population distribution of fluorescence and scatter signals received from a large number of cells or particles (several tens of thousands). Advanced statistical evaluation of these data is crucial to the establishment of separation criteria for sorting or analysis.

Multiple fluorescent dyes can be used simultaneously in a flow cytometer. This allows for what is termed "multiparameter analysis." Using this complex capability, it is possible to separate very complex mixtures of cells or particles so that particles of very specific nature can be identified electronically and then physically sorted as discussed below. An example of several fluorescence histograms is shown in Fig. 4. In this figure, it can be seen that there is a relatively complex association of multiple dyes so that it is necessary to apply advanced mathematical analysis to separate the actual fluorescence signals from each dye from the mixed signals.

Cell Sorting—The principle of cell sorting as already mentioned was developed by Fulwyler (14) to sort red cells. Kamentsky (12) and Dittrich and Gohde (17) also devised cell separation technologies to analyze cells of interest. The electrostatic technology has matured and evolved into the technique of choice for virtually all current commercial cell sorters. Much of the technology related to fluorescence detection of cellular populations was implemented by Bonner *et al.* (18) in the early 1970s.

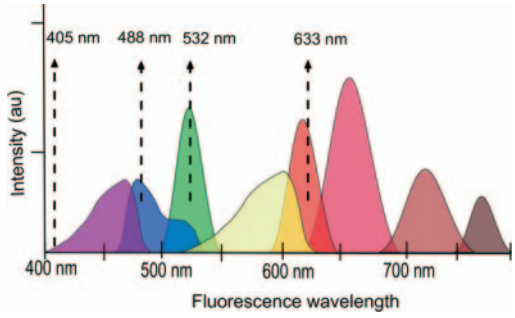

FIG. 4. **Example of spectral overlap occurring when several fluorescent organic dyes are used simultaneously.** *Lines* indicate common excitation wavelengths of available laser sources used in flow cytometry. *au*, absorbance units.

The principle of electrostatic sorting is based on the ability to identify a cell of interest, identify its physical position in the jet with a high degree of accuracy, break the jet into droplets, place a charge on the stream at a precise time to charge the droplet containing the cell of interest, and finally deflect the charged droplets containing the desired cell into a vessel (Fig. 5).

van den Engh recently defined the technology of high speed sorting (19) and discussed in detail the complex issues involved. The speed and accuracy of a sorting are based on several key factors. First, for cell sorting the stream must be vibrated using a piezoelectric device to generate droplets. But as mentioned in the initial discussion a fully stable laminar flow is still required for accurate analysis. It is thus mandatory to match the nozzle diameter, sheath flow rate (pressure), and droplet generation frequency as well as high speed electronics to obtain stable droplet generation and thus high speed cell sorting. Kachel *et al.* (20) characterized the principle that governs the generation of droplets whereby the wavelength of the undulations is $\lambda = v/f$ where λ is the undulation wavelength, v the stream velocity, and f is the modulation frequency.

The system is optimized for maximum droplet generation when $\lambda = 4.5d$ (d = exit orifice = stream diameter). Thus the

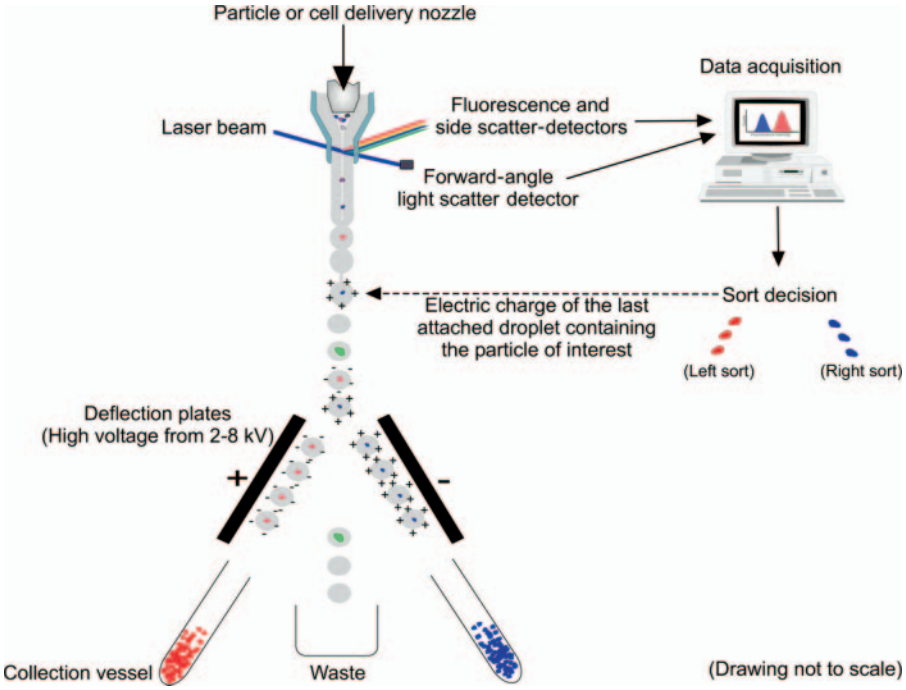

FIG. 5. **The principle of electrostatic cell sorting based on droplet formation.** Charged droplets containing particles or cells are deflected to specific containers. The criteria for deflection is based on the phenotypic characterization defined previously. Pure populations of individual phenotypes can thus be obtained.

