ELSEVIER

Journal of Immunological Methods 243 (2000) 77–97

JIM
Journal of
Immunological Methods

www.elsevier.nl/locate/jim

A practical approach to multicolor flow cytometry for immunophenotyping

Nicole Baumgarth[a,1], Mario Roederer[b,*]

[a]*Department of Genetics, Stanford University Medical School, Beckman Center B-007, Stanford, CA 94305-5318, USA*
[b]*Department of Stomatology, UCSF, San Francisco, CA, USA*

Abstract

Through a series of novel developments in flow cytometry hardware, software, and dye-chemistry it is now possible to simultaneously measure up to 11 distinct fluorescences and two scattered light parameters on each cell. Such advanced multicolor systems have a number of advantages over current two- and three-color flow cytometric measurements. They provide a large amount of novel information for each sample studied, an exquisitely accurate quantitation of even rare cell populations, and allow identification and characterization of novel cell subsets. In particular, this technology is proving crucial to identifying functionally homogeneous subsets of cells within the enormously complex immune system; such identification and enumeration is important for understanding disease pathogenesis. However, multicolor flow cytometry comes with a new and sometimes difficult set of technical problems that must be overcome by users to derive meaningful results. In this manuscript, we describe the basic aspects of multicolor flow cytometry, including the technical hurdles and artefacts that may occur, and provide some suggestions for how to best overcome these hurdles. While inspired by the 11-color technology that we currently use, these principles apply to all flow cytometric experiments in which more than one fluorescent dye is used.

Keywords: Flow cytometry; Immunofluorescence; Immunophenotyping; Multiparameter flow cytometry

1. Introduction

Since the earliest application of flow cytometry to the study of cells, there has been a drive to increase the number of distinct measurements for each cell. This developmental effort blossomed in the late 1990s, when the number of independently measurable ‘colors’ (each color corresponds to a distinct fluorescence-based measurement of a cellular protein or function) increased from five to 11 (Roederer et al., 1997; Bigos et al., 1999).

The success of this developmental effort was due to the coordinated development of new hardware, new fluorochromes, and new software analysis tools

Abbreviations: APC, allophycocyanin; A595, Alexa 595; CasB: Cascade Blue; CasY, Cascade Yellow; Cy5.5APC, Cy7APC, Cy5.5, Cy7 conjugates of allophycocyanin; Cy5PE, Cy5.5PE, Cy7PE, Cy5, Cy5.5, and Cy7 conjugates of phycoerythrin; Cy5.5PerCP: Cy5.5 tandem conjugate of PerCP; FITC, fluorescein; PBP: phycobiliprotein (i.e., PE or APC); PE, phycoerythrin; PerCP: peridinin chlorophyll protein; PFC, polychromatic flow cytometry; PMT, photomultiplier tube; TR, Texas Red; TRPE, Texas Red-conjugated phycoerythrin; SA, streptavidin

*Corresponding author. Tel.: +1-415-514-0395; fax: +1-415-476-4204.

E-mail address: roederer@drmr.com (M. Roederer).

[1]Present address: Center for Comparative Medicine, University of California, Davis, CA 95616, USA.

0022-1759/00/$ – see front matter
PII: S0022-1759(00)00229-5

that significantly increase the quality and quantity of measurements. This increase comes with a price, however, as this new technology has its own set of technical problems and difficulties that users must be aware of and must overcome in order to derive meaningful results. Nonetheless, once these hurdles are overcome, this new technology is well worth the effort, as the information obtained from the measurements is not only novel but could not be obtained otherwise using standard flow cytometric techniques.

As we outline below, setting up multicolor flow cytometry is not simply achieved by adding new reagents to existing reagent combinations, but requires a more involved process of quality control, optimization, and 'debugging'. Therefore, to distinguish multicolor flow cytometry with its unique benefits and technical problems from current standard flow cytometric technologies (i.e., using four or fewer fluorescent dyes), we refer to it as 'polychromatic flow cytometry' (PFC) and will use the term throughout this manuscript.

