Fundamental Immunology, Second Edition, edited by William E. Paul, Raven Press Ltd., New York © 1989.

LAH #316

29

Flow Cytometry and Fluorescence-Activated Cell Sorting

David R. Parks, Leonore A. Herzenberg, and Leonard A. Herzenberg

Department of Genetics, Stanford University, Stanford, California 94305

Over the last 10 years, flow analysis and cell sorting have evolved from esoteric technologies carried out in a few specialized laboratories to become standard research tools used extensively in immunology and in an expanding range of other disciplines. One of the original lines of development in fluorescence-activated cell sorting (FACS)[1] was specifically directed toward applications in immunology (1,2), and immunologic research is still a major source of the challenges that promote new developments in the field. However, attention has also been focused recently on FACS applications in molecular biology studies, and these studies have now begun to have a substantial impact on FACS development.

In flow analysis systems, cells in suspension are carried through a sensing region, usually illuminated by at least one focused laser beam. Light scatter and fluorescence signals collected from each cell (as it passes the sensing region) are evaluated and recorded in a computer system. These signals may also be used to identify cells of interest which can then be sorted and used in further investigations.

The power of FACS analysis comes from making *quantitative multiparameter measurements on statistically large numbers of individual cells*. Cell analysis and sorting at rates up to several thousand cells per second are typical. Essentially any cellular property detectable with a fluorescent label or antibody can be measured with the FACS, and that measurement can be combined with measurements of several other properties to provide a sophisticated characterization of the cell (and, collectively, of cell populations present in a sample).

FACS yields a large number of selected cells in a relatively short time and provides a high degree of flexibility in the criteria used to specify which cells are to be sorted. Since the mechanism for cell sorting is independent of the measurement values, sorting criteria can be defined for any desired combination of signal levels in the various measurements. Thus particular cells or cell populations can be identified by multiparameter analysis, sorted according to criteria established by the analysis, and then used for functional testing or physical observation.

[1] We use the long-standing acronym FACS for fluorescence-activated cell sorter or sorting where it fits better than its synonyms. None of the material in this chapter is specific to FACS, a trademark of one manufacturer of flow cytometry instruments.

These characteristics set flow cytometry apart from other cell characterization and isolation techniques, which mainly involve manipulation of cell populations rather than individual cells, are usually based on selection of a single parameter and, at best, provide a semiquantitative evaluation of that parameter.

With typical instruments, FACS techniques work well for particles in the 1- to 30-μm size range. Most applications are limited to samples that can be maintained in single-cell suspension.

This chapter is designed primarily to assist the reader in understanding, interpreting, and critically evaluating published information derived from flow cytometry studies. However, it is also constructed to provide an overall framework for the planning and execution of flow cytometry experiments and to give the reader a basis for evaluating the potential usefulness of flow cytometry in his or her own studies. Thus it effectively complements two recent reviews (3,4), which offer a more comprehensive view of flow cytometry methodology for the committed practitioner. Some areas not well covered in available references are presented here in more detail than those for which convenient references can be given.

HISTORICAL OVERVIEW OF FLOW CYTOMETRY IN IMMUNOLOGY

The fluorescence-activated cell sorter (FACS) was developed in the early 1970s to meet a growing need for more specific cell separation methods to facilitate genetic and functional studies on viable lymphocyte subpopulations. By that time, studies using immunofluorescence microscopy had already demonstrated that lymphocytes could be broadly subdivided according to whether they had surface immunoglobulin (Ig) or surface Thy-1 molecules, that is, into what we now call B and T cells. In addition, functional studies with cells from sources such as bone marrow, spleen, and thymus had established convincing correlations between the presence of the visually defined T and B cell subpopulations and the ability of the cell population to express or transfer particular immune functions.

Physical and immunocytotoxic methods had been developed that were useful for depleting B or T cells from lymphocyte preparations, for example, passage through nylon wool, passage through affinity columns, or killing with anti-Thy-1 and complement. These methods, however, did not provide highly enriched samples of individual lymphocyte subpopulations (because target cells were specifically depleted rather than specificially enriched). The development and commercial proliferation of FACS instruments has led to their widespread use in studies investigating the functions of lymphocyte subpopulations that are identified, sorted, and characterized by their expression of cell surface markers detected with fluorescent labeled (mostly monoclonal) antibodies, individually or in combinations.

One of the best known and perhaps the most important FACS applications to date derives from the identification of the so-called helper and suppressor T cell subsets in human peripheral blood ($CD4^+$ and $CD8^+$ T cells, respectively) (5,6). The demonstration that reductions in $CD4^+$ T cell frequencies correlate with the onset of clinical AIDS and with the severity of the disease manifestations has now led to extensive use of FACS instruments to monitor the course of the disease, to determine when to initiate treatment, and to evaluate responses to established and experimental therapies.

Other FACS applications of importance in basic and clinical immunology include the charting of T cell and B cell developmental pathways, the identification of functional distinctions among the lymphocytes in these pathways, the typing of leukemias in humans and the "tailoring" of therapies for these leukemias, the recognition of lymphocyte subpopulation frequency changes in various diseases in the mouse and human, and, most recently, the development of methods for studying the regulation of gene expression in lymphocytes, that is, by measuring the expression of a bacterial "reporter" gene (*E. coli* β-galactosidase). Several of these applications are discussed later in this chapter.

FUNDAMENTALS OF FLOW CYTOMETRY INSTRUMENTATION

A variety of commercial and custom-made cell sorters and flow analyzers exist, but most of their basic functions can be described with reference to a generic diagram as shown in Fig. 1.

The fluid flow system carries cells to be analyzed through the sensing zone and provides the cell-containing drops which are manipulated to effect cell sorting. Cell-free sheath fluid is supplied to the nozzle, producing a liquid jet that intersects one or more laser beams. The cell sample is injected into the center of the sheath flow and forms the core of the liquid jet so that the cells are constrained to a region near the center of the jet. Although cells occur at random positions in the sample stream, they cross each laser beam in a few microseconds, making it unlikely that two cells will appear simultaneously and be measured together.

When two or more lasers are in use, their beam spots on the liquid jet are normally separated by a short distance so that as a cell passes through the beams, signals from different lasers are separated both in space and time. This is an important aid in making distinct measurements on different dyes. The separation in space allows separate apertures to be used to select fluorescence light from the laser–jet intersections, while the separation in time is useful in the electronic evaluation of the signals. Several different types of lasers have been used for flow cytometry, but the principal ones in routine use are argon ion lasers (with useful wavelengths of 350 to 360, 458, 488, and 514 nm) and tunable dye lasers (pumped by argon ion lasers) operating in the 580- to 620-nm range.

In most nonsorting flow cytometers and in some sorters the laser beams intersect the cell stream in an enclosed flow cell. This configuration has some advantages espe-

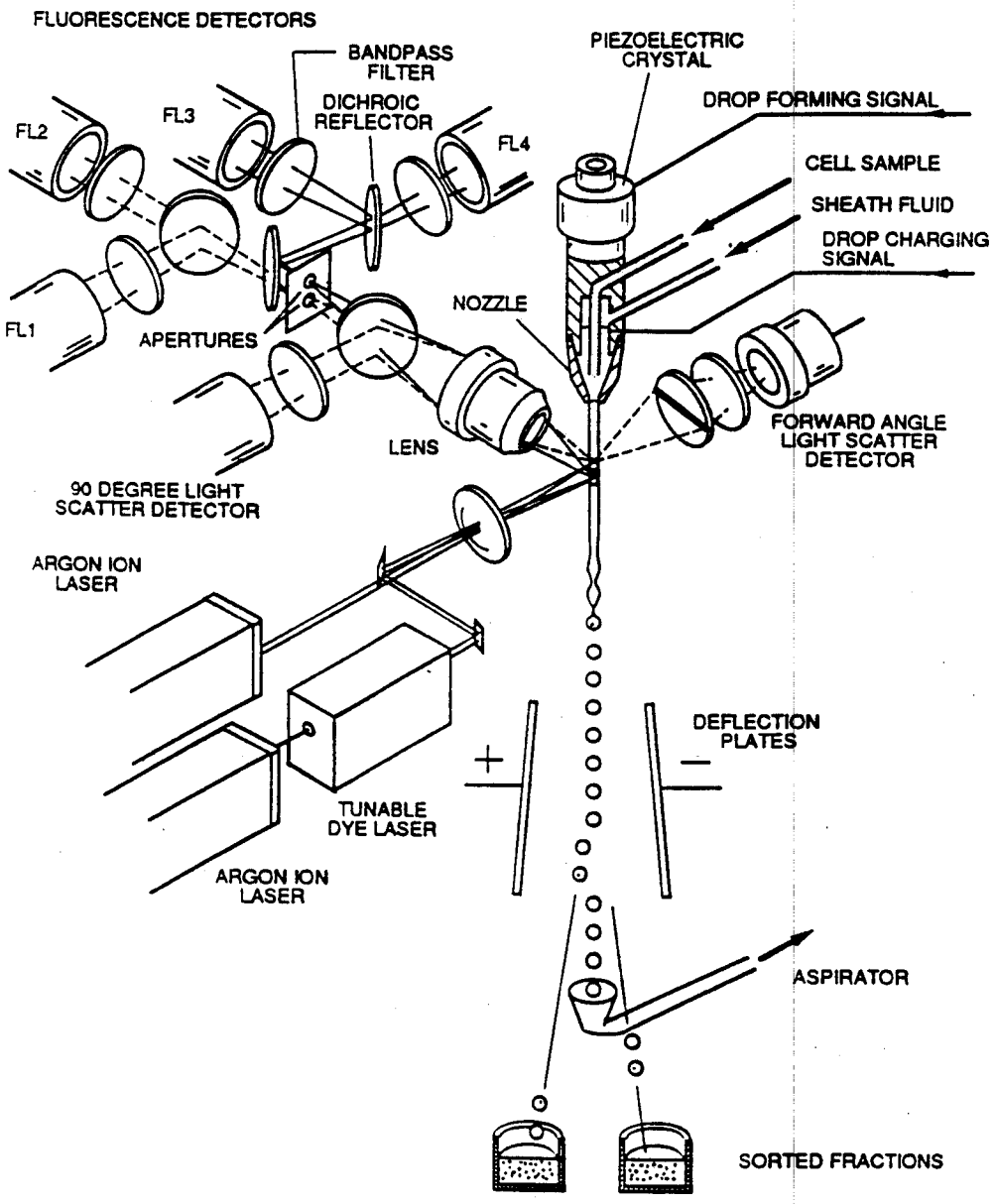

FIG. 1. Diagram of a typical two laser cell sorter (fluid, mechanical, and optical elements).

cially in allowing the use of more efficient fluorescence collection optics. However, keeping flow cell surfaces clean so that accurate measurements are made can be a problem.

As cells pass through the focused laser beam(s), light is scattered in all directions and fluorescent molecules are excited and emit light whose spectrum is determined by the nature of those molecules. Scattered light, collected over one or more angular ranges (frequently at forward angles and near 90°), is detected to produce electronic signals corresponding to the magnitude of the scattered light. Fluorescence is usually collected by a lens at 90° to the jet and laser beam directions. Dichroic reflector filters divide the collected light into different wavelength ranges and direct the light toward separate detectors for each fluorescence signal that is to be measured. Bandpass filters in front of each detector reject scattered laser light (which is usually much more intense than the fluorescence) and further select exactly what range of wavelengths will be detected.

Photomultiplier tube (PMT) detectors convert light into electronic signals whose magnitude is proportional to the intensity of the light. PMTs are normally used to detect fluorescence signals since they provide low noise detection and amplification of small amounts of light. (Silicon photodiodes can be used instead of PMTs to detect strong signals like forward-angle light scatter.) The amplification or gain of a PMT is a strong function of the applied voltage so that typically a 10% change in voltage will result in about a twofold change in gain. Thus relatively small differences in voltage imply large differences in signal levels.

Each of the light scatter and fluorescence electronic signals is amplified and evaluated. Depending on the system, integrated signals or just the peak value of the pulses

may be measured, but in most cases the results are quite similar. Digitized versions of these results are passed to a computer for storage and analysis. Sorting decisions defining the disposition of the cells are also derived from the measurement results.

To effect sorting, the nozzle assembly is vibrated at a constant frequency by a piezoelectric crystal driven by a drop-forming signal. This causes the liquid jet to be broken into uniform drops at a constant distance from the nozzle. Under these conditions, the time between detection of a cell in the first laser beam and the incorporation of the cell into a drop is well defined. If the cell has been selected for sorting, one or more drops that should contain the cell are charged by applying a potential in the range of 100 volts to the fluid inside the nozzle. The electrical conductivity of the jet is such that the potential is effectively transmitted to any drops that break free while the potential is applied. The drops fall between deflection plates held at a constant potential difference of several thousand volts. Positively charged drops are drawn in one direction, negatively charged drops are drawn in the other, and uncharged drops are undeflected.

In the following sections we describe FACS measurements, data analysis, and cell sorting in more detail particularly as they apply to important considerations in immunologic studies.

LIGHT SCATTER ANALYSIS

When a cell passes through a laser beam, some of the laser light is scattered. The intensity (and polarization) of the scattered light at different angles contains a great deal of information about the size, shape, and internal structure of the cell. Because of the complexity of light scatter distributions, however, it is not possible to convert a few measurements into a detailed "picture" of the cell. On the other hand, light scatter measurements in appropriate angular ranges are very useful for distinguishing different types of cells or cells in different states. Flow cytometers commonly provide one light scatter measurement taken over a range of moderate forward angles (e.g., 2° to 15° from the laser beam) and another in a region near 90°.