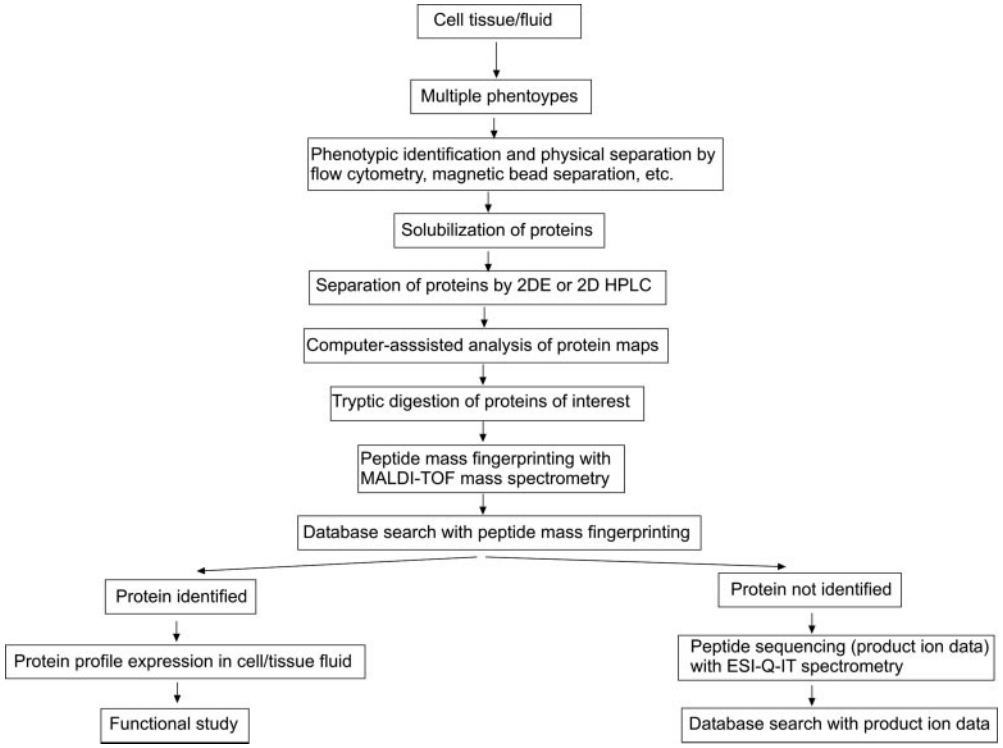

FIG. 6. **General strategy for proteome analysis by 2D gel electrophoresis or chromatography, MS, and database searching.** Q-IT, quadrupole-ion trap.

optimal generation frequency is given by $f = v/4.5d$. If a system can accommodate this optimal droplet formation, as demonstrated by Pinkel and Stovel (21), the jet velocity is proportional to the square root of the jet pressure. Thus, if an optimal system sorts 20,000 events/s (20,000 Hz) such that each drop is separated by 4.5 stream diameters and flowing at 10 m/s it means that the drops are 200 μm apart from each other. Of course the diameter of the droplets must decrease as the number of droplets generated increases, with the obvious conclusion that the speed of high speed sorters will eventually be partially regulated by the size of the particle to be sorted and the velocity to which the stream can be taken without damaging the particles/cells. This is particularly important for fragile cells. In general, however, regular sorting procedures do not inflict damage to cells. Over the past 40 years, cell sorting has been used to isolate functionally normal cells. This feature of flow cytometry is critical to ensuring that isolation of a phenotypically distinct population for protein profiling did not preferentially damage certain cells with the result that the profiled population represented only robust cells.

High speed sorters are essentially sorters that are designed to operate at sort speeds in the range of 20,000–100,000 particles/s. To achieve this sorting rate, higher pressures must be applied on both sample and sheath streams. When exceeding a sorting rate of 40,000 cells/s, the key issue becomes analysis time, which becomes the limiting factor because complex analysis must precede the sorting decision. The primary issue is thus the high pressure that must be used to create very high speed droplet formation as mentioned in van den Engh (19). For instance, generating droplet frequency of 250,000/s requires a jet pressure approaching 500 p.s.i., a significantly higher value than can be designed safely in most systems. Thus, for limits of around 100 p.s.i., a droplet rate of around 100,000/s is closer to the realistic maximum. This then is the real limitation to current high speed sorting systems.

As mentioned earlier, flow cytometry addresses analysis of single particles. It means that particles must flow in single file through the interrogation point. Thus it is crucial to ensure that no measurements take place when two cells or more pass through the interrogation point at or near the same time. Poisson statistics enter the equation at this point because it is impossible to predict exactly when any particle will pass the interrogation point. There is, however, a relationship between particle concentration and coincident detection events. Based on current available techniques it is possible to sort and purify a very large number of cells (10^6–10^8) from complex mixtures in a reasonable time (hours). Flow cytometry is a mature, well defined technology and is ideal for separation of mixtures of cells where complex phenotypes exist.