Flow cytometers capable of collecting data for more than three or four colors are now becoming more prevalent, as manufacturers have recognized the significant demand for the types of analysis afforded by this technology. Given the bewildering array of fluorochromes, lasers, hardware, and software that might be used in PFC, we outline here the fundamental requirements, interactions, and problems associated with setting up this technology, so that users can make educated decisions about instrument requirements and the design of their experiments. Finally, we provide some examples in which this technology has been of particular benefit. We hope to provide with this brief review some practical tips and encouragement for those thinking of expanding their current flow cytometric measurements. The benefits of true multicolor flow cytometry make this technique a particularly useful and probably soon an irreplaceable tool for the study of cell biology and immunology.

2. Fluorochromes

The ability to measure multiple parameters of each cell is limited by the number of fluorochromes that can be simultaneously measured. The 11-color PFC that is currently in routine use at Stanford uses dyes excited by three different laser lines. The excitation and emission spectra of these dyes and the filters that were chosen to collect the emitted light from these dyes are shown in Fig. 1.

2.1. Characteristics of useful fluorochromes

When designing experiments for the flow cytometer that include the use of new dyes, careful consideration must be given to the choice of fluorochromes. Desirable fluorochromes for cytometric technologies have several properties: they (i) are biologically inert; (ii) have high cell-associated fluorescence intensities ('bright'); (iii) exhibit little spectral overlap amongst each other; and (iv) for immunophenotyping, are easily conjugated to monoclonal antibodies.

2.1.1. Biological inertness

Most of the fluorochromes that are currently in use are biologically inert: i.e., they do not affect the cells, nor do they bind directly to cellular elements. There are, however, exceptions to this: the most common example is that of the 'background' binding of Cy5PE (and other Cyanine–PBP tandem dyes) to monocytes and B cells. This background binding is variable between species and can be extremely high in some instances. For example, Cy5PE binds strongly to B cells in mice with autoimmune disorders (e.g., non-obese diabetic mice). While there are methods available to reduce this background, it is of concern when the particular cell types being studied are those that interact 'nonspecifically' with the fluorochrome.

2.1.2. Cell-associated fluorescence intensity

With regard to the high fluorescence intensities or the 'brightness' of a fluorochrome, it should be noted that the characterization of a fluorescence signal as 'bright', i.e., the difference between the unstained and the stained cells, is still empirical. 'Bright' signals result from fluorochromes with the following characteristics: (1) a high extinction coefficient; (2) a high quantum yield; (3) an emission spectrum overlapping as little as possible with cellular auto-