The applications of light scatter measurements include the following:

1. *Cell detection.* Light scatter (either forward or 90°) provides reliable detection of all cell-sized objects regardless of fluorescence, so it is used in most flow cytometry to initiate measurement on all signal channels.
2. *Cell size evaluation.* Light scatter in all directions is affected by cell size, but it is not easy to obtain a true cell size measure since forward and 90° scatter also include contributions due to the state and structure of the cell. Among similar cells, larger signals in either range imply larger cell size, but considerable caution should be exercised in such interpretation. Scatter at very small angles (less than 2°) is more closely related to size, but signals in this range are not detected on most instruments due to the difficulty of making measurements so close to the unscattered laser beam.
3. *Cell structure evaluation.* The internal structure of a cell has a great effect on the amount of light scattered in the 90° range. Among equal-sized cells those with more granularity scatter more light. This effect is quite pronounced in human blood where the forward light scatter distribution of granulocytes overlaps those of lymphocytes and monocytes, but the granulocytes are easily resolved due to their much larger signals in 90° scatter.
4. *Live/dead discrimination.* For cell surface staining experiments and most other flow cytometry applications it is desirable to exclude dead cells from consideration. Cells that were dead at the start of an experiment tend to stain differently than live cells, introducing confusion and errors in interpretation. Viewing lymphocytes or cells from tissue culture on an ordinary microscope, dead cells tend to look ragged edged and grainy compared to live cells. This is reflected in FACS light scatter measurements where dead cells usually give lower signals in forward light scatter and larger signals in 90° scatter than the corresponding live cells. The difference in scatter signals obtained from live versus dead cells often makes it possible to define a region in a forward versus 90° light scatter display that includes the live cells of interest and excludes most dead cells and debris (which can thus be excluded from sorted populations). Dead cell exclusion can be made even stronger using propidium iodide fluorescence as described later.

FLUORESCENCE ANALYSIS

Fluorescence is the process where a molecule absorbs light, moves into an excited state, and undergoes a transition back to the lower state with the emission of a photon of light. Some energy is lost in internal transitions within the states so that the emitted photon normally has less energy (i.e., a longer wavelength) than the original excitation. The spectral shift makes it possible to measure very low levels of fluorescence, because the emitted light is measured at wavelengths where there is little or no background light. Under optimal conditions, very small numbers of reagent molecules carrying fluorescent markers (often only a few thousand) can be detected.

There are various types of fluorescent reagents used in flow cytometry. These include antibodies and other ligands coupled to fluorescent molecules, intrinsically fluorescent DNA labeling compounds, molecules whose fluorescent properties are a function of calcium concentration, pH, and so on. Fluorescent reagents having suitable differences in their excitation and emission spectra can be used together in multiple labeling experiments. For example, routine stainings are currently carried out with a cocktail of up to four different immunofluorescent reagents, each of which is then measured independently.

As an example of dyes that can be used in combination, Fig. 2 shows excitation and emission spectra for a set of five dyes (fluorescein, r-phycoerythrin, propidium iodide, Texas Red, and allophycocyanin) that are used routinely as immunofluorescence labels in flow cytometry and that

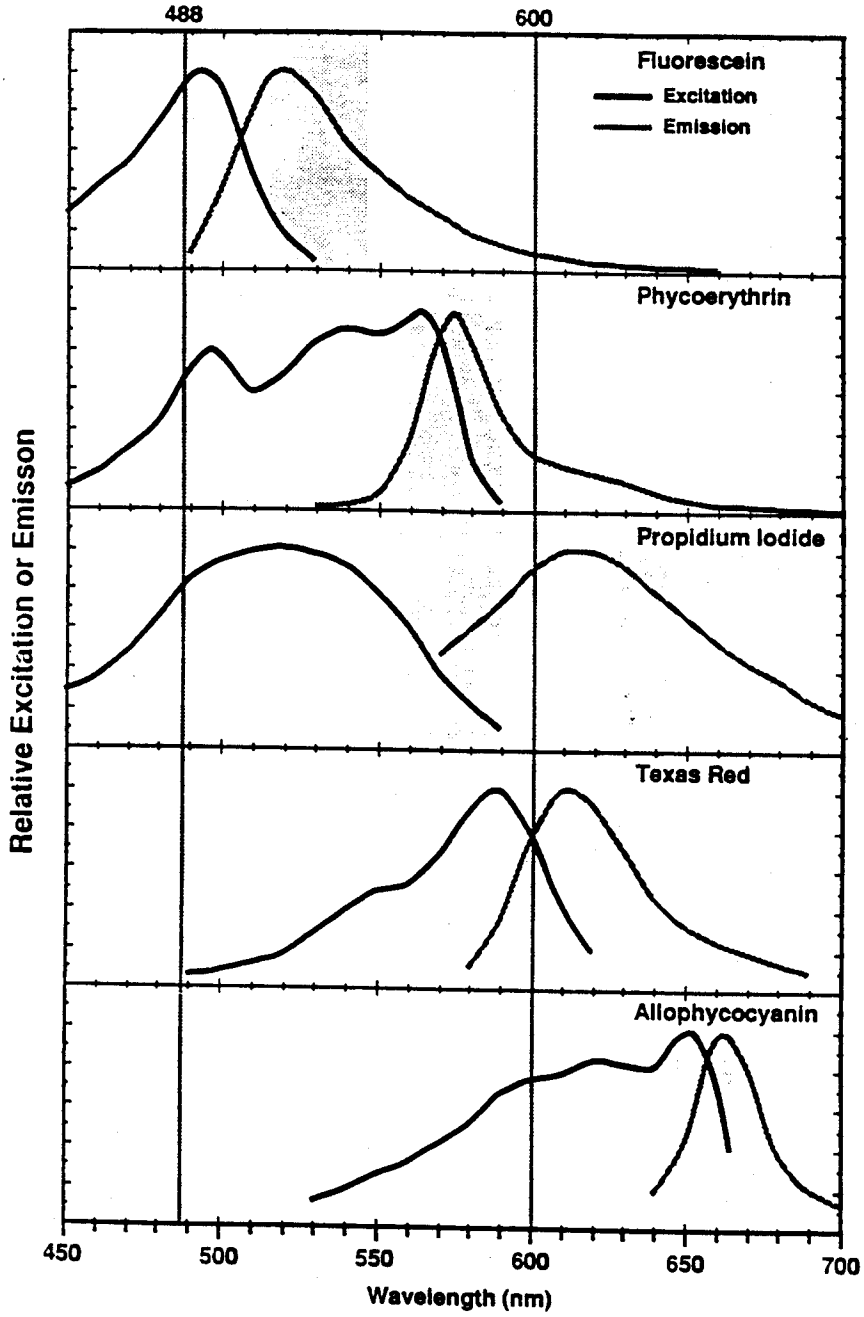

FIG. 2. Excitation and emission spectra of dyes commonly used for immunofluorescence in flow cytometry. The lines at 488 and 600 nm mark, respectively, typical argon ion laser and tunable dye laser wavelengths used for immunofluorescence excitation. The shaded bands indicate the detector acceptance ranges used to measure each dye. Propidium iodide fluorescence can be collected in either of the two marked regions.

may be used together for four-color immunofluorescence. Fluorescein and Texas Red are low molecular weight chromophores which are coupled directly to antibodies or other ligands. R-phycoerythrin and allophycocyanin are highly fluorescent auxiliary photosynthetic pigments of red algae and cyanobacteria and consist of proteins that carry a number (up to 34) of open-chain tetrapyrrole (bilin) chromophores. Fluorescein and r-phycoerythrin are excited efficiently by the 488-nm line of the argon ion laser. Texas Red and allophycocyanin can be excited together by a dye laser operating at about 600 nm.

With appropriate optical filters each of these four dyes can be detected in preference to the others on a photomultiplier detector; however, due to the overlaps in the emission spectra (see Fig. 2), each detector will usually receive some signal from each of the dyes excited by a particular laser (i.e., the fluorescence detector systems do not actually measure the amounts of different dyes but only the amounts of light in different color ranges). Nevertheless, if the emission spectra of the two dyes are sufficiently different, it is possible to subtract out the overlapping contributions electronically to obtain "independent" dye measurements. For example, fluorescein fluorescence is predominantly green but includes some yellow emission. Therefore a fluorescein-labeled cell will give not only a "green" signal but also some signal on the "yellow" detector optimized for measuring phycoerythrin. Since the relative amounts of green and yellow in the fluorescein fluorescence are constant, the green signal can be used to estimate the fluorescein-de-

rived yellow signal. Subtracting this from the total yellow signal leaves the actual phycoerythrin-derived contribution. Most flow cytometers have provision for such "two-color compensation," which gives processed signals that are effectively single-dye measurements. Correct evaluation of double-labeled cells depends on the accuracy of these adjustments.

Compensation adjustments are unnecessary when two lasers are used with matched dyes, for example, 488-nm laser excitation with fluorescein and 600-nm excitation with Texas Red or allophycocyanin. Under these conditions, two fully independent fluorescence signals are obtained. This is advantageous when it is necessary to evaluate small amounts of one reagent in the presence of large amounts of the other.

Autofluorescence

In many FACS experiments, especially those involving immunofluorescence and/or large cultured cells, cell autofluorescence is the main source of "background" signal which limits the sensitivity and accuracy of the reagent measurements. Although autofluorescence contribution can be minimized by proper choice of optical filters, in the end it is a real part of the signal and must be taken into account in evaluating data. When autofluorescence is bright and has a consistent spectrum that is different from the reagent dye spectrum, it is possible to correct for autofluorescence in the primary signal by treating it as a second dye and carrying out dual fluorescence compensation as described previously (7). In cases where it is applicable, this clearly improves the quality of the reagent dye measurement.

Typical cell autofluorescence decreases at longer wavelengths of excitation or emission. In particular, autofluorescence is less of a problem for dye laser excitation around 600 nm than for 488-nm excitation. Cell fixation tends to increase autofluorescence. A fixation procedure currently used to minimize fixation-related autofluorescence employs 1% paraformaldehyde buffered at pH 7.4 for 30 to 120 min (lower pH will decrease the fluorescence of fluorescein and some other dyes).

Propidium Iodide as a Dead Cell Marker

Propidium iodide (PI) is a small molecule that binds to nucleic acids and is very effectively excluded by mammalian cells with intact cell membranes. When exposed to low concentrations of PI, dead cells become strongly fluorescent and can be excluded from further consideration without detectably affecting live cells. The excitation and emission spectra of PI bound to DNA are shown in Fig. 2. With appropriate detector configurations, PI measurements for dead cell exclusion are compatible with any or all of the immunofluorescence dyes shown in the figure.

The dead cell selectivity of PI is lost when cells are fixed, but a photoactivated ethidium compound has been developed which binds covalently to DNA (8). After unbound material is washed away, the dead cells remain marked even after fixation.

IMMUNOFLUORESCENT REAGENTS AND CELL STAINING

The full array of dyes, reagents, and staining procedures used in flow cytometry is too varied and expanding too rapidly to be reviewed here. This section therefore provides a brief description of the monoclonal antibody reagents and immunofluorescence staining methods typically used in immunology and related fields. Detailed information on these and other reagents and techniques used for flow cytometry studies can be found in several articles in *The Handbook of Experimental Immunology* (9).

Monoclonal antibodies have been very important in expanding the use of flow cytometry, particularly multicolor analysis. Their specificity, large-scale availability, and reproducibility make them particularly suitable for quantitative measurements. Fluorescent-labeled monoclonal antibodies can be used either as one-step (direct) staining reagents or as second-step (indirect) reagents that reveal the presence of unlabeled antibodies that have previously been allowed to bind to cells. Directly coupled monoclonal staining reagents are convenient to use and provide adequate fluorescence signals for most purposes. Polyclonal antibodies and multistep staining procedures sometimes yield higher signal levels, but specificity problems are often greater than with directly labeled monoclonals so that the ratio of signal to background may be decreased rather than increased. In our experience, phycoerythrin conjugates often give the best ratio of specific signal to background.

Direct staining is advantageous when the number of reagents in use is small and sufficient amounts of each of the monoclonal antibodies are available to allow preparation of the fluorochrome-coupled reagents. Two-step staining is practical when a single fluorescent reagent is required to reveal the binding of a number of different first-step antibodies or when the first-step antibodies are only available in unpurified form, for example, as hybridoma supernatants. The generality afforded by this method makes it quite useful, particularly for one-color immunofluorescence studies. Multiparameter studies with this method are more difficult because a second-step antibody that binds to more than one of the first-step reagents will yield ambiguous results.

The two-step staining methods typically used in multiparameter studies are designed to assure the specific binding of the second-step reagent to only a single first-step reagent. The biotin–avidin binding system is the most commonly used of these methods. Biotin is a small molecule that can be conveniently coupled to (first-step) antibodies. Avidin binds strongly to biotin and can be labeled with phycobiliproteins or with small fluorochromes like Texas Red.

Low molecular weight fluorescent dyes like fluorescein

and Texas Red are usually provided in active forms that couple covalently to antibody protein upon mixing. Macromolecular dyes such as the phycobiliproteins (r-phycoerythrin, allophycocyanin) must be treated with activating agents prior to coupling to antibody molecules, which must also be preactivated so that the two proteins will link to each other when mixed. (Care must be exercised with this type of protein–protein coupling to avoid excessive cross-linking and formation of large complexes.)

Cell preparation and staining procedures vary according to the species and cell type studied. In general, media, temperature, and so on should be selected to keep the cells in the best possible condition during staining, washing, and post-stain handling of samples. In multicolor studies, all direct-staining and first-step reagents can be mixed and added to the cells together. Incubations for 20 min on ice are typical for lymphocyte staining. After incubation the cells are washed to remove unbound reagent and resuspended in media. Second-step reagents (if any) can then be added and the incubation cycle repeated. After the last wash, cells are resuspended in medium (containing PI, if appropriate).

If the cells are not to be analyzed immediately or if they present a biohazard, they may be fixed after washing. Reliable fixation procedures have been well worked out for human lymphocytes; however, some modifications may be required for use with other species or cell types. In general, optimized fixation procedures have relatively little effect on fluorescence patterns (except for PI, which can no longer be used for dead cell discrimination). Some degradation of light scatter signals may occur, particularly if the cells are held for more than 18 hr.