PRINCIPLES OF PROTEOMICS

Although numerous variations have been proposed in the field, a typical proteomic study comprises the following steps (Fig. 6).

TABLE I
Overview of protein separation/fractionation technologies

Technology	Optimal protein mass range	Detection dynamic range	Number of fractions/spots	Technical issues
	kDa			
2DE	10–200 (proteins)	1000	3000	Low replicability and difficult quantization
Automated 2D HPLC	>5 (proteins and peptides)	100	2500	Limited to UV absorption detection (unless coupled with MS)

- Solubilization of proteins from the sample (tissue or cell population).
- Fractionation of the protein mixture using 2D[3] gel electrophoresis (2DE) or multidimensional LC.
- Computer-assisted analysis of protein patterns (maps) and identification of the proteins of interest by MS and searching of databases MS spectra.

We will discuss briefly these steps to outline limitations of current proteomic technologies and possibilities of its expansion using a cytomic approach.

Solubilization—Treatment of biological sample involves cell lysis or another disruptive step to yield a suspension of cells, organelles, or fragments of subcellular structures. Proteins are extracted from this suspension with a combination of chemicals (22, 23). Detergents (*e.g.* CHAPS, SDS, and Tween) are used to aid solubilization of membrane proteins and their separation from liquids. Chaotropes (*e.g.* urea and guanidine) are added to solubilize hydrophobic proteins. Reductants (*e.g.* dithiothreitol and mercaptoethanol) help to prevent oxidation and break disulfide bonds. Enzymes (*e.g.* DNase and RNase) may be used to decompose contaminating nucleic acids, lipids, and carbohydrates. Protease inhibitors are added to prevent protein degradation by endogenous proteases. The resulting extract may represent the total protein composition of the analyzed biological sample. Alternatively the extraction step may be used to prefractionate the protein mixture and eliminate unwanted components. An aliquot of the extract is subjected to separation (or fractionation) using 2DE or LC.

Protein Separation and Fractionation—2DE is a powerful separation technique that allows simultaneous resolution of thousands of proteins. The high resolution capability of 2DE stems from the fact that the first and second dimensions are based on two independent protein characteristics. The first dimension of 2DE is IEF during which the proteins are separated based on their isoelectric point. In the second dimension, the proteins are separated orthogonally by SDS-PAGE according to their molecular weight. After separation, proteins in 2D gels are visualized by staining, commonly with a Coomassie Blue stain, or with a modified silver stain that is compatible with subsequent MS analysis (24, 25). The 2D gels are digitized, and the resulting gel images are qualitatively and quantitatively analyzed with specialized software programs. In this manner, proteins can be quantified, and spot patterns in multiple gels can be matched and compared. Statistical analysis can be performed on groups of features (spots) in sets of gels, and variations, differences, and similarities can be evaluated. LC constitutes an alternative to the 2DE technique of separation of protein and peptide mixtures. The most widely used variant is HPLC. The advantage of this technique is the diversity of separation modes available. These include the following: reverse phase and hydrophobic interaction chromatography, which separates proteins according to their hydrophobicity; strong cation and anion exchange, which takes advantage of the net positive or negative charge to separate proteins; and size exclusion, which separates proteins according to their molecular weight (size). Different separation modes can be combined in series to give greater resolution of protein mixtures. This "tandem HPLC" is one of the most effective tools in analytical proteomics. During the separation proteins are detected with UV absorbance. A summary of the HPLC and gel electrophoresis protein separation techniques is given in the Table I.

Identification—Proteins resolved by 2DE or HPLC can be identified based on unique attributes that are measured by MS. These attributes are determined from analysis of peptides generated by proteolytic digestion of the protein of interest. The most commonly used enzyme for protein digestion is trypsin, which cleaves the protein at the C-terminal side of lysine and arginine. Two specific protein attributes can be obtained by MS analyses of proteolytic digests. The first protein attribute is the so-called peptide mass fingerprint. Peptide mass fingerprinting involves determination of the masses of all peptides in the digest. The second attribute includes fragmentation of selected peptides inside the mass spectrometer into series of sequence-diagnostic product ions. From these product ions, a portion of the amino acid sequence of the peptide (a "sequence tag") can be deduced; alternatively, uninterpreted product ion spectra can be used directly for protein identification.

Problems—Proteomic analysis with either HPLC or 2DE involves extraction of protein content from a population of cells. One should note that because 90% or so of the protein mass in a given cell type comes from only 10% of the pro-

[3] The abbreviations used are: 2D, two-dimensional; 2DE, two-dimensional gel electrophoresis.

teins, supersensitive technologies are required to even detect those proteins whose levels are very low but whose biological potency may be high, such as G-protein-coupled receptors, transcription factors, and kinases (26). Many proteins in humans, mice, and flies can be reduced to a few percent of their wild-type levels without debilitating phenotypic consequences, and the amount of an individual protein can vary significantly between different wild type individuals. The extent to which this natural between-individual variation contributes to phenotype is not well documented. Consequently taking into account resolution of fractionation techniques at least 10^6 cells are necessary to construct a reliable proteome profile. However, as the cells are lysed, it is obvious that the readout of these experiments is an average for protein activation states across the cell population(s). Such averaging may obscure significant biology, such as is likely in cancer. This disease, although often classified as genetic, is, in a functional sense, a proteomic one; genetic mutations can modify protein signaling pathways and thereby create a survival advantage for the cell because they force it to ignore negative inhibitory signals or perpetually send it false positive signals.