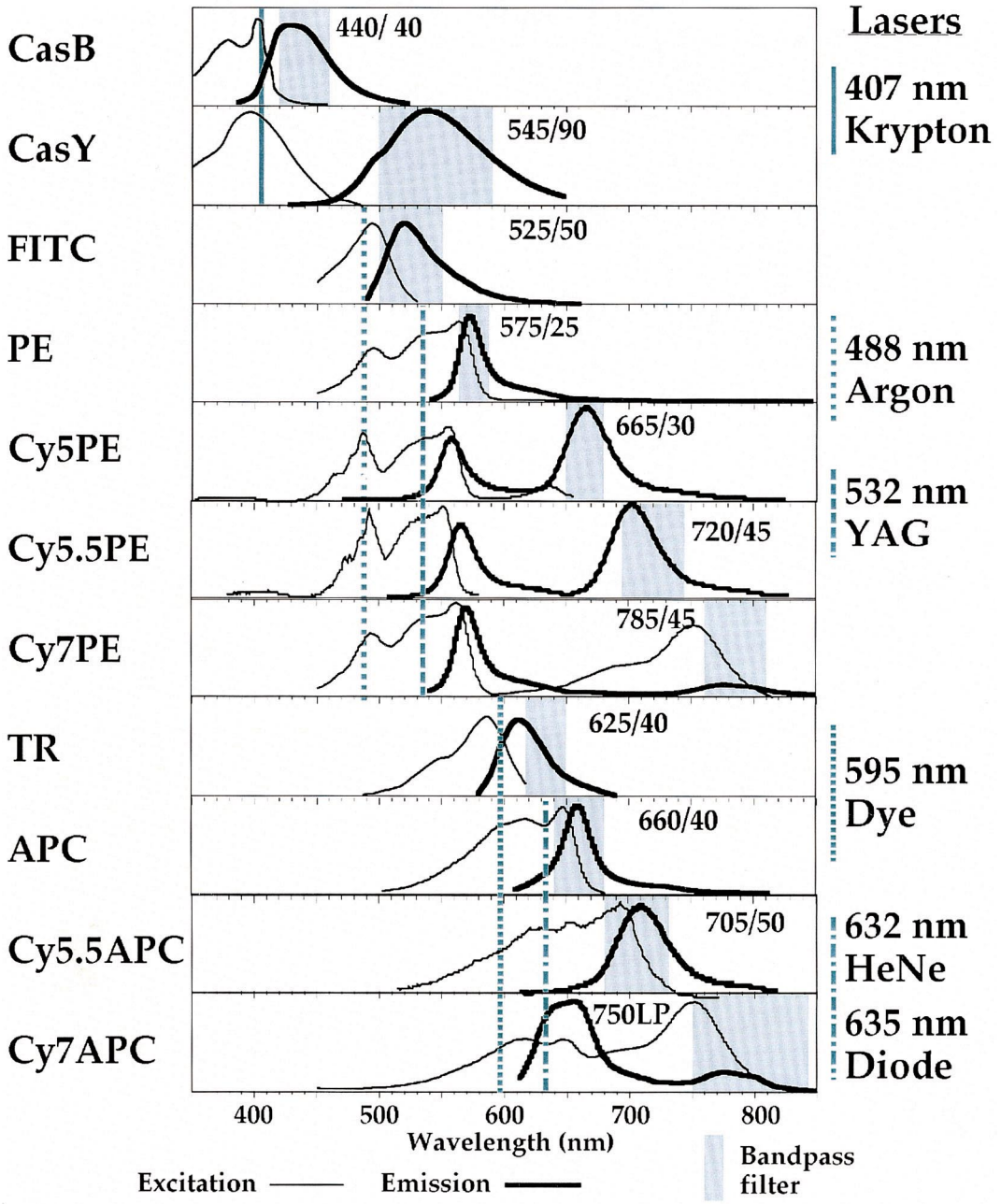


Fig. 1. Shown are excitation and emission spectra of the 11 dyes that we currently use for polychromatic flow cytometry (PFC). Indicated are the wavelengths of the different laser lines used for their excitation (407, 488 and 595 nm; other laser lines that can be used for some of the dyes are shown for reference), and the filters chosen for optimal light collection and minimal spillover. The fluorochromes, laser lines and filters are currently in routine use on a modified three-laser hybrid Becton-Dickinson/Cytomation flow cytometer, described previously (Roederer et al., 1997; Bigos et al., 1999).

fluorescence; (4) measurable with high sensitivity detectors; and (5) the ability to conjugate multiple fluorochromes to each detecting unit (e.g., a high ratio of fluorochrome to antibody).

The differences in brightness can be quantitated experimentally by conjugating the same monoclonal antibody to the various fluorochromes at optimized ratios. Fig. 2 shows histograms of human peripheral blood lymphocytes that were stained with 11 different conjugates of an anti-CD8 monoclonal antibody, one conjugate for each of the 11 fluorochromes that we currently use in PFC. As shown in this figure, the 'brightness' is affected not only by the intensity with which the positive cells are stained, but also by the background of the negative cells. Although some of the conjugates appear to stain much 'brighter' than others, all of these fluorochromes are useful for clearly distinguishing the positively stained cells from the unstained cells.

It is important to point out that the brightness of a fluorochrome will differ depending on the instrument used. For example, conjugates of antibodies to PE generally result in much brighter staining when cells are analyzed with typical benchtop cytometers using flow cells as compared to cytometers using jet-in-air detection. This is due to a number of differences in the optical paths that affect the efficiency with which emitted light is collected by the detectors; however, other differences (notably the laser power) also influence sensitivity. Therefore, reagent 'brightness' is a relative measurement and must be assessed by users on any given instrument.