STANDARDS AND CALIBRATION

The job of particle standards in flow cytometry is to ensure that data taken at different times and possibly on different instruments can be compared quantitatively. Ideal standard particles would look like cells to the instrument, would have fluorescences with the same spectra as the dyes used to stain cells, would have a high degree of uniformity, and would be stable over time. Unfortunately, no single standard fulfills all these criteria, and the choice of actual standards depends on which factors are most important in a particular situation.

For day-to-day standardization of a particular instrument, we have found that small polystyrene microspheres ("beads") loaded with appropriate dyes have good uniformity and stability, which are critical factors in this application. For comparisons between different instruments, particles actually carrying the dyes of interest may be needed. (The limitations in stability and uniformity characteristic of such particles are not so important for these comparisons, at least over a limited time span.)

In a typical procedure for day-to-day standardization, the standard particle sample is run after an appropriate warm-up period and the optical alignment is adjusted to optimize signal amplitude and uniformity. A check on signal uniformity (e.g., coefficient of variation) at this point indicates whether normal data quality is being achieved. If so, photomultiplier voltages or amplifier gains are adjusted to bring the signal amplitudes to exactly their standard levels (correcting for minor differences in laser power, optical element cleanness, etc. from day to day). The instrument is then ready to run cells.

CELL SORTING

Some FACS experiments are primarily cell sorting oriented and the sorted cells are the main product. In other experiments, the data from FACS analysis are all that is desired and the cells themselves are discarded. Usually, however, even in the most analysis-oriented flow cytometry project, a point will come where it is necessary (or desirable) to isolate some subpopulation of the cells for further examination or experimentation. Cell sorting makes it possible to select practically any type of cell that can be defined by the measured parameters.

The gating processes described later for data analysis are also applicable to defining sorting criteria. Cells can be classified for sorting using any combination of the measured parameters that the instrument at hand is able to evaluate; however, the amount of computation that can be done to arrive at the classification is limited by the need to have the decision available before the cell is incorporated into a drop (several hundred microseconds after the initial measurement).

With the drop-forming oscillator on, the time required for a cell to go from the laser sensing area to the point where drops are forming is fixed by the jet velocity and the distance to the break point. Charging a drop, or to be safe more than one drop, at the time that a previously measured cell should appear at the break point results in the deflection of the drop containing the cell. The deflection plates produce a constant electric field which deflects charged drops left or right (depending on whether they are charged positive or negative) and leaves uncharged drops undeflected. In normal sorting, purity of the sorted fractions is maintained by "coincidence" logic that prevents any drop from being sorted if there is a possibility that it contains an undesired cell. This results in some loss of desired cells, but usually purity is more important.

The rate at which cells can be sorted is fundamentally limited by the drop formation rate in the jet. Obviously, two cells in the same drop cannot be sorted into different fractions. The practical limit is a function of the drop formation rate, the number of drops allocated for sorting each cell, and the acceptable loss rate. For "two-drop" sorting, a cell rate about 15% of the drop rate gives a good compromise between speed and efficiency. With a 30-kHz drop frequency this corresponds to 4,500 cells per second or 1.5×10^7 cells per hour. Taking coincidence losses and so on into account gives an effective total cell flow of about 10^7 per hour. Cell aggregation can cause a dramatic decrease in cell recovery, since only single cells can be properly sorted. The majority of cells in what looks like a slightly aggregated sample may well be in the aggregates.

Cell sorting can be routinely carried out under sterile conditions, usually for the purpose of growing the sorted cells in long-term culture. To obtain clonal populations derived from selected cells, single cells fitting the selection criteria are sorted directly into individual culture wells. This procedure is convenient when rare cells are being selected, but it is so much more efficient than limiting dilution cloning that it is also used routinely to obtain subclones where no fluorescence selection is involved.

DATA COLLECTION

A variety of computer systems that collect, store, and analyze FACS data are currently available. The data received by these systems from all flow cytometry instruments are essentially the same. In essence, the analog electronic signals from the FACS sensors, which correspond to the amount of scattered or fluorescent light detected when a single cell passes through the laser beam, are amplified (in linear or logarithmic fashion) and translated into digital brightness levels, typically at 8- to 10-bit resolution (256 and 1,024 levels, respectively). The digitized readings are then recorded by the computer.

The simplest and usually most desirable way to store FACS data is list mode, particularly when more than two measurements are made on each cell. The measurement values for each incoming cell are recorded so that the full set of measurements and even their order of arrival are preserved. Once list mode data have been stored, they can be processed in various ways to obtain viewable displays and numerical evaluations, and they can be repeatedly reprocessed to obtain optimal results. The focus during data collection is therefore limited to assuring that the acquired data are as good as possible, since it is not necessary to analyze the data while they are being collected.

The main limitation on list mode data recording is the large amount of storage space required (e.g., 4 parameter data on 10,000 cells involves storing 40,000 values). The time required to retrieve a selected data list and sort through it to obtain results for a particular analysis is also a limiting factor. However, with the decreasing cost of disk and other storage and the increasing power of affordable computers, these problems are becoming less important. The expanding applications of multiparameter analyses, on the other hand, make list mode data more desirable.

In some cases it is possible to improve the efficiency of list mode data storage by screening the data to exclude events that are known to be uninteresting. This is commonly done by "gating" on light scatter measurements so that data values are stored only for objects whose light scatter signals are in the selected range (often designed to include all single viable cells). All measurements including the light scatter may be recorded for the accepted objects. When a propidium iodide measurement is used to exclude dead cells, the PI signal is often not recorded for even the viable cells.

Histogram (rather than list mode) storage for data is used when list mode storage is not available or when an analysis is so well understood that irreversible decisions can be made safely during the data collection. For a one-parameter histogram, the number of values ("channels") in which data will be stored is related to the digitizing resolution (e.g., 256 for 8-bit data, 1,024 for 10-bit data) and is thus independent of the number of cells analyzed. The numerical value stored in each "channel" is the number of cells whose digitized brightness corresponds to that channel value.

If multiparameter measurements are being made with the FACS and a single-parameter histogram is being collected, the measurements not being stored in the histogram are ignored or used to "gate" the values that are stored so that only a subpopulation of cells are allowed to contribute to the stored data. In some systems, histograms of several parameters can be acquired simultaneously. However, in this process the interrelation among the measurements on each cell (which is retained in list mode) is lost. Histogram data storage is compact and the data are easily processed for display and numerical evaluation, but it offers no opportunity for reanalysis and asking further questions.

Data can also be stored in two-parameter histograms, which are usually recorded at low resolution (typically 6 bits in a 64 × 64 array) to keep storage requirements low. This results in coarse steps in the data, especially in logarithmic data, and even at 64 × 64 one data set takes as much storage space as sixteen 256-channel one-parameter data sets. Again, any measured parameters in addition to the two being recorded must be ignored or just used to gate the incoming data.

The usefulness of FACS data is greatly enhanced by associating additional information with the numerical data values stored in the computer. (The provisions for such entries vary among computer systems.) Identification of the cell source, the reagents and dyes used, the amplifier gain settings, and similar information make it possible to reconstruct the experiment months or years later with minimal reliance on external notebooks and memory.

DATA DISPLAY AND EVALUATION

Several forms of data display and numerical evaluation are commonly employed in analyzing FACS data. These include single-parameter displays (e.g., histograms) which plot cell frequencies as a function of the amount of signal collected by a single sensor, contour plots and similar displays that show cell frequency distributions as a function of paired measurements on two FACS sensors, and sample/subsample statistics such as the percentage of cells falling within specified ranges of sensor readings and means or medians of signal levels.

These data forms are frequently used in concert to provide different views of the data from a single sample. Current flow cytometry computer/software systems provide capabilities in each of these areas, but there are substantial differences between systems in the options provided. Furthermore, there can be marked differences between

plots produced from a single data set when different analytic and statistical options are used. Thus considerable care must be taken when reading (or publishing) FACS data to ascertain (or state) the basis for computing a given data display. Failure to do so can result in significant errors in interpretation of the data.

In the sections that follow, we outline the methods used to generate the common FACS data displays and discuss some problems in data interpretation. Finally, we illustrate several of these points with data drawn from our recent B cell lineage and subpopulation studies (10).

Single-Parameter Displays (Histograms)

The simplest way to display one-parameter digital FACS data is to plot the number of cells with a particular signal level as a function of the signal level. For example, with 8-bit data, the x axis is divided into 256 brightness ranges (channels) and the number or frequency of cells falling in each channel is plotted in the y direction. The resulting bar graph is usually drawn with only the tops of the bars connected to form a broken line (see Fig. 3). The area under any part of this line is proportional to the total frequency of cells in that signal level region, making it possible to visually estimate the relative frequencies of two populations.

The ability of a raw data histogram to represent the true distribution of light scatter or fluorescence is limited by the discontinuous steps introduced in digitization and by statistical fluctuations due to the finite number of cells on which data are actually recorded. The effects of statistical variation on single-parameter data displays can greatly be reduced by using an appropriate density estimation func-

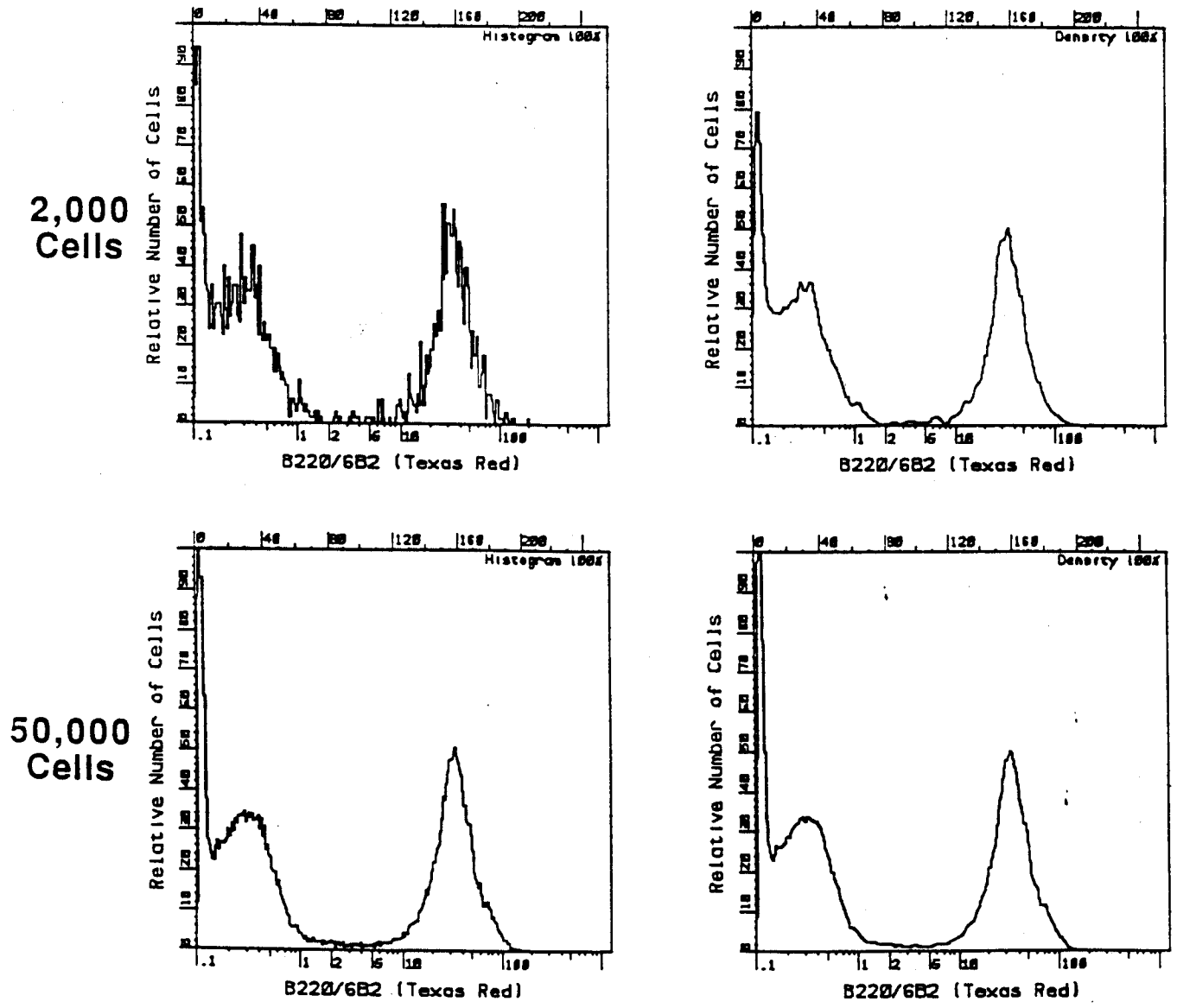

FIG. 3. Examples of raw data histograms based on 2,000 and 50,000 cells (**left panels**) and corresponding plots where a density estimation function has been applied to the same data (**right panels**).

tion to compute (histogram-like) curves that provide a better estimate of the true distribution than the raw data (11). For most purposes, such density estimation plots are preferable to raw data histograms. The results are illustrated in Fig. 3. The derived plot for 2,000 cells is essentially the same as a raw data plot with 50,000 cells and gives a much better estimate of the true distribution than the 2,000-cell raw data.

Two-Parameter Displays: Dot Displays, Perspective Plots, and Contour Plots

Typical two-parameter FACS data contain much more information than can be shown on two single-parameter histograms. A full two-parameter display has a place for cells with any combination of the two measurement values and a way to indicate the frequency of cells in any region.

Dot displays are the simplest two-parameter displays to construct: a two-dimensional region is defined in which the x dimension represents the signal for one parameter and the y dimension represents the signal for the other; and a dot is placed at the x-y point that corresponds to the measurements for each cell in the data set (see Fig. 4). The cell frequency in any region is represented simply by the number of dots in the region.