One should note that tumor tissues are complex systems where different cell types and proteins are present. For instance, solid tumors may comprise serum proteins, blood vessels and blood cells, and mesenchymal, endothelial, and epithelial cells as well as necrotic material (27). Furthermore development of some tumors, like liver fibrosis, may depend on interaction of certain cell types with extracellular matrix (28). Unfortunately biopsy material from tumors generally consists of a mixture of different cell types (29). This problem may be alleviated using laser capture microdissection to extract cells of the desired type (3). However, only a few thousand cells may be isolated at a time using this method. Furthermore laser capture microdissection relies on visual (histopathological) identification of cells and thus may give ambiguous results. An alternative is to use cell lines (27). However, care has to be taken to ascertain that *in vitro* cell culture corresponds to an *in vivo* situation.

The cellular proteome is constantly fluctuating depending on not only the cellular microenvironment but also cell cycle. Most current cancer therapeutics are directed at protein targets. Hence it is not surprising that efficiency of several such agents (*e.g.* bleomycin, cisplatin, etoposide, and vincristine) depends on the phase of the cycle (30). It is conceivable that the response of a malignant cell to chemotherapy depends on its proteomic characteristics at different stages of the cycle.

EXAMPLE OF CYTOMIC APPROACH TO PROTEOMICS: THE STUDY OF HEPG2 CELLS

Hepatocellular carcinomas are one of the most common cancers worldwide and a leading cause of death in Africa and Asia (31). These tumors are caused mainly by two viruses: hepatitis B virus and hepatitis C virus. Other hepatocarcinogens include aflatoxins (32). Several studies of protein expression in these cancers have been undertaken (33, 34). This work led to the detection of several transformation- and proliferation-associated protein variants, but it was not possible to identify and hence assign functions to these proteins (34) because of the inability to isolate sufficient proteins for microsequencing. This problem was overcome using preparative gel electrophoresis based on IEF-IPG combined with SDS-PAGE (32, 35). It has been established that marked differences in protein expression exist between normal liver and each of the transformed cell lines, HepG2, FOCUS, Huh-7, and SK-Hep1 (35, 36). For instance, expression of caveolin was up-regulated in HepG2 cells, whereas cytochrome p450 reductase was down-regulated. Application of high resolution gel electrophoresis in combination with mass spectrometry led to generation of comprehensive protein maps of several forms of hepatocellular carcinoma (32). Nonetheless these studies did not take into account different phases of the cell cycle. To illustrate the importance of this factor for proteome

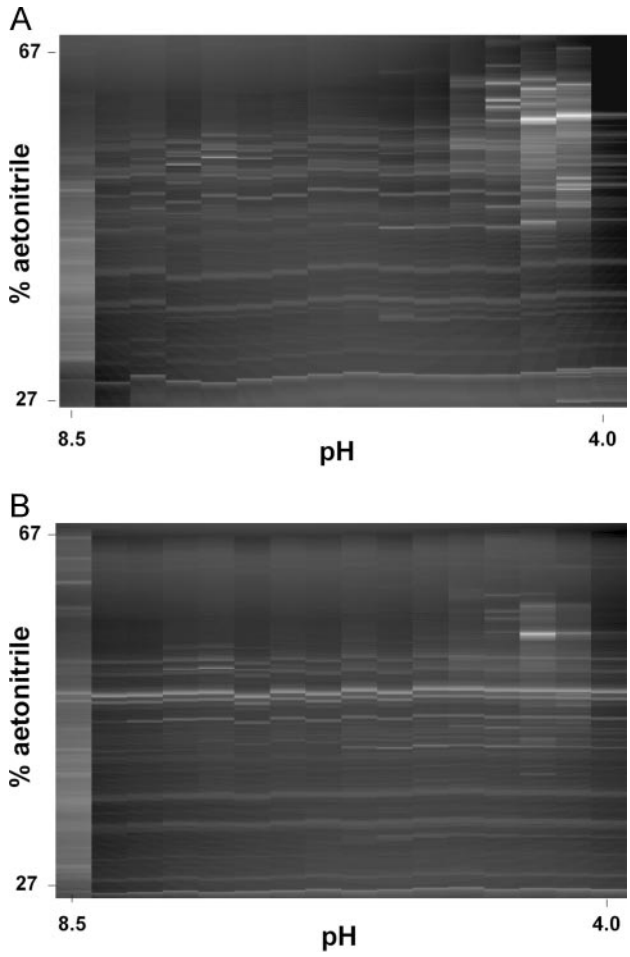

FIG. 7. **Protein maps of HepG2 cells in G_0/G_1 (interphase, *A*) G_2/M (mitosis, *B*).** Absorbance is shown in grayscale with *black* representing 0 and *white* representing 0.2 units. Total (integral) absorbance (214 nm) of these two profiles was normalized to 13.5×10^3 units.