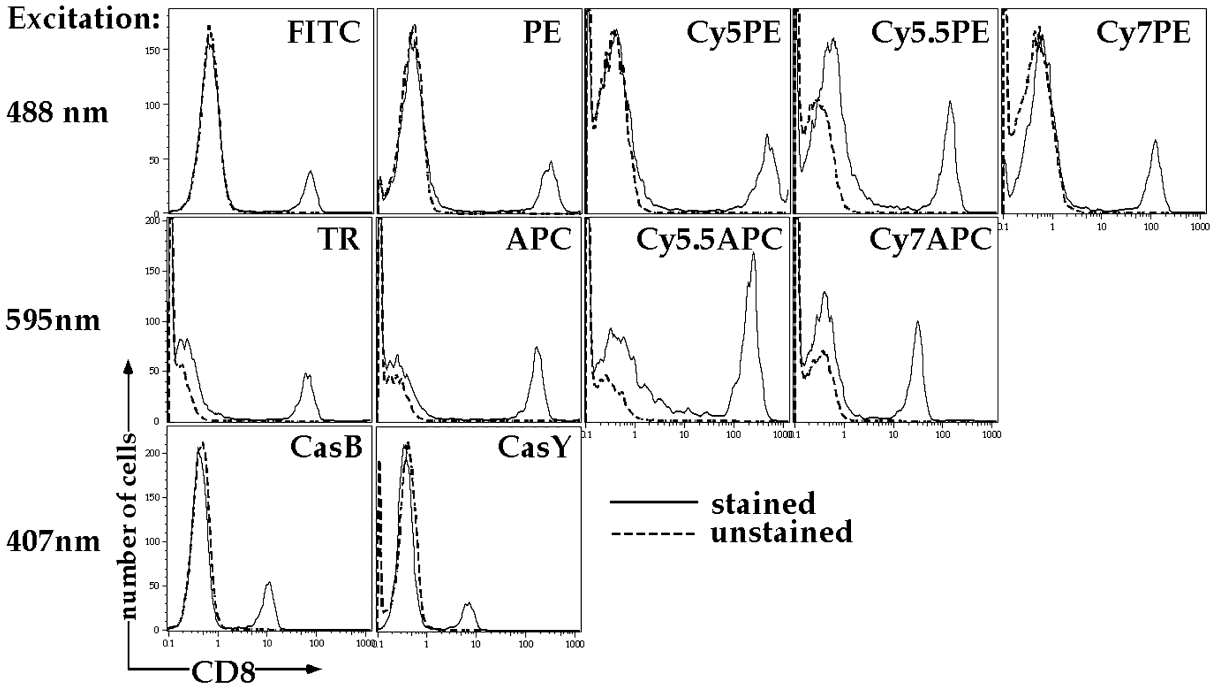


Fig. 2. An anti-human CD8 monoclonal antibody was conjugated at optimal fluorochrome–antibody ratios to each of the indicated fluorochromes (Roederer, 1997a,b). Human peripheral blood lymphocytes, purified by gradient centrifugation over Ficoll, were stained with each of the indicated anti-CD8 conjugates and with an antibody to the pan-T cell marker CD3, conjugated either to FITC (for CasB and CasY graphs) or to CasB (all others). Shown are histograms of cells that stained positively for CD3. The dashed line indicates levels of staining for cells stained with anti-CD3 only, and the solid line indicates the anti-CD8 fluorescence for cells stained with both antibodies. Empirically, the relative 'brightness' of any of these fluorophores can be estimated based on the separation of CD8 positive and negative cells. The apparent difference in the number of cells between the positive and negative peaks for these dyes is a visual artefact of the scaling (note that the 'CD8-dim' cells appear to be more prevalent in those plots where the CD8 peak is higher). There was no difference in the percent dim or bright-positive cells for any of distributions.

2.1.3. Spectral overlaps

The spectral overlaps that exist between dyes currently represent the biggest hurdle in PFC analysis. Due to the similarities and overlaps in the emission spectra of different dyes (see Fig. 1), it is not possible to choose emission filters that uniquely measure only one of the dyes in a multicolor experiment. The appropriate choice of filter, however, can greatly reduce collection of light from other fluorochromes (i.e., reduce 'spillover'). Note from Fig. 1 that a set of up to four dyes excited by three different lasers can be chosen that exhibit essentially no spectral overlap; thus, four- or fewer-color experiments can be carefully designed to avoid the problems caused by spillover. However, the necessity of having three lasers for a four-color experiment renders this an impractical solution.

Because of the spectral overlap, each fluorochrome will contribute a signal to several detectors, therefore the contribution in detectors not assigned to that fluorochrome must be subtracted from the total signal in those detectors. This process, termed 'compensation', is discussed in detail below.