Dot displays are available as immediate readout monitors on some flow cytometers and can be generated from list mode data with some software systems. Although they do not lend themselves to good quantitation, especially in high-density regions where the dots merge, they can provide a quick view of two-parameter data and give an easy way to identify rare events in low-frequency regions of the display.

Perspective plots and contour plots are derived in effect by constructing a three-dimensional surface to represent the data and then reducing the three-dimensional surface to a two-dimensional form that can be displayed on a screen or printed. The values of the two parameters being displayed are shown as a two-dimensional region as in a dot display. The height of the surface in the third dimension represents the frequency of cells for each combination of the two-parameter values. The result is something like a land surface where hills represent regions of high cell density, the valleys represent low cell density, and the number of cells in an x-y region corresponds to the volume under the surface in that region.

Perspective plots are derived by covering the surface with a grid ("fishnet") and viewing the result from some particular direction (see Fig. 4). When rotated on a computer screen, the three-dimensional image can be visualized, and a well-chosen single view can usually convey the overall shape of the distribution. The ability to directly visualize and compare the sizes of "hills" in perspective plots makes them attractive for certain purposes; however, they tend to be less suitable for defining the boundaries of particular cell populations or the x-y locations of particular features.

Contour plots are the two-parameter displays most generally used in data analysis and publications. In these plots, the three-dimensional surface described previously is marked with "contour" lines that indicate where the surface crosses specified cell frequency levels. The contour plot is a view looking straight down onto the surface so that the height is represented only by the contour lines projected onto the two-parameter base (see Fig. 5). In its simplest form, where the contours are drawn at fixed intervals, this display is much like a geologic survey topographic map where the lines represent land surface elevation. There are, however, several different ways to choose the contour levels, and these can produce quite different displays of the same data.

Four types of contour plots, each based on a different procedure for distributing contour lines, are now in common use. We refer to these as equal interval, probability,

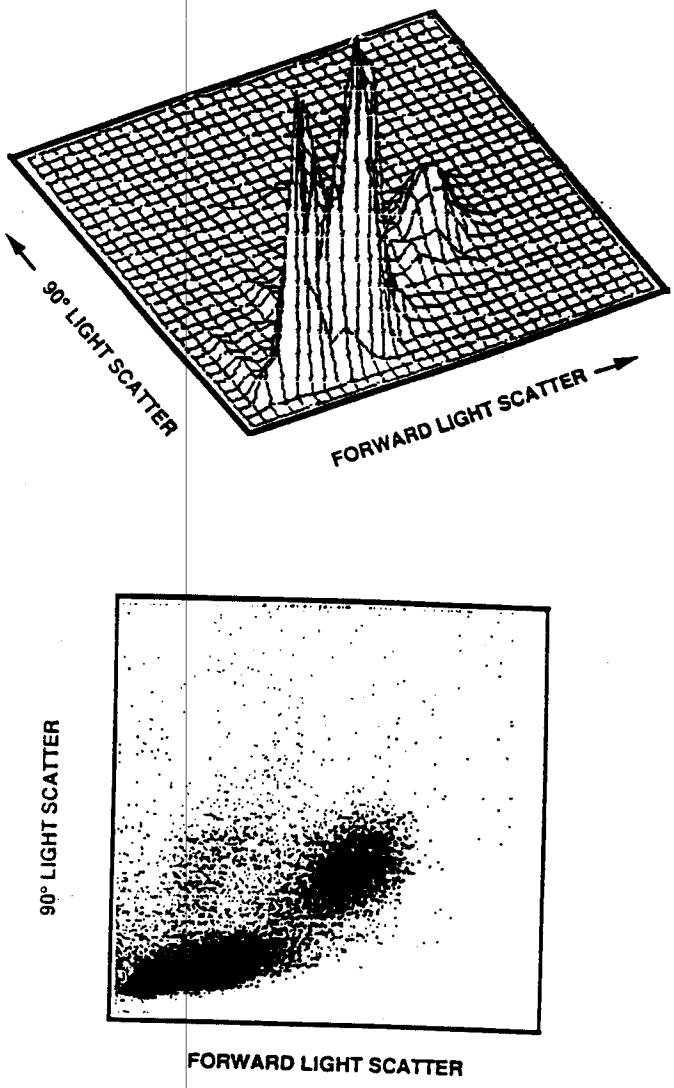

FIG. 4. Examples of two-parameter displays: a dot plot (**lower panel**) and a perspective plot (**upper panel**) of the same data. The measurements were forward light scatter and 90° light scatter on a "lymphocyte" preparation from human blood. From low light scatter to high light scatter the three major groups consist of dead cells, live lymphocytes, and live monocytes.

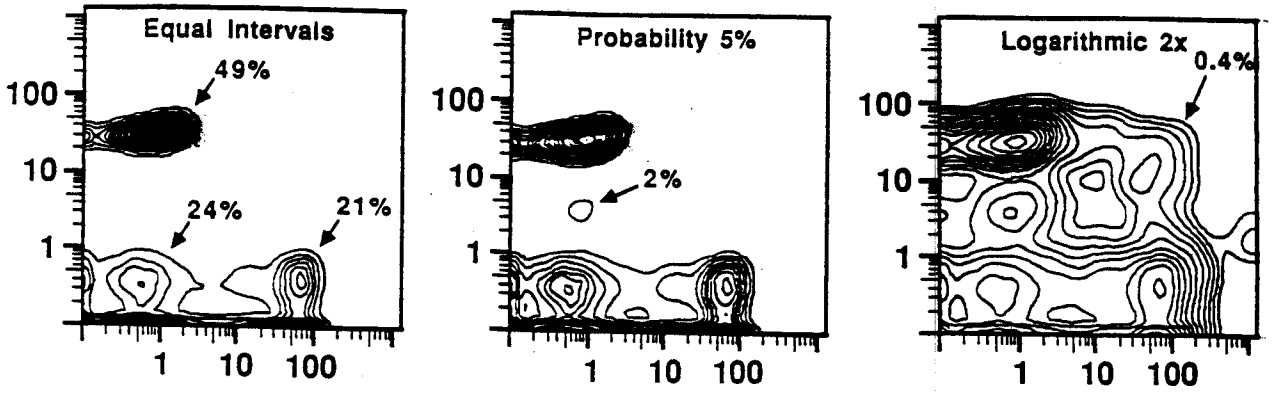

FIG. 5. Equal interval, probability, and logarithmic contour displays of a single set of two-color immunofluorescence data. The cell sample was a human peripheral blood lymphocyte preparation stained for CD8 (Leu-2a-Texas Red, *horizontal axis*) and CD4 (Leu-3-phycoerythrin, *vertical axis*). The three main subpopulations include 49% ($CD4^+$, $CD8^-$), 24% ($CD4^-$, $CD8^-$), and 21% ($CD4^-$, $CD8^+$) of the lymphocytes. The CD4-low$^+$, $CD8^-$ subpopulation seen in the logarithmic plot and marked by a single contour in the probability plot includes 2% of the lymphocytes, and the $CD4^+$, $CD8^+$ subpopulation seen only in the logarithmic plot includes 0.4%.

logarithmic, and free-style plots. These are described next and the first three are illustrated in Fig. 5.

In equal interval plots, contours are drawn at equal spacing in the cell frequency dimension. This method for defining contour levels is simple but usually not optimal for presenting flow cytometry data. When close contour spacing is used to reveal low-frequency populations, the contours in high-frequency areas may be too closely spaced to be drawn separately. Also, equal interval plots can be misleading since the number of contours in a region reflects the height of a peak rather than the number of cells it encompasses (i.e., its volume).

Probability plots, in contrast, provide a visual image that tends to reflect the number of cells under each peak. These contours are distributed so that an equal number of cell data points will fall between each pair of adjacent contours. Therefore the number of contours drawn in a given area is related to the number of cells in that area, and a randomly chosen cell in the distribution will have an equal probability of falling between any pair of adjacent contours. (The plot takes its name from this characteristic.) The approximate proportionality between cell frequency and "ink on the paper" makes probability plots preferable to equal interval plots when visual comparisons of the cell population frequencies are important.

Logarithmic contour plots allow visualization of very small populations of cells without "blacking out" high-frequency regions with large numbers of contours. Contours on logarithmic maps are drawn at intervals whose spacing decreases exponentially from high-frequency levels downward. Thus, if we use the common twofold steps for log contours and find the first contour located halfway down a peak, the next contour will be three-quarters of the way down, the next will be seven-eighths of the way down, and so on. Several logarithmic plot contours may surround a subpopulation that would not be visible on a typical equal interval or probability plot of the same data (see Fig. 5).

The final contour plot on our list we have dubbed "free-style." With some flow cytometry software the user can specify contour levels individually so that there is no necessary pattern in the selection. In general, these plots are reliable for locating populations but may give quite misleading impressions of the relative frequencies of cell populations. Thus published free-style plots should always be accompanied by a clear statement indicating the contour intervals that were selected, and readers should be wary of conclusions that are based on visual estimates of frequencies from these plots.

In summary, probability and logarithmic plots are generally the most useful types of two-parameter data display probability plots because they offer a relatively well-balanced view of the data and logarithmic plots and because they provide greater detail in low-frequency regions. Equal interval plots tend to miss low-frequency populations or black out high-frequency areas when optimized for low frequencies. Free-style contour plots can be dangerous since the selection of the contour intervals is arbitrary.

Like histograms, contour plots should be constructed on data that have been processed using an appropriate density estimation function to reduce the influence of statistical variations on the contours.

GATED DATA AND SEQUENTIAL MULTIPARAMETER ANALYSIS

Gating is used to investigate the characteristics of subpopulations of cells present in a sample. Gates, defined as upper and lower boundary values for a given parameter or as a two-parameter region, are used to define subpopulations for further analysis. Histograms, contour plots, and numerical evaluations can all be obtained for gated populations and, in fact, are rarely computed on ungated data.

In immunologic studies, initial light scatter gates are frequently set to include most lymphocytes and to exclude most monocytes, macrophages, and other nonlymphoid cells. This is possible because experience has shown that lymphocytes tend to be the cell populations in such mixtures that scatter the least amount of laser light at both forward and large angles. Therefore by drawing boundaries around the cell populations that are low in light scatter, the characteristics of the lymphocyte population can be examined with relatively little interference from other kinds of cells in the sample. Similarly, if cells are stained for surface Ig, an additional gate can be set on the appropriate signal channel to exclude (or include) Ig-bearing cells and thus to restrict analysis to T (or B) lymphocytes in the sample.

Interpreting gated data is straightforward unless cells in the gated populations occur at very low frequency. The minimum number of cells required to obtain reliable contour plots and histograms depends on the type of display and the processing methods used, if any. (As indicated previously, appropriate density estimation functions can minimize the number of cells needed.)

LOGARITHMIC VERSUS LINEAR SIGNAL PROCESSING AND DATA PRESENTATION

Most flow cytometers provide both linear and logarithmic (log) signal amplification, one of which is usually more appropriate than the other for a particular application. The choice of which to use has important implications for the display and analysis of FACS data. Theoretically, logarithmic and linear data presentations contain the same information about the cells. For example, a gating region defined on linear data can be matched by a region on log data (see Fig. 6). However, in practice, log and linear data are not readily interconverted.

Linear amplification is appropriate for signals that have a limited dynamic range and for signal distributions that include narrow peak components such as cellular DNA measurements. Logarithmic processed data can cover a wide dynamic range at constant relative resolution. For example, if four decades (10,000-fold range) of logarithmic data are recorded in 256 channels, each channel represents about 4% more signal than the previous one. Linear data in 256 channels have 4% or better resolution over only the one-decade (10-fold) range above channel 25, but the data have much better than 4% resolution in the upper-channel range. Thus when collecting log data it is seldom necessary to make gain adjustments to accommodate different samples, and cell populations at very different signal levels can be evaluated in a single data collection. Another useful property of logarithmic data is that the shape of a cell-staining distribution stays the same when equivalent reagents of different brightnesses are used.

Certain biologic measurements are particularly well suited to logarithmic scaling. For example, the distribution of a cell surface antigen among cells of a particular population is often approximately log–normal. A linear histogram of such a distribution shows a skew or "tail" on the bright side, but a logarithmic display of the same distribution will be a symmetric peak (Fig. 6). In circumstances where a cell giving half the typical signal for the population is considered to be as "untypical" as one giving twice the typical signal, a logarithmic display in which these two cells are equally spaced from the center of the distribution is the proper one. Frequently it is easier to identify and delineate subpopulations of cells in log–log two-parameter displays than in linear–linear displays (see Fig. 6).

The implications of logarithmic and linear data in statistical evaluations are discussed in the next section.

SAMPLE AND SUBSAMPLE STATISTICS

In flow cytometry, the usual objective in calculating sample statistics is to condense selected characteristics of a cell population (such as its relative frequency, typical signal level, and the shape of the distribution) into a few numbers to facilitate tabulation and comparison of results.

The main uncertainties in the evaluation of subpopulation statistics come from inherent statistical limitations and from problems in delineating the subpopulations. A minimum requirement for precise evaluations is that enough cells be represented in the data. Table 1 shows the minimum difference between two subpopulation frequency estimates that would be considered statistically significant. The value is a function of both the measured frequency range and the number of cells analyzed. With this kind of guidance we can normally assure that enough cells are measured to provide reliable results.

Determining which signals should be included and which excluded from a population is a fundamental problem for which there is often no absolute answer. In flow cytometry, even when all the cells in a sample are considered to be part of one population, there are usually some signals due to dead cells, debris, and so on which should be excluded from evaluations. When a sample includes populations that are not completely resolved in the measurements, even an optimal delineation (i.e., "gating" as described previously) will assign some cells to the wrong population and produce biased results.