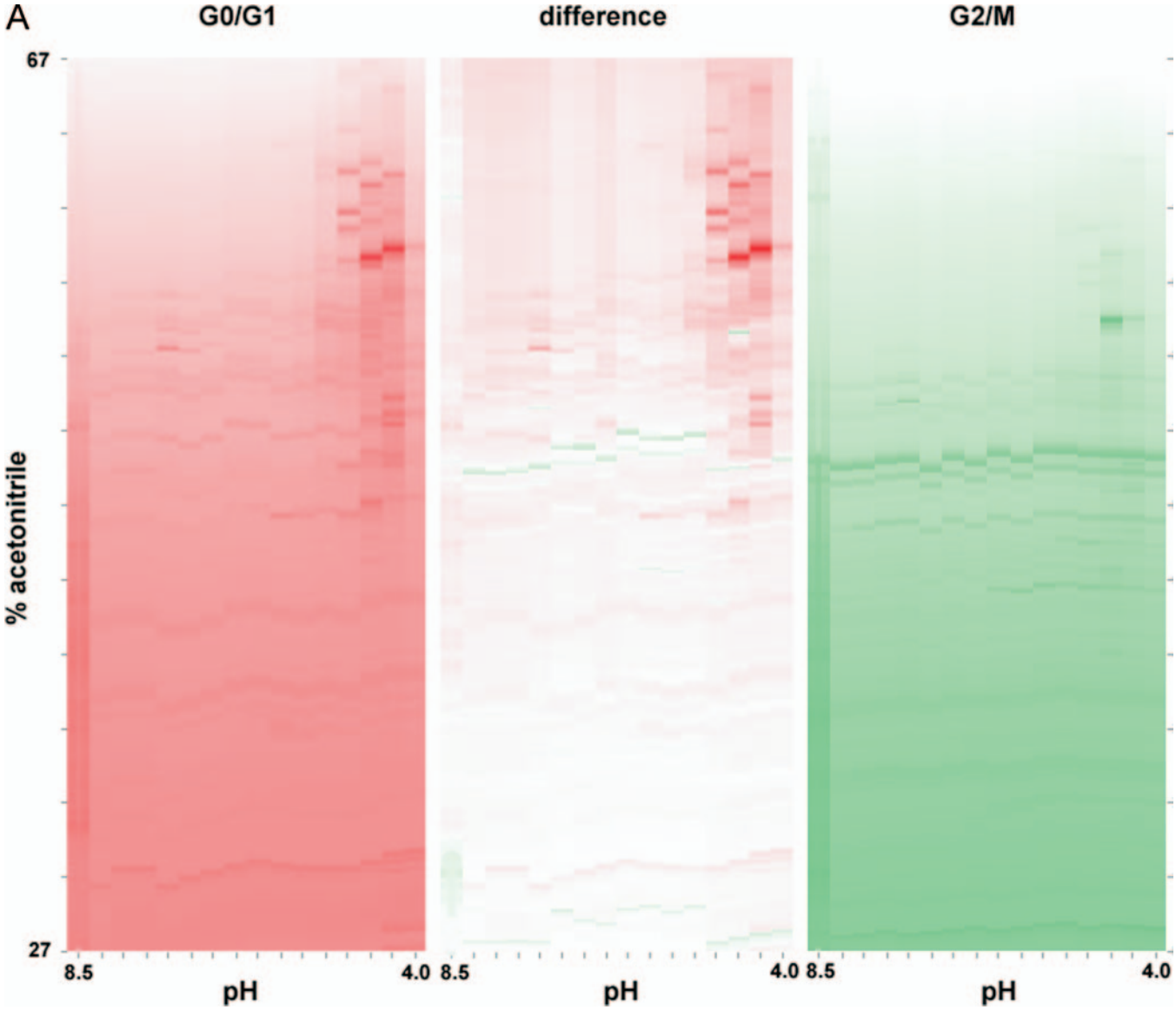

FIG. 8. **Comparison of proteome profiles for HepG2 cells in G_0/G_1 (*red*, *left*) and G_2/M (*green*, *right*).** The differential map (*middle*) shows the proteins present in different amounts in these two populations (indicated with their respective colors). Complete maps are shown in *A*, whereas *B* (next page) shows the region where the highest differences are detected.

studies we compared protein maps of HepG2 cells in G_1/G_0 and G_2/M cell cycle phases. The cells were cultured as described previously (37) and harvested in logarithmic growth phase. The cells were suspended in PBS (pH 7.4), stained using ethidium bromide (10 μg/ml final concentration), and sorted using the Epics Altra (Beckman Coulter) according to their DNA content. The two populations of HepG2 cells were lysed (Beckman Coulter protocol) to solubilize the total protein content. A typical sorting yielded 5×10^6 cells in G_0/G_1 phase and 1×10^6 cells in G_2/M phase, respectively. These two populations of HepG2 cells were lysed in a hypotonic buffer containing chaotropes (urea and thiourea) and a nonionic detergent (*n*-octyl glucopyranoside) to solubilize the total protein content (as recommended by the manufacturer of the protein fractionation system). The solubilized proteins were fractionated using HPLC implemented in a commercial protein fractionation system (PF2D, Beckman Coulter). The fractionation was executed in two steps. First, the proteins were separated according to their pI using chromatofocusing (linear pH gradient from 8.5 to 4.0, room temperature). Second, the protein fractions collected after the chromatofocusing were separated using reverse phase (linear hydrophobicity gradient from 0 to 100% of acetonitrile in water, 55 °C). The amount of protein in the fractions collected after the second step was determined by absorbance measurement at 214 nm. The results for G_0/G_1 (Fig. 7*A*) and G_2/GM (Fig. 7*B*) cell populations are shown in the form of 2D protein maps (hydrophobicity *versus* pI).