2.1.4. Conjugation to antibodies

For many applications of flow cytometry, a further requirement for a fluorochrome is that it can be readily conjugated to antibodies. The 11 dyes we commonly use (Figs. 1 and 2) meet this criterion. There are, however, varying degrees of difficulty involved in carrying out the conjugation reactions. It is likely that any laboratory carrying out flow cytometric analyses with five or more dyes will need to conjugate antibodies, as commercial conjugates are not yet available for a number of these dyes. Even if all of the dyes become available as antibody conjugates, it is unlikely that every combination of antibody and fluorochrome necessary for each user will be available.

Detailed procedures for conjugating the dyes can be found on the web (Roederer, 1997a) or from commercial vendors who supply conjugation kits. Briefly, the procedures for conjugating these dyes fall into three classes. (1) Small organic molecules (FITC, TR, A595, A430, CasB, and CasY) are conjugated to antibodies in relatively simple, short reactions that require only one or two column separations (or other desalting procedures). (2) Single protein molecules (PE, APC) require a slightly more complex procedure, but reactive fluorochromes can be easily prepared, are stable for long periods of time, can be used in a simple conjugation procedure. (3) Tandem dyes (Cy5PE,[2] Cy5.5PE, Cy7PE, Cy5.5APC, Cy7APC, Cy5.5PerCP), require the use of more complex procedures which involve first the generation of the reactive tandems that can then, in a second step, be used for conjugation to antibodies. Generating the reactive tandems requires careful testing of different conditions for each batch of tandem dye to optimize the ratio of the two dyes used. However, once the chemistry has been optimized, a large batch of reactive dye can be prepared for use in many conjugations.

2.2. Other fluorochromes.

New fluorochromes for immunophenotyping applications are constantly being developed. At the time of writing, several dyes exist that are useful for immunophenotyping applications and which can be used in addition to or instead of some dyes listed in Fig. 1.

2.2.1. PerCP

This is a chlorophyll-like protein that can be directly conjugated to antibodies. It has an emission spectrum similar to Cy5PE, but with no excitation in the red, and hence very little spillover into the APC detector compared to Cy5PE. On the other hand, it has several disadvantages: it is easily 'bleached' by lasers, limiting its utility to low-power instruments such as benchtop instruments; it is less bright than Cy5PE; and it is only available in conjugates from a single vendor (Becton-Dickinson, San Jose, CA).

2.2.2. Cy5.5PerCP

A relatively new tandem conjugate of PerCP that has excitation and emission spectra very similar to Cy5.5PE. Unlike PerCP, it can be used in high power laser instruments as it does not 'bleach' easily, but is also only available in conjugated form through Becton-Dickinson.

[2]Cy5PE is the same as 'Tricolor', Cy7APC is the same as 'Allo-7' and 'PharRed', and Cy5.5PerCP is the same as 'TruRed'.

2.2.3. Alexa 430, 595

A series of new dyes termed 'Alexa dyes' from Molecular Probes (Eugene, OR) includes two that we have tested for their usefulness in PFC. A430 is an excellent replacement for CasY, in that it has less non-specific interaction with CasB and is otherwise roughly similar in 'brightness' (Baumgarth, unpublished). A595 is spectrally similar to Texas Red (TR) but does not have its non-specific binding characteristics, leading to a somewhat better separation of positive and negative signals than that seen with many TR conjugates.

2.2.4. TRPE

This dye, also called 'Red 613' or 'ECD' is excited by a 488 nm laser line and can replace Cy5PE in a multicolor staining combination, as it has considerably less overlap into the APC channel than Cy5PE. It is however less bright than Cy5PE and it has somewhat more overlap into the PE channel; it also displays considerably higher background binding typical of TR conjugates.

There are a number of other dyes that are commonly used to directly stain cellular components such as DNA (e.g., Hoechst, propidium iodide), cellular membranes (e.g., CFSE, nile red), lysosomes (e.g., CFDA), and dyes that can measure metabolic activity of the cell (e.g., rhodamine 123). It is beyond the scope of this manuscript to cover fluorochromes that are not used for immunophenotyping. For a detailed description of these therefore refer to (Shapiro, 1994). Importantly, however, the use of any combination of dyes, irrespective of the measurement for which it is used, will need to address the same issues such as reagent 'brightness', spillover, and compensation as described here for antibody conjugates.