In practice, the use of log or linear presentation may influence the selection of gating levels to divide populations that are not fully resolved, since the low point of the valley between populations will correspond to different signal levels in log and linear displays. In fact, some distributions that show a valley between populations in log presentation have no valley in linear (see Fig. 6).

Subpopulation Frequencies

Evaluating the fraction of cells in a subpopulation is easy if the subpopulation is well defined by the data, but

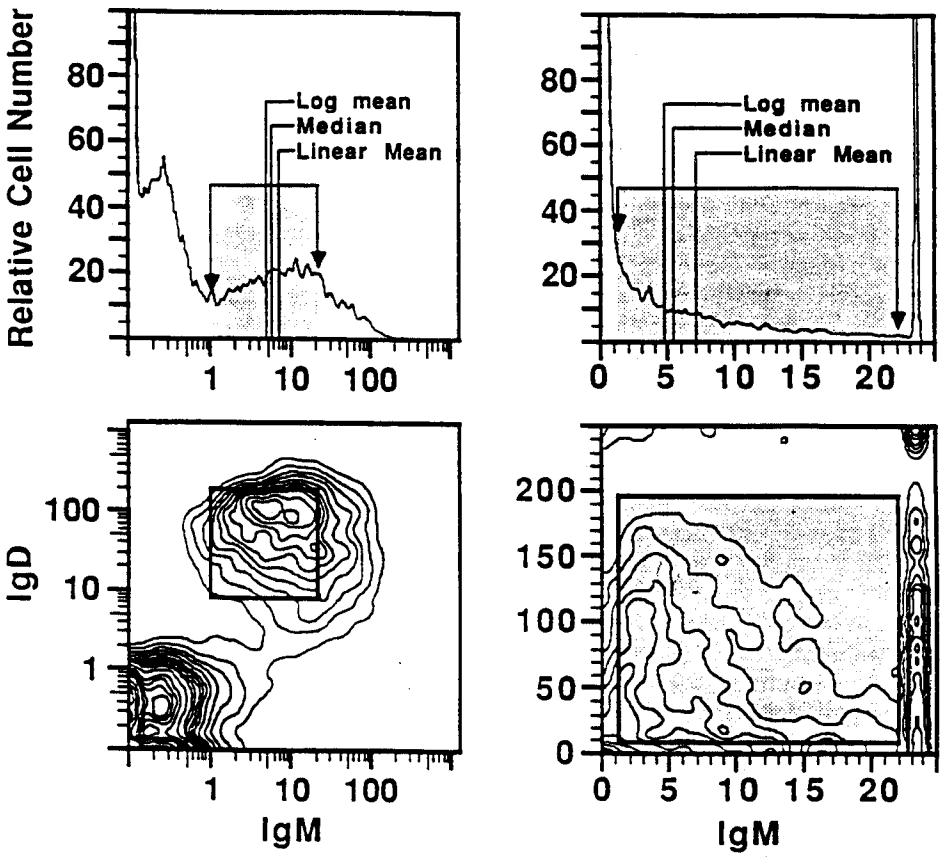

FIG. 6. Comparison of logarithmic (**left panels**) and linear (**right panels**) amplification and display of two-color immunofluorescence measurements taken on one cell sample. The *shaded* regions within the *arrows* in the matched logarithmic and linear displays mark the same actual signal ranges. The means and median were calculated for cells in the included region. The peak near the right edge of the linear displays represents all signals above 23 units. Peaks at the low edges of the logarithmic displays represent all signals below 0.1 units.

TABLE 1. *Effect of cell sample size on the significance of population frequency measurements*

Measured frequency (f)	Total number of cells analyzed (N)						
	1,000	2,000	5,000	10,000	20,000	50,000	100,000
50%	4.40%	3.11%	1.97%	1.39%	0.98%	0.62%	0.44%
30%	4.04%	2.85%	1.80%	1.27%	0.90%	0.57%	0.40%
20%	3.52%	2.49%	1.57%	1.11%	0.79%	0.49%	0.35%
10%	2.64%	1.86%	1.18%	0.83%	0.59%	0.37%	0.26%
5%	1.92%	1.36%	0.86%	0.61%	0.43%	0.29%	0.19%
1%	0.87%	0.62%	0.39%	0.27%	0.19%	0.12%	0.08%

This table presents the differences in population frequency measurements required for significance at the 95% level for various measured frequencies and cell sample sizes.

Table Calculation Model: Consider a cell population in which a particular subpopulation has a frequency f. If two data sets of N cells are taken from the population, and the observed frequency of the subpopulation in each data set is evaluated, we would expect that 95% of the time the difference between the two observed frequencies would be less than Q, the value listed in the table for cell number N and frequency f.

This implies that if we take data on two populations (perhaps of cells subjected to different treatments) and evaluate a subpopulation frequency in each data set, the frequencies must differ by at least Q to provide any evidence that the samples are indeed different. For example, to believe that frequencies of 19 and 20% indicate a real difference, they would have to come from data on at least 20,000 cells (the 1% difference is greater than the 0.79% significance level at 20% measured frequency).

difficulties occur when positive staining and control (negative) populations are not well resolved. If such an unresolved signal distribution is composed of two subpopulations, that is, some of the cells are positive for the labeled characteristic and some are negative, any cutoff selected to define positive cells will underestimate their true frequency (after subtraction of the control signals above the cutoff), and there may be no good way to estimate the error. In cases where one believes that all the cells are part of a single population but there is overlap between stained and control signal distributions (e.g., when a clonal population is stained for a low-density cell surface molecule) *stating a percentage of positive cells is inappropriate*, and an estimate of the typical signal levels (see next section) for stained and control populations is more reasonable.

Some flow cytometry software provides "quadrant statistics" in which a pair of lines, one vertical and one horizontal, are used to divide a two-parameter display into four regions (quadrants) and the fraction of cells in each region is calculated. This is quick and convenient for preliminary analysis but generally does not give optimal delineation for accurate or publishable population estimates, since quadrants often include peripheral events that are not part of the primary population.

Signal Level Evaluations

Accurate signal level evaluations on linear data depend on the linearity of the amplifier and the accuracy of gain ratios. These should generally be good, but either can be defective. Estimates of mean signal levels may also be in error because the limited dynamic range of the linear data collection often results in some cells having offscale signal values which cannot be evaluated properly.

Problems also exist for evaluating signal levels collected with logarithmic amplifiers. Such amplifiers are never perfectly logarithmic, so the accuracy of signal level estimates and of signal ratios from one point to another depends on how well the actual characteristics of the amplifier are known. For most purposes, estimates assuming true logarithmic behavior with a well-measured value for average channels per decade are adequate, but some logamps deviate significantly from true logarithmic character (David R. Parks et al., unpublished results).

The mean and median are commonly used statistics to represent the typical signal level in a population. The arithmetic mean is appropriate when comparing FACS signals from a population of cells to a corresponding bulk measurement from the population, for example, comparison of the mean amount of a surface protein measured by FACS and the mean amount found in an extract of the population. A simple mean calculated on logarithmic data is effectively a geometric mean of the original signals. For data distributions where logarithmic presentation is appropriate, the geometric mean probably represents the typical signal level better than the arithmetic mean.

The median (50th percentile point) has several advantages over means for characterizing populations in flow cytometry data. The median falls at the same signal level in logarithmic or linear data. It is also a more robust statistic than the mean in that including a few data values far outside the main population, or excluding outlying values that are really part of the population, normally has much less effect on the median than on the mean.

Population Variation or Uniformity

There are various ways to describe the shape of a distribution, but in flow cytometry only the population uniformity or width of the distribution is commonly described numerically. If more detail is needed, a histogram or other display is used. The coefficient of variation (CV, defined as the standard deviation divided by the mean) is frequently used to evaluate the uniformity of DNA staining distributions and can be applied to immunofluorescence staining, for example. However, the standard deviation and CV are even more sensitive than the mean to the inclusion or exclusion of marginal data. Percentile-based population width statistics, such as the interquartile range or ratio (from the 25th to the 75th percentile) or the 10 to 90% range, have distinct advantages for describing flow cytometry data since they are easily computed and visualized and are not so sensitive to details of population delineation.

DATA DISPLAY AND ANALYSIS ILLUSTRATIONS: MURINE B CELL SUBPOPULATIONS

Virtually any issue of an immunology journal will contain at least one example of how FACS analysis or sorting is used to define the characteristics of a T or B lymphocyte subpopulation, to investigate the origin of such subpopulations, or to chart the influence of disease processes or drug treatments on subpopulation representation. In the following section, we present several illustrative examples of FACS analysis and data display applied to characterizing a B lymphocyte subpopulation. To minimize distractions, we have drawn all the examples from a single set of studies (our work with murine Ly-1 B cells) and will refrain from commenting on the biologic significance of the FACS findings. Several publications from our laboratory and elsewhere provide more information on the biologic properties of the Ly-1 B lineage and on the flow cytometry based studies from which these examples were drawn (10,12–14).

Figure 7 shows an analysis of peritoneal cells from a BALB/c mouse. In the upper left panel, forward and large-angle light scatter gating regions have been constructed to yield two subgroups of lymphocytes (large and small). In most analyses, a single lymphocyte gating region would include both the large and small lymphocyte populations; however, in this analysis, we take advantage of small differences in light scatter to distinguish B cell subpopulations that express substantially different levels of surface IgM and IgD (revealed by the fluorescence dis-

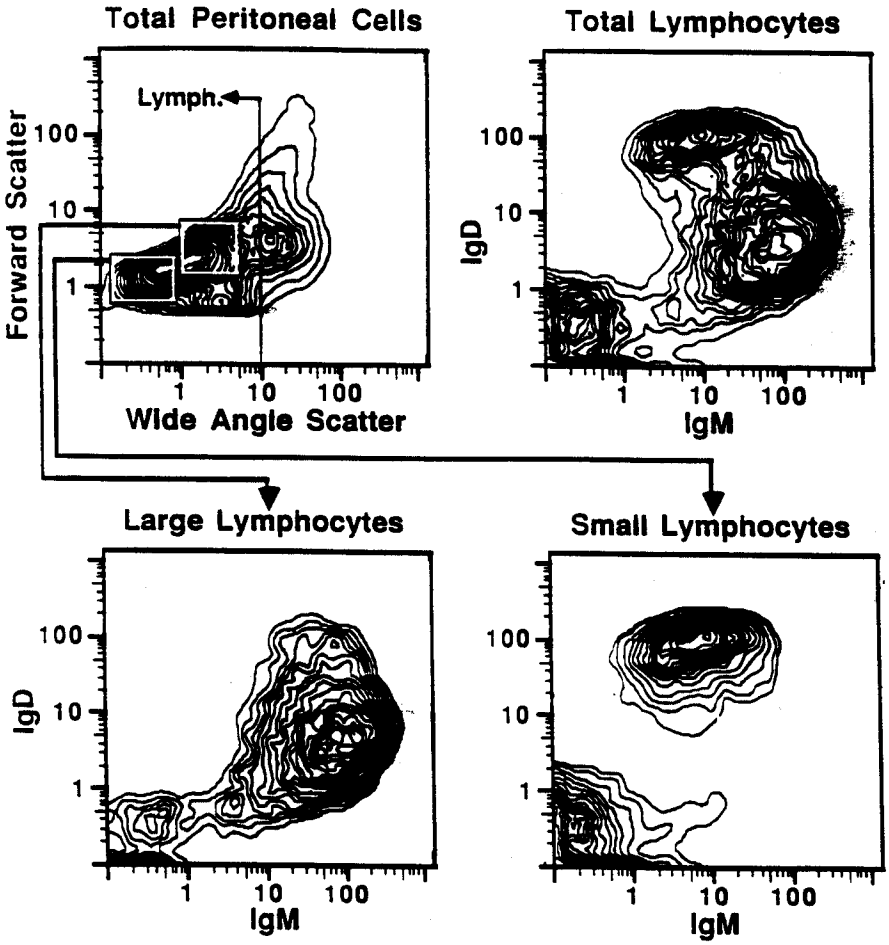

FIG. 7. Conventional and Ly-1 B cells are distinguished by forward and large-angle light scatter.

tributions for the gated populations shown in the lower panels of the figure). Ly-1 B cells, previously known to have high levels of IgM and low levels of IgD, are demonstrated by these data to be included primarily in the large lymphocyte group.

In this figure, as in most of our published work, data are presented in probability contour plots in which 5% of the cells in the sample are between each pair of contours and the data are processed with a standard density estimation function (11). Dead cells, which scatter less light than live lymphocytes and stain brightly with propidium iodide, are routinely gated out during data collection in our studies.

Note the roughly 50-fold range of IgM expression among the B cells in both the large and small lymphocyte groups and the 20-fold expression ratio between mid-peak cells of the two groups (Fig. 7, lower panels). In our experience, differences of this magnitude occur frequently in the expression of surface antigens used for subpopulation analyses. These analyses take full advantage of the four-decade range of the logarithmic amplifiers available on many (although not all) flow cytometry instruments.

Figure 8 shows a typical cell surface phenotype analysis, in this case for peritoneal Ly-1 B lineage cells (bounded by the grey box in each panel). These data are from three-color FACS analyses in which the five reagents shown were used in various combinations to allow cross-checking of the results obtained with each set. Data are plotted for IgM versus each of the other four reagents, since IgM expression tends to be very bright on Ly-1 B cells and thus offers a good visual "anchor." In essence, data in this figure show that Ly-1 B cells tend to be bright for IgM, dull for IgD, dull for Ly-1, dull for B220/6B2, and dull but clearly positive for MAC-1. Conventional B cells, in contrast, tend to be dull for IgM, bright for IgD, bright for B220/6B2, and negative for MAC-1 and Ly-1. (The third major subpopulation revealed in this figure is bright for Ly-1, negative for all the other markers tested, and is almost exclusively composed of T cells.)