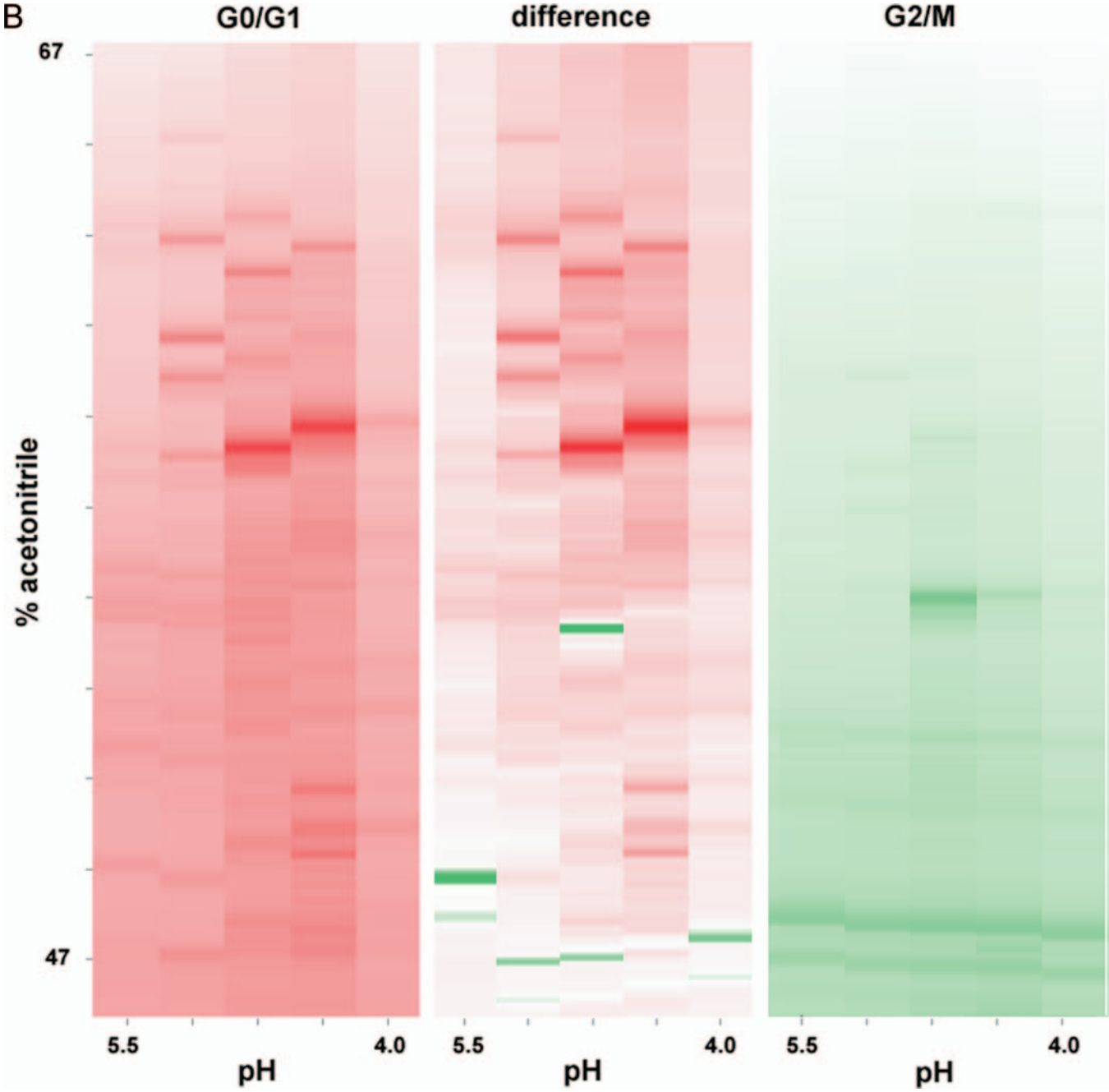

FIG. 8— *continued*

One may note that the protein maps corresponding to cells in different phases of the cycle are similar. However, the degree of similarity seems different in the regions corresponding to low pI proteins (5.5–4.0 pH units) than in the regions of high and neutral pI (8.5–5.5 pH units). To verify this notion a difference profile was constructed following normalization of the protein maps (Fig. 8, *A* and *B*).

Proteins characterized by low, neutral, and high pI represented similar fractions of the total protein content in the G_0/G_1 and G_2/M cells (Table II). The number and characteristics of absorbance peaks corresponding to proteins in the high and neutral pI region were similar in these two populations of cells. However, in the low pI region both the number and the position of the peaks differed in these populations (Fig. 8*B* and Table II). It is interesting to note that these proteins are characterized by high hydrophobicity (Fig. 8*B*). The biological significance of these data has not yet been established. Nonetheless one may postulate that taking such differences into account may improve validity of studies of cancer etiology and drug discovery. What is clear, however, is

TABLE II

Comparison of protein maps of HepG2 cells in G_0/G_1 and G_2/M phases of cell cycle

The total absorbance of each profile was normalized to unity. The peaks were considered matching if they exhibited 50% overlap. Total normalized absorbance of the matching peaks is indicated. Rel. abs., relative absorbance.

	pI					
	High (8.5–7.0)		Moderate (7.0–5.5)		Low (5.5–4.0)	
	Peaks	Rel. abs.	Peaks	Rel. abs.	Peaks	Rel. abs.
G_0/G_1	33	0.10	28	0.05	48	0.85
G_2/M	29	0.15	26	0.10	42	0.75
G_0/G_1 and G_2/M (match)	28	0.22	25	0.12	25	1.10

the importance of either phenotypic or functional separation of cells or preferentially both.

FUTURE PROSPECTS

An enormous effort has been undertaken to develop a database of lymphoid proteome profiles. The database contains 2D patterns and derived information pertaining to polypeptide constituents of unstimulated and stimulated mature T cells and immature thymocytes, cultured T cells and cell lines that have been manipulated by transfection with a variety of constructs or by treatment with specific agents, single cell-derived T and B cell clones, cells obtained from patients with lymphoproliferative disorders and leukemia, and a variety of other relevant cell populations. The database has experienced a substantial expansion in 2D patterns that it contains, currently numbering 9167 individual 2D patterns (30).

Except for already available drug targets, it is not yet known which components of proteomic profiles are biologically relevant for human disease networks in different individuals or which are excellent therapeutic targets for a given disease. Hence the first task is a diagnostic one: to obtain the proteomic profiles of normal and diseased tissues and to biologically ascertain which protein combinations are the key contributors to these two categories in the specific genetic background of an individual. As an intermediate step in utilizing the technologies identified in this communication, it will be necessary to correlate phenotypes with proteomic display in the context of clinical studies and drug discovery. This will require the identification of specific functional molecules linked to important processes such as cell cycle or apoptosis or a variety of signaling molecules for example. Identifying clinically important proteins will drive the linkage between cytomics and proteomics. Once these relevant markers are identified other technologies such as multiplexing, protein microarrays, and protein identification tools will be required to facilitate functional relationships to the phenotypic classification.

Development of single-cell profiling of phosphoprotein networks using the phosphospecific antibody- and flow cytometry-based strategy that was described in the recent study by Irish *et al.* (38) represents an important conceptual advance in our thinking for using molecular markers to create personalized molecular medicine. An ideal tumor marker should be specific for that particular type of cancer, produced only by it and not by any nonmalignant conditions. Flow cytometry combined with proteomics may help to identify and validate such markers.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: Cytometry Laboratories, Hansen Life Sciences Research Bldg., Rm. B050, Purdue University, 201 South University St., West Lafayette, IN 47907-2064. E-mail: jpr@flowcyt.cyto.purdue.edu.

REFERENCES

1. Valet, G. K., and Hoeffkes, H. G. (2004) Data pattern analysis for the individualised pretherapeutic identification of high-risk diffuse large B-cell lymphoma (DLBCL) patients by cytomics. *Cytometry A* **59,** 232–236
2. Beranova-Giorgianni, S. (2003) Proteome analysis by two dimensional gel electrophoresis and mass spectrometry: strengths and limitations. *Trends Anal. Chem.* **22,** 273–281
3. Patterson, D. S., and Aebersold, R. H. (2003) Proteomics: the first decade and beyond. *Nat. Genet.* **33,** 311–323
4. Shapiro, H. M. (1983) Multistation multiparameter flow cytometry: a critical review and rationale. *Cytometry* **3,** 227–243
5. Robinson, J. P. (2004) Multispectral cytometry: the next generation. *Biophotonics Int.* **11,** 36–40
6. Robinson, J. P., Durack, G., and Kelley, S. (1991) An innovation in flow cytometry data collection and analysis producing a correlated multiple sample analysis in a single file. *Cytometry* **12,** 82–90
7. Bassoe, C. F., Li, N., Ragheb, K., Lawler, G., Sturgis, J., and Robinson, J. P. (2003) Investigations of phagosomes, mitochondria, and acidic granules in human neutrophils using fluorescent probes. *Cytom. B Clin. Cytom.* **51,** 21–29
8. Moldavan, A. (1934) Photo-electric technique for the counting of microscopical cells. *Science* **80,** 188–189
9. Papanicolaou, G. N., and Traut, R. (1941) The diagnostic value of vaginal smears in carcinoma of the uterus. *Am. J. Obstet. Gynecol.* **42,** 193–206
10. Coons, A. H., Creech, H. J., and Jones, R. N. (1941) Immunological properties of an antibody containing a fluorescent group. *Proc. Soc. Exp. Biol. Med.* **47,** 200–202
11. Kamentsky, L. A., Melamed, M. R., and Derman, H. (1965) Spectrophotometer: new instrument for ultrarapid cell analysis. *Science* **150,** 630–631
12. Kamentsky, L. A., and Melamed, M. R. (1967) Spectrophotometric cell sorter. *Science* **156,** 1364–1365
13. Sweet, R. G. (1965) High frequency recording with electrostatically deflected ink jets. *Rev. Sci. Instrum.* **36,** 131–136
14. Fulwyler, M. J. (1965) Electronic separation of biological cells by volume. *Science* **150,** 910–911
15. Shapiro, H. M. (1994) *Practical Flow Cytometry*, 3rd Ed., Wiley-Liss, New York
16. Shapiro, H. M. (2003) *Practical Flow Cytometry*, 4th Ed., Wiley-Liss, New York
17. Dittrich, W., and Gohde, W. (1969) Impulse fluorometry of single cells in