Figure 9 shows an unusual, four-color analysis designed to define and measure the frequency of the very small population of Ly-1 B cells in the spleen. The markers used in this analysis (Ly-1, IgM, IgD, and B220/6B2) are all expressed at characteristic levels on Ly-1 B cells but are inadequate to define the splenic Ly-1 B population unless used in concert.

The ungated analysis in the lower left panel of Figure 9 shows the typical spleen cell staining pattern obtained with anti Ly-1 and anti-IgM. A large population of T cells expressing high levels of Ly-1 but no detectable IgM is visible along the y axis and an equally large population of (conventional) B cells expressing low levels of IgM and no detectable Ly-1 is visible along the x axis. This B cell population is essentially contiguous with a smaller pop-

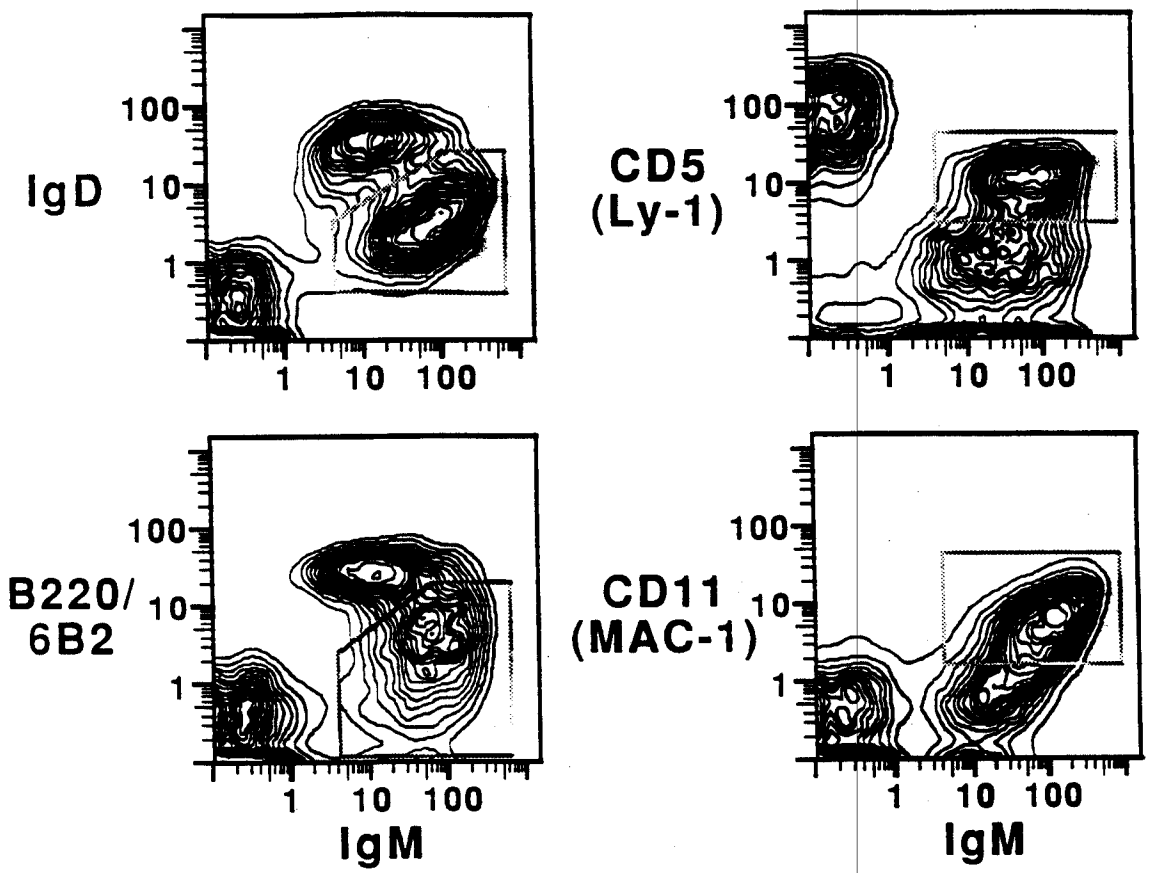

FIG. 8. Cell surface phenotype of peritoneal Ly-1 B lineage cells.

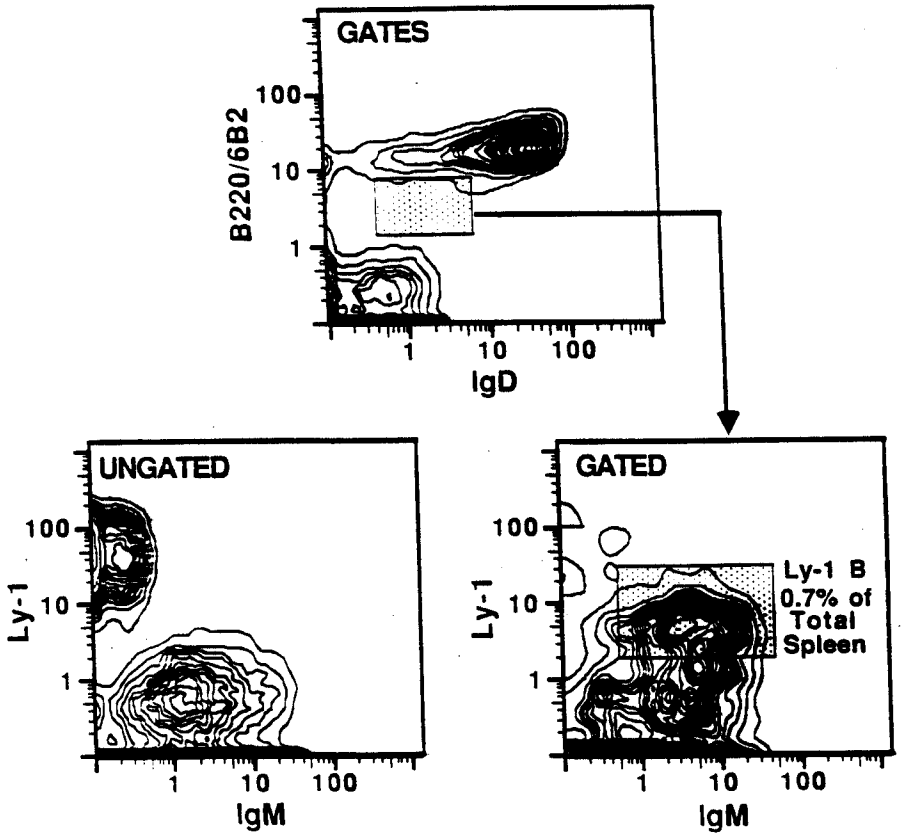

FIG. 9. Identification of splenic Ly-1 B cells by four-color FACS analysis.

ulation of "null" cells that do not express either IgM or Ly-1 but are sometimes autofluorescent enough to overlap in brightness with the dullest IgM-bearing cells. (Although not shown in the figure, these latter populations are readily distinguished by three- or four-color analysis since IgD tends to be quite bright on splenic B cells that express relatively little IgM but is not detectable on IgM-negative cells.)

To reveal the splenic Ly-1 B cells, gates were selected for IgD and B220 dimensions to include only those cells expressing these markers at levels typical for peritoneal Ly-1 B cells (as defined by previous analyses such as those shown in Fig. 8). Light scatter gates were set to include all lymphocytes and to exclude dead cells, macrophages, and other very large cells. (PI gating could not be used to exclude dead cells in this analysis because the measurement channel that is usually available for PI was used for one of the four immunofluorescence measurements.)

The Ly-1 and IgM phenotype of the spleen cell population that falls within these gates is shown in the contour map in the lower right hand panel of Fig. 9. This population, which in total represents roughly 2% of total splenic lymphocytes, is not detectable in the ungated contour map. It contains a subpopulation (bounded by the box) that represents roughly 0.7% of splenic lymphocytes and whose levels of IgM and Ly-1 expression are equivalent to the typical levels expressed by Ly-1 B cells in the peritoneum.

The logarithmic contour plots shown in Fig. 10 define a miniscule subset of splenic Ly-1 B cells (0.05% of total splenic lymphocytes) that contains the cells responsible for one of the well-known Ly-1 B functional activities, that is, the production of autoantibodies to bromelain-treated mouse erythrocytes (BrMRBC). These autoantibody-producing cells were identified some time ago in splenic Ly-1 B populations sorted with IgM and Ly-1 gates similar to those shown in the lower left panel of Fig. 9 and have now been shown to be included among the very small number of spleen cells that bind phosphatidylcholine (PtC), the target antigen on BrMRBC (15,16).

Figure 10 shows that PtC-binding cells in spleen, visualized by the use of logarithmic contour maps, express the typical Ly-1 B surface phenotype. The upper left-hand panel in the figure shows a typical probability contour plot for spleen cells stained with anti-IgM and with fluorescein-containing liposomes that have a high density of PtC groups on the surface (A. M. Stall and G. Haughton, unpublished results). The number of spleen cells that bind PtC liposomes is below detectability on this plot. However, as the upper right panel shows, a logarithmic contour map produced with the same data clearly reveals a very small population of IgM-bearing cells that stain brightly with the liposomes and express the same level of surface IgM as typical Ly-1 B cells. Similarly, the levels of several other surface markers on these PtC-binding cells are the same as those measured for the Ly-1 B population as a whole (compare Fig. 10 to the boxed regions of Fig. 8).

The final figure in this section (Fig. 11) demonstrates the use of IgM and IgD allotypes to identify the origin of cells found in chimeric mice. The chimeras were produced by lethally irradiating CBA/Ca animals (Igh-C^a allotype) and reconstituting their immune system with a mixture of CBA/Ca bone marrow and allotype congenic CBA · Igh^b (Igh-C^b allotype) peritoneal cells (PerC). The contour maps shown are produced from three-color analyses of peritoneal cells taken from chimeric animals several months after irradiation and reconstitution.

The left panels of Fig. 11 show that cells that express Igh-6^a, the IgM allotype of the peritoneal cell (PerC) donor, express the typical Ly-1 B phenotype (dull for IgD, Ly-1, and MAC-1). Furthermore, as expected, the IgD on these cells is of the PerC donor allotype (Igh-5^a). In contrast, cells expressing the bone marrow donor allotype, Igh-5^b and Igh-6^b (shown in the right panels of Fig. 11) express the typical phenotype for conventional B cells, that is, dull IgM, bright IgD, and no detectable Ly-1 or MAC-1. These data are consistent with evidence from our previous studies (17), demonstrating that conventional B cells in the chimera are derived from Ig-negative precursors in bone marrow whereas Ly-1 B cells are derived from self-renewing Ig-positive cells in the peritoneum.

The examples selected here are drawn from experiments conducted over several years during which different reagent lots were used and various modifications were made to the FACS instrument. Nevertheless, the surface staining distributions measured on the various types of lymphocytes are quite consistent. This stability is primarily due to the use of standardized procedures for reagent production, cell staining, cell analysis, and FACS instrument calibration (see previous discussion). The importance of using such standardized reagents and instruments for FACS subpopulation studies cannot be emphasized too strongly.

EXAMPLES ILLUSTRATING THE RANGE OF APPLICATIONS OF FLOW CYTOMETRY IN IMMUNOLOGIC RESEARCH

Cell surface immunofluorescence studies (such as those used to illustrate the data display and analysis section of this chapter) have been the most widespread applications of flow cytometry in immunology, but a variety of other flow cytometric techniques are also important in immunologic studies. In the following sections we describe a few approaches that we think are becoming increasingly valuable in immunology.

Indo-1/Ca^{2+} Response as a Probe for Cell Activation

A number of fluorescent probes are now available for monitoring the physiological state of cells, among which dyes that respond to pH or free calcium concentration are of particular interest. These dyes have been used to monitor cellular processes such as activation and response to hormones.

For flow cytometric evaluation, probes that respond by

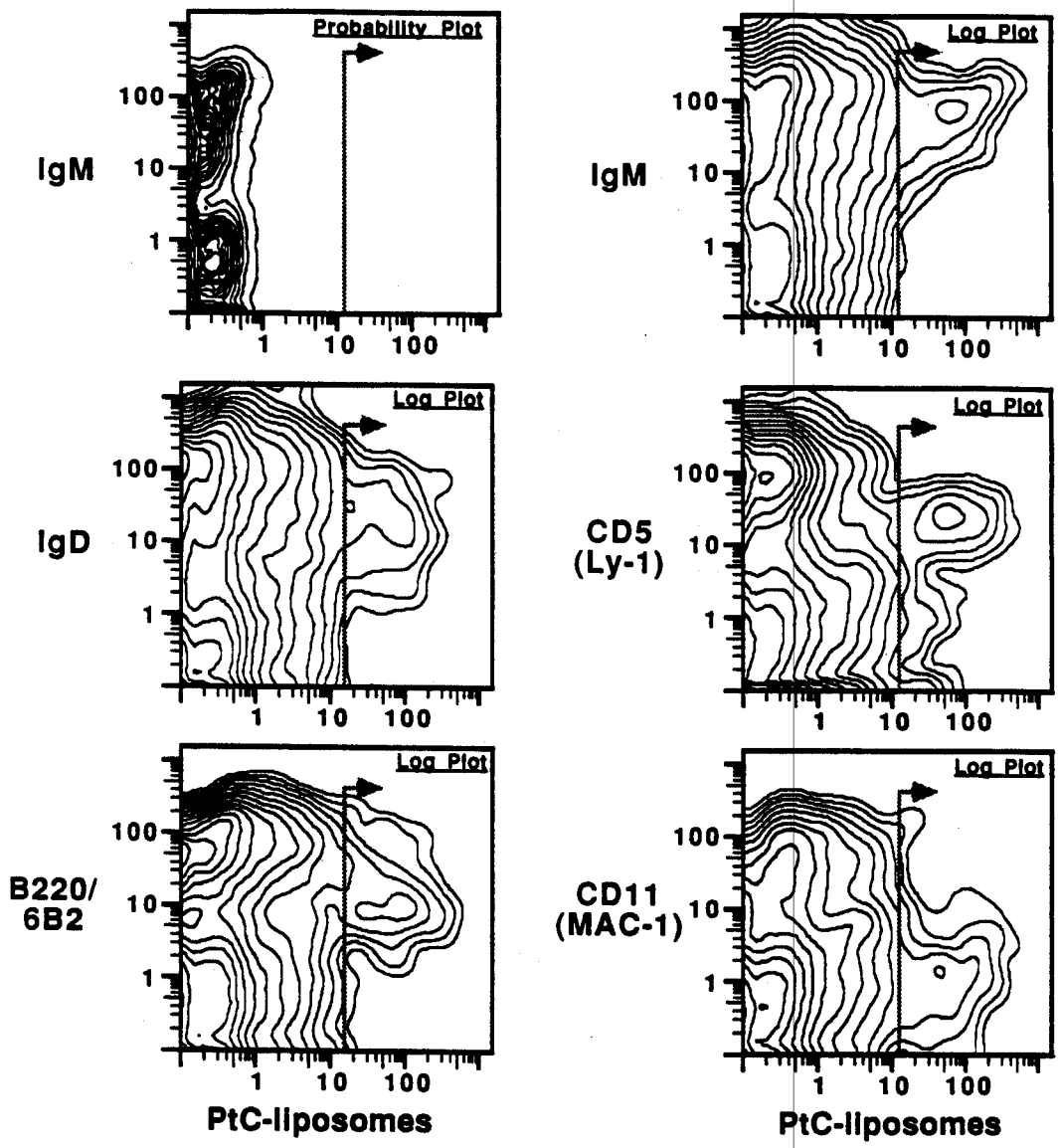

FIG. 10. Cell surface phenotype of phosphatidylcholine–liposome binding (anti-BrMRBC) cells in the spleen.