suspension. *Z. Naturforsch. Sect. B Chem. Sci.* **24**, 360–361
18. Bonner, W. A., Hulett, H. R., Sweet, R. G., and Herzenberg, L. A. (1972) Fluorescence activated cell sorting. *Rev. Sci. Instrum.* **43,** 404–409
19. van den Engh, G. (2000) High speed cell sorting, in *Emerging Tools for Single Cell Analysis: Advances in Optical Measurement Technologies* (Durack, G. and Robinson, J. P., eds) pp. 21–48, Wiley-Liss, New York
20. Kachel, V., Kordwig, E., and Glossner, E. (1977) Uniform lateral orientation of flat particles in flow through systems caused by flow forces. *J. Histochem. Cytochem.* **25,** 774–780
21. Pinkel, D., and Stovel, R. (1985) *Flow Chambers and Sample Handling.* Flow Cytometry: Instrumentation and Data Analysis. Dean, P. N., Ed. Academic Press, New York
22. Herbert, B. (1999) Advances in protein solubilisation for two-dimensional electrophoresis. *Electrophoresis* **20,** 660–663
23. Rabilloud, T., and Chevallet, M. (2000) *Solubilization of Proteins in 2D Electrophoresis*, Vol. 9, Springer, Berlin
24. Liebler, D. C. (2002) *Introduction to Proteomics: Tools for the New Biology*, Humana Press, Totowa, NJ
25. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* **68,** 850–858
26. Gabor Miklos, G. L., and Maleszka, R. (2001) Integrating molecular medicine with functional proteomics: realities and expectations. *Proteomics* **1,** 30–41
27. Sarto, C., Deon, C., Doro, G., Hochstrasser, D. F., Mocarelli, P., and Sanchez, J. C. (2006) Contribution of proteomics to the molecular analysis of renal cell carcinoma with an emphasis on manganese superoxide dismutase. *Proteomics* **1,** 1288–1294
28. Friedman, S. L. (2000) Molecular regulation of hepatic fibrosis, an integrated cellular response to tissue injury *J. Biol. Chem.* **275,** 2247–2250
29. Masters, J. R., and Lakhani, S. R. (2000) How diagnosis with microarrays can help cancer patients. *Nature* **404,** 921
30. Hütter, G., and Sinha, P. (2001) Proteomics for studying cancer cells and the development of chemo resistance. *Proteomics* **1,** 1233–1248
31. Schafer, D. F., and Sorrell, M. F. (1999) Hepatocellular carcinoma. *Lancet* **353,** 1253–1257
32. Seong, T. C., Liang, R. C., Leow, C. K., and Chung, M. C. (2001) Hepatocellular carcinoma: from bedside to proteomics. *Proteomics* **1,** 1249–1263
33. Kudofuku, T., and Sato, T. (1985) Detection of the changes in cellular proteins in regenerating rat liver by high-resolution two-dimensional electrophoresis. *J. Chromatogr.* **343,** 51–58
34. Zeindl-Eberhart, E., and Rabes, H. M. (1992) Variant protein patterns in hepatomas and transformed liver cell lines as determined by high resolution two-dimensional gel electrophoresis (2DE). *Carcinogenesis* **13,** 1177–1183
35. Wirth, P. J., Hoang, T. N., and Benjamin, T. (1995) Micropreparative immobilized pH gradient two-dimensional electrophoresis in combination with protein microsequencing for the analysis of human liver proteins. *Electrophoresis* **16,** 1946–1960
36. Sanchez, J. C., Appel, R. D., Golaz, O., Pasquali, C., Ravier, F., Bairoch, A., and Hochstrasser, D. F. (1995) Inside SWISS-2DPAGE database. *Electrophoresis* **16,** 1131–1151
37. Bernas, T., Zarebski, M., Cook, P. R., and Dobrucki, J. W. (2004) Minimizing photobleaching during confocal microscopy of fluorescent probes bound to chromatin: role of anoxia and photon flux. *J. Microsc.* **215,** 281–296
38. Irish, J. M., Hovland, R., Krutzik, P. O., Perez, O. D., Bruserud, O., Gjertsen, B. T., and Nolan, G. P. (2004) Single cell profiling of potentiated phosphoprotein networks in cancer cells. *Cell* **118,** 217–228
39. Reynolds, O. (1883) An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Phil. Trans. R. Soc. Lond.* **174,** 935–982