changing their fluorescence spectrum are much more useful than those that just change in fluorescence signal level since fluorescence spectrum shifts can be measured on a cell-by-cell basis. The spectrum shift probes include BECEF [2',7'-bis-carboxyethyl-5(6)-carboxyfluorescein] (18) for pH and Indo-1 for Ca^{2+} (19). When excited by ultraviolet light, free Indo-1 fluoresces in the blue (peak wavelength 485 nm), but Indo-1 bound to calcium fluoresces violet (peak wavelength 405 nm). Inside a cell, free and Ca-bound Indo-1 are in an equilibrium with the intracellular free Ca^{2+}, so the balance between blue and violet fluorescences is a measure of free Ca^{2+} concentration.

When lymphocytes are activated, a cascade of biochemical events leads to proliferation and possibly differentiation. The proliferation assays commonly used to monitor activation are carried out several days after the initial stimulus and require that a whole series of events has occurred in the meantime. On the other hand, changes in free Ca^{2+} may be observed within seconds of the application of a stimulus, making it possible to investigate early events in activation even if the process is not carried through to proliferation (20).

T cell clones have been produced which respond to the presence of other cells that carry specific cell surface molecules and do not respond to cells lacking those molecules. Indo-1 has been used to demonstrate that the initial steps in activation occur quite promptly when appropriate cells are brought into contact (P. Nghiem, unpublished results). In the experiment illustrated in Fig. 12, Indo-1-labeled responder cells were mixed with appropriate or inappropriate stimulators, centrifuged briefly to bring the cells into contact, incubated at 37°C for 30 sec, resuspended, and run on the FACS immediately.

With appropriate stimulators, most of the responder cells showed a large and transient increase in violet/blue

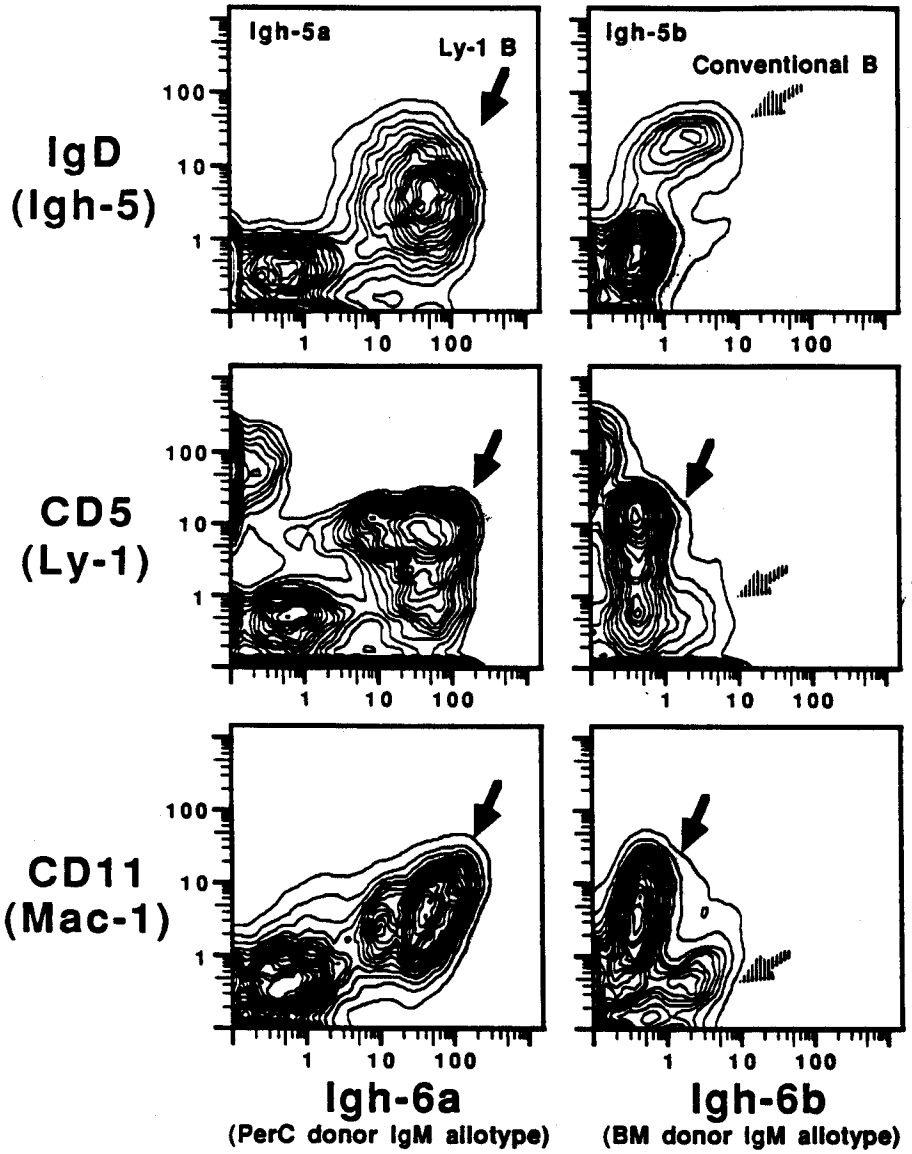

FIG. 11. Peritoneal cells selectively reconstitute the Ly-1 B lineage in BM/PerC chimeras.

ratio, indicating a temporary increase in free Ca^{2+} (Fig. 12). Specific responders in a mixture of T cell clones have been enriched by sorting cells with high violet/blue ratio after exposure to stimulator cells. This procedure may lead to more rapid and efficient isolation of new antigen specific T cell clones as well as better understanding of activation processes.

Gene Cloning, Expression, and Regulation

The ability to identify, sort, and clone rare cells has made the FACS increasingly useful in molecular and genetic studies with mammalian cells in culture. FACS instruments have been used, for example, to isolate Ig isotype switch variants and antibody affinity mutants, to clone genes for lymphocyte surface molecules, and to study regulatory elements that control gene expression. Most of this work has been done with genes that encode or control the expression of Ig or other cell surface molecules detectable with typical fluorochrome-coupled antibody reagents. However, FACS techniques have recently been developed that make it possible to study gene regulatory elements and hormones and other factors that control the expression of genes encoding intracellular enzymes whose activity is measurable in individual viable cells as a function of the amount of an internal fluorescent dye generated by enzymatic cleavage of a fluorogenic substrate (21). Some of these applications are discussed next.

FACS isolation of rare variants from cultured cell lines frequently requires several sort-and-grow cycles to progressively increase the frequency of variants within the population. Cloned variant lines can then be obtained by sorting individual cells (with the appropraite FACS phenotype) into culture wells and screening the resultant clones for expression of variant characteristics.

The initial FACS studies in somatic cell genetics mainly relied on one-color methods in which individual fluorochrome-coupled reagents were used to sort marker-positive transfectants from a negative parent cell line or vice versa (22) or to isolate variants that no longer express

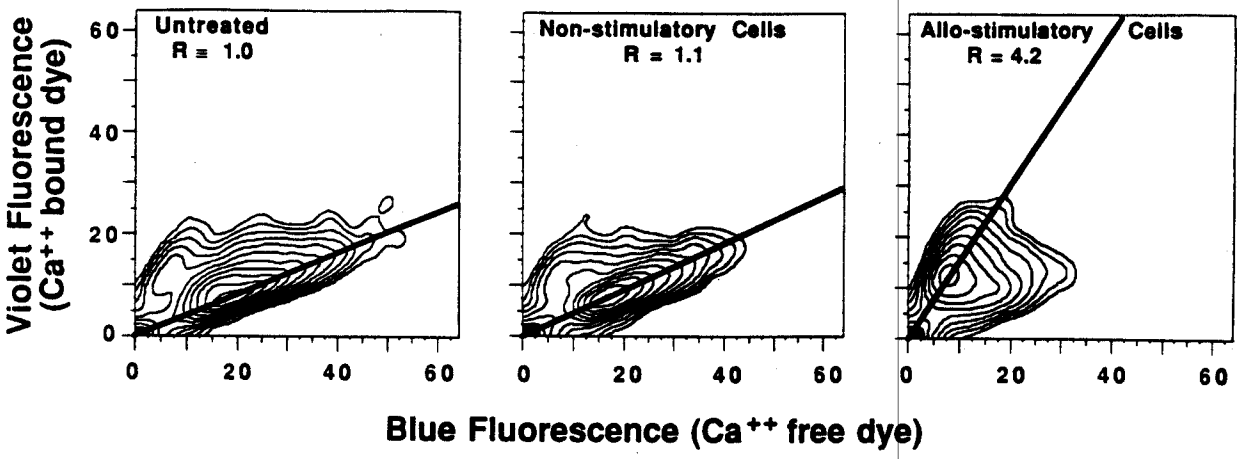

FIG. 12. Response of an Indo-1-labeled alloreactive T cell clone to appropriate and inappropriate "stimulator" cells. Indo-1 fluorescence is predominantly blue in the absence of Ca^{2+} and shifts to the violet when Ca^{2+} is bound. The violet/blue ratio (R, defined here to be 1.0 for the resting cells shown in the **left panel**) is indicative of the Ca^{2+} concentration to which the dye is exposed with higher ratios corresponding to higher Ca^{2+} concentrations. Clone B10 cells proliferate when exposed to clone MAK cells but not when exposed to clone MAR cells. Bringing B10 and "stimulator" cells into contact by centrifugation has little effect when clone MAR is used (**middle panel**) but with clone MAK a large increase in free intracellular Ca^{2+} concentration is observed (**right panel**). This effect occurs on a time scale of minutes rather than the several days needed for a proliferation assay. The figure panels all show logarithmic contours at twofold intervals.

particular MHC-encoded epitopes but still express the (MHC-encoded) surface molecules on which those epitopes are expressed (23). Recent methods for isolating rare mutants, however, are more sophisticated and often use two or more reagents to increase the selectivity of the procedure. For example, variants expressing an altered form of a surface molecule can now be sorted efficiently by using one fluorescent reagent to monitor the amount of the molecule on each cell and a second reagent (coupled to a different fluorescent label) to detect cells where structural alterations of the molecule result in the loss of a determinant or a change in antigen-binding properties.

These multiparameter FACS methods have been used to isolate antigenic "loss variants" that either no longer express a particular cell surface antigen or have lost a particular epitope (antigenic determinant) but still express the molecule on which the epitope was carried in the parent cell line. Similar methods have been used with hybridomas to isolate rare spontaneous mutants that either show a change in the binding of cell surface antibody to antigen or show a shift in the relative binding of two closely related haptens. Functional, structural, sequencing, and computer modeling studies with the antibodies produced by affinity and fine specificity variants such as these have yielded surprising insights into the ways in which antibody structure determines antigen-binding properties (24).

FACS analysis and sorting strategies have also been used to clone several genes encoding lymphocyte surface molecules and to study the expression of these genes when introduced into lymphoid and nonlymphoid cell lines. Genes coding for the murine and human CD5 and CD8 lymphocyte surface antigens, for example, were cloned by transfecting L cells with cellular DNA and using the FACS to isolate and clone rare cells ($1/10^4$ to $1/10^5$) expressing the genes. After several weeks in culture, further FACS selection was applied to select cells that expressed amplified levels of CD5 (or CD8).

Once isolated, these amplified clones were used as a source of mRNA for selecting cDNA clones coding for CD5 (or CD8). The identity of the isolated cDNA clone was verified by putting it in an expression vector, transfecting it back into L cells and showing that the transfected cells then expressed the surface marker (CD5 or CD8). Finally, the cDNA clones were used to isolate and sequence genomic clones coding for each of the markers.

The expression of the transfected CD5 and CD8 genes was amplifiable in about half the marker-positive L cell lines obtained from the initial transfection series. The mechanisms underlying this amplification appear similar to those involved in the amplification of the gene that codes for dihydrofolate reductase (DHFR), an enzyme whose content in cells can be measured by the stoichiometric binding of methotrexate-fluorescein, a fluorescent, noncleavable substrate analog (25). In both systems, the FACS can be used to select cell lines whose expression of the gene product increases progressively with selection and falls when the selected lines are grown for some time in the absence of FACS (or other) selection. Furthermore, in both systems, the amplified genes may be located in an amplifiable chromosomal site; however, they are often found on acentric chromosomes ("double-minutes") whose increase (or decrease) in frequency during selection is responsible for the observed changes in the overall expression of the gene product (26,27).

To look more closely at the molecular and cellular mechanisms regulating gene expression, FACS methods have now been developed that measure the expression of

a bacterial "reporter" gene under the control of various kinds of regulatory elements. In essence, recombinant DNA methods are used to introduce the *E. coli* gene coding for β-galactosidase (*lacZ*) into mammalian cells growing in culture and the FACS is used to identify, sort, and clone cells expressing various levels of β-galactosidase activity. This activity is revealed in individual cells by using the FACS to detect intracellular fluorescein generated by the cleavage of the fluorogenic β-galactosidase substrate, fluorescein-di-β-D-galactopyranoside (that can be loaded into the cells without decreasing viability) (21).

This technique is quite versatile since many different cell types can be infected or transfected with a wide range of viral or recombinant DNA constructs to introduce *lacZ*, either alone or in association with cloned mammalian regulatory elements (promoters, enhancers) that control the expression of physiologically important genes. With this method, cloned cell lines can be isolated in which the level of *lacZ* expression reflects regulatory activity controlling the integrated *lac-Z*-containing construct due, for example, to integration near active nearby promoter and enhancer regions, or to alterations of the intracellular level of regulatory molecules (trans-acting factors) that control the expression of genes associated with particular enhancers or promoters (28).

The use of this *lacZ* "reporter gene" system also makes it possible to identify promoter and enhancer sites that undergo a change in status during differentiation, that is, off to on or vice versa. This approach is illustrated by an experiment in which *lacZ* was introduced into 70Z pre-B cells on a transcriptionally disabled viral construct carrying a functional promoter but no enhancer sequences. With this construct, *lacZ* should only be expressed in cells in which the construct has become integrated into a chromosomal site near an active enhancer, and expression should cease in cells whose *lacZ* construct is situated near a stage specific enhancer that becomes inactive when the cells are triggered to differentiate.

The 70Z pre-B cell line was chosen for these experiments because it can be stimulated with LPS to differentiate from the pre-B cell state in which it normally grows to a more mature B cell state expressing high levels of surface Ig. This differentiated state is maintained indefinitely in the presence of stimulating levels of LPS; however, the cells return to the pre-B state once the LPS is removed. Thus to identify stage specific enhancers, the transcriptionally disabled (enhancerless) *lacZ* construct was introduced into 70Z cells growing in the pre-B state (without LPS) and *lacZ*-expressing cells were selected, grown, and tested for loss of *lacZ* expression following stimulation with LPS (28).

As might be expected, most of the 70Z cells that expressed *lacZ* in the absence of LPS continued to do so after LPS stimulation, indicating that the integration sites of the active *lacZ* constructs were mostly near enhancers that are equally active in pre-B and B cells. Nevertheless, one clone was isolated that terminates *lacZ* expression and initiates high-level surface Ig expression when stimulated with LPS and recovers *lacZ* expression and terminates Ig production when the LPS is removed (see Fig. 13). Thus, in this cell line, *lacZ* appears to have integrated next to a truly stage specific enhancer that loses activity when the cells differentiate to mature B cells and regains activity when the cells return to the pre-B state (28).

The use of the FACS in these kinds of molecular studies clearly opens a broad new approach to the use of recombinant methods in mammalian cells. Hopefully, the merging of these technologies will prove as fruitful as the earlier merger of monoclonal antibody and FACS methods, which now constitute key tools in the clinical laboratory and in the fight against AIDS.

ACKNOWLEDGMENTS

The authors thank Alan Stall for data used in Figs. 3 and 5 to 11 and for artwork on the figures. We thank

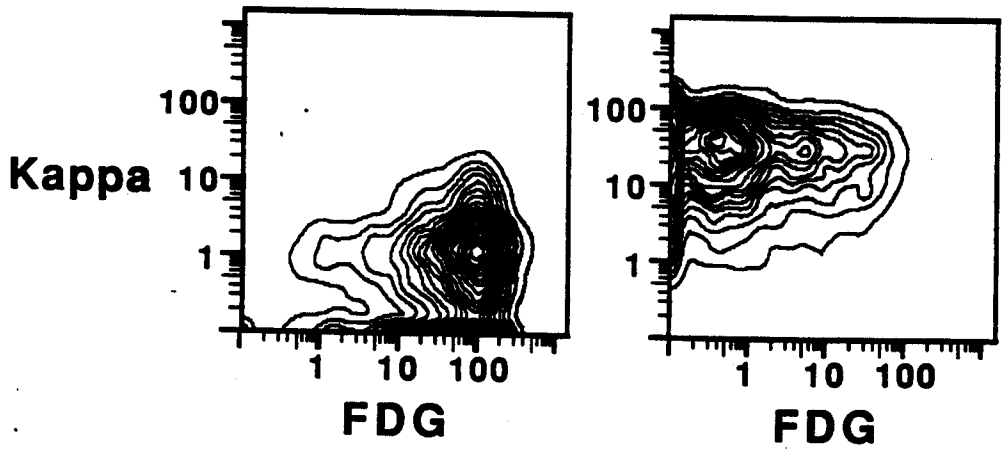

FIG. 13. Reciprocal regulation of *lacZ* and κ light chain expression following LPS induction of 70Z/3 cells. Samples of clone 7e17, a *lacZ*-expressing derivative of 70Z/3, were treated with 10 μg/ml LPS for 24 hr (**right panel**) or kept in normal media (**left panel**) at equivalent cell density. The 7e17 cells were loaded with FDG and maintained near 0°C for 2 hr during which they were stained for cell surface κ light chains using anti-κ-biotin and avidin–Texas Red. Phenylethylthio-β-D-galactoside (PETG, 1 mM) was added to halt FDG cleavage, and the samples were analyzed on a two-laser FACS.

Richard Stovel for Fig. 1, Lisa Peterson for the spectra in Fig. 2, Diether Recktenwald for the plots in Fig. 4, Paul Nghiem for the Indo-1 data in Fig. 12, William Kerr and Garry Nolan for the *lacZ* data in Fig. 13, and Martin Bigos for the calculations for Table 1. Data shown in Figs. 3 and 5 to 13 were collected and plotted using the FACS/DESK flow cytometry software developed by Wayne Moore at Stanford University. Histogram or contour plot images produced on a VAX computer were transferred to a Macintosh computer where the axes and labels were recomposed.

REFERENCES

1. Hulett, H. R., Bonner, W. A., Barrett, J., and Herzenberg, L. A. (1969): Cell sorting: Automated separation of mammalian cells as a function of intracellular fluorescence. *Science*, 166:747–749.
2. Bonner, W. A., Hulett, H. R., Sweet, R. G., and Herzenberg, L. A. (1972): Fluorescence activated cell sorting. *Rev. Sci. Instrum.*, 43:404–409.
3. Parks, D. R., and Herzenberg, L. A. (1984): Fluorescence activated cell sorting: Theory, experimental optimization and applications in lymphoid cell biology. In: *Methods in Enzymology* Vol. 108, edited by S. P. Colowick and N. O. Kaplan, pp. 197–241. Academic, Orlando, FL.
4. Parks, D. R., Lanier, L. L., and Herzenberg, L. A. (1986): Flow cytometry and fluorescence activated cell sorting. In: *The Handbook of Experimental Immunology*, 4th ed., edited by D. M. Weir, L. A. Herzenberg, C. C. Blackwell, and L. A. Herzenberg, pp. 29.1–29.21. Blackwell Scientific, Edinburgh.
5. Ledbetter, J. A., Evans, R. L., Lipinski, M., Cunningham-Rundles, C., Good, R. A., and Herzenberg, L. A. (1981): Evolutionary conservation of surface molecules that distinguish T lymphocyte helper/inducer and cytotoxic/suppressor subpopulations in mouse and man. *J. Exp. Med.*, 153:310–323.
6. Reinherz, E. L., and Schlossman, S. F. (1980): The differentiation and function of human T lymphocytes. *Cell*, 19:821–827.
7. Alberti, S., Parks, D. R., and Herzenberg, L. A. (1987): A simple method for subtraction of cell autofluorescence in flow cytometry. *Cytometry*, 8:114–119.
8. Jensen, C. P., and Muirhead, K. A. (1988): A novel method for live/dead discrimination in fixed samples. *Cytometry [Suppl.]*, 2:77.
9. Weir, D. M., Herzenberg, L. A., Blackwell, C. C., and Herzenberg, L. A. (eds.) (1986): *The Handbook of Experimental Immunology*, 4th ed. Blackwell Scientific, Edinburgh.
10. Herzenberg, L. A., Stall, A. M., Lalor, P. A., Sidman, C., Moore, W. A., Parks, D. R., and Herzenberg, L. A. (1986): The Ly-1 B cell lineage. *Immunol. Rev.*, 93:81–102.
11. Moore, W., and Kautz, R. (1986): Data analysis for flow cytometry. In: *The Handbook of Experimental Immunology*, 4th ed., edited by D. M. Weir, L. A. Herzenberg, C. C. Blackwell, and L. A. Herzenberg, pp. 30.1–30.11. Blackwell Scientific, Edinburgh.
12. Herzenberg, L. A., Stall, A. M., Braun, J., Weaver, D., Baltimore, D., Herzenberg, L. A., and Grosschedl, R. (1987): Depletion of the predominant B cell population in immunoglobulin mu heavy chain transgenic mice. *Nature*, 329:71–73.
13. Stall, A. M., Farinas, M. C., Tarlinton, D. M., Lalor, P. A., Herzenberg, L. A., Strober, S., and Herzenberg, L. A. (1988): Ly-1 B cell clones similar to human chronic lymphocytic leukemias (B-CLL) routinely develop in older normal mice and young autoimmune (NZB-related) animals. *Proc. Natl. Acad. Sci. USA*, 85:7312–7316.
14. Lalor, P. A., Stall, A. M., Adams, S., Herzenberg, L. A., and Herzenberg, L. A. (1989): Feedback regulation of murine Ly-1 B cell development. *Eur. J. Immunol.* (in press).
15. Mercolino, T. J., Arnold, L. W., and Haughton, G. (1986): Phosphatidyl choline is recognized by a series of Ly-1$^+$ murine B cell lymphomas specific for erythrocyte membranes. *J. Exp. Med.*, 163:155–165.
16. Mercolino, T. J., Arnold, L. W., Hawkins, L. A., and Haughton, G. (1988): Normal mouse peritoneum contains a large population of Ly-1$^+$ (CD5) B cells that recognize phosphatidyl choline. Relationship to cells that secrete hemolytic antibody specific for autologous erythrocytes. *J. Exp. Med.*, 168:687–698.
17. Hayakawa, K., Hardy, R. R., Herzenberg, L. A., and Herzenberg, L. A. (1985): Progenitors for Ly-1 B cells are distinct from progenitors from other B cells. *J. Exp. Med.*, 161:1554–1568.
18. Musgrove, E., Rugg, C., and Hedley, D. (1986): Flow cytometric measurement of cytoplasmic pH: A critical evaluation of available fluorochromes. *Cytometry*, 7:347–355.
19. Grynkiewicz, G., Poenie, M., and Tsien, R. Y. (1985): A new generation of Ca^{2+} indicators with greatly improved fluorescence properties. *J. Biol. Chem.*, 260:3440–3445.
20. Nisbet-Brown, E., Cheung, R. K., Lee, J. W. W., and Gelfand, E. W. (1985): Antigen-dependent increase in cytosolic free calcium in specific human T-lymphocyte clones. *Nature*, 316:545–547.
21. Nolan, G. P., Fiering, S., Nicolas, J.-F., and Herzenberg, Leonard A. Fluorescence activated cell analysis and sorting of viable mammalian cells based on beta-D-galactosidase activity after transduction of *E. coli LacZ*. *Proc. Natl, Acad. Sci. USA*, 85:2603–2607.
22. Kavathas, P., and Herzenberg, L. A. (1983): Stable transformation of mouse L cells for human membrane T cell differentiation antigens, HLA and B_2-microglobulin: Selection by fluorescence activated cell sorting. *Proc. Natl. Acad. Sci. USA*, 80:524–528.
23. Holtkamp, B., Cramer, M., Lemke, H., and Rajewsky, K. (1981): Isolation of a cloned cell line expressing variant H-2K^k: A method that permits selection of structural variants of cell surface molecules by fluorescence-activated cell sorting. *Nature*, 289:66–68.
24. Panka, D. J., Mudgett-Hunter, M., Parks, D. R., Peterson, L. L., Herzenberg, L. A., Haber, E., and Margolies, M. N. (1987): Variable region framework differences result in decreased or increased affinity of variant anti-digoxin antibodies. *Proc. Natl. Acad. Sci. USA*, 85:3080–3084.
25. Johnston, R. N., Beverley, S. M., and Schimke, R. T. (1983): Rapid spontaneous dihydrofolate reductase gene amplification shown by fluorescence activated cell sorting. *Proc. Natl. Acad. Sci. USA*, 80:3711–3715.
26. Kaufman, R., Brown, P. C., and Schimke, R. T. (1979): Amplified dihydrofolate reductase genes in unstably methotrexate-resistant cells are associated with double minute chromosomes. *Proc. Natl. Acad. Sci. USA*, 76:5669–5673.
27. Kavathas, P., and Herzenberg, L. A. (1983): Amplification of a gene coding for a human T cell differentiation antigen. *Nature*, 306:385–387.
28. Kerr, W. G., Nolan, G. P., and Herzenberg, L. A. (1989): *In situ* detection of transcriptionally-active chromatin and genetic regulatory elements in individual viable mammalian cells. *Immunology* (in press).