2.3. Fluorochrome combinations

Out of the dozen or more fluorochromes that can be used for immunophenotyping (by virtue of their spectral and chemical properties discussed above), it is important to select the best combination for a particular experiment under consideration. This is true whether the experiment requires four or 11 colors. To a large extent, the constraints on which fluorochromes can be used will depend on the availability of laser lines on the machine in use (i.e., is a 595-nm dye laser line available, or a 407-nm violet-enhanced krypton laser line?). The useable set of dyes also depends on the number of detectors that are available on the instrument to collect the excited light (i.e., how many colors can be collected that are excited by the first, second, and third laser line?). Once those constraints have been determined, there can still be a wide range of possible combinations. Table 1 lists a number of such combinations, given some constraints of laser and detector availability often seen in currently available machines.

3. Hardware

3.1. Lasers

One of the single largest component costs of cytometry instrumentation is the cost of the lasers. Newer diode lasers are becoming prevalent and these can be significantly cheaper than the older gas ion lasers. When considering a solid-state laser, it is important to choose one that has a long life span and provides a consistent power output, which therefore excludes the very cheap solid-state lasers that are currently available. In addition, for current stream-in-air instrumentation it is desirable to have at least 50 mW of power for each laser line in use, since the fluorescence signal (and thus sensitivity) increases with laser power. For most purposes, it is not necessary to have much more power than this, because most fluorochromes will saturate (are maximally excited) at 100–150 mW. Indeed, additional power may actually *reduce* relative signal to background staining, because the fluorophore will saturate at much lower powers than background fluorescence. Table 2 lists some of the currently available lasers and the lines that are often used in flow cytometry, their approximate costs and the dyes that they can excite.

3.2. Optical setup

Sensitive detection of fluorochromes requires selection of appropriate filters that are placed before each detector, or photomultiplier tube (PMT). Filters must be selected so as to collect as much emitted

Table 1
Sample fluorochrome combinations

Dyes[a]	Lasers[b]	Combinations[c]
4	2 (488+647)	FITC, PE, Cy5.5PE or Cy7PE, APC
		FITC, PE, Cy5PE, Cy5.5APC or Cy7APC
		FITC, PE, Cy5PE, APC
		FITC, PE, TRPE, APC
	1 (488)	FITC, PE, Cy5PE, Cy5.5PE or Cy7PE
5	2 (488+595)	FITC, PE, Cy5.5PE or Cy7PE, TR, APC
		FITC, PE, Cy5PE, TR, Cy5.5APC or Cy7APC
		FITC, PE, Cy5PE, TR, APC
	2 (488+632)	FITC, PE, Cy7PE, APC, Cy5.5APC
		FITC, PE, Cy5.5PE, APC, Cy7APC
		FITC, PE, Cy5PE, Cy5.5APC, Cy7APC
		FITC, PE, Cy5PE, APC, Cy5.5APC or Cy7APC
		FITC, PE, Cy5PE or Cy7PE, APC, Cy7APC
	1 (488)	FITC, PE, Cy5PE, Cy5.5PE, Cy7PE
6	2 (488+633)	FITC, PE, Cy5.5PE, Cy7PE, APC, Cy5.5APC or Cy7APC
		FITC, PE, Cy5PE, Cy5.5PE or Cy7PE, Cy5.5APC, Cy7APC
		FITC, PE, Cy5PE, Cy5.5PE or Cy7PE, APC, Cy5.5APC or Cy7APC
	2 (488+595)	FITC, PE, Cy5.5PE, Cy7PE, TR, APC
		FITC, PE, Cy5PE, Cy7PE, TR, Cy5.5APC
		FITC, PE, Cy5PE, Cy5.5PE, TR, Cy7APC
		FITC, PE, Cy5PE, Cy5.5PE or Cy7PE, TR, APC
	1 (488)	FITC, PE, Cy5PE, TRPE, Cy5.5PE, Cy7PE
7	2 (488+632)	FITC, PE, Cy5.5PE, Cy7PE, APC, Cy5.5APC, Cy7APC
		FITC, PE, Cy5PE, Cy5.5PE or Cy7PE, APC, Cy5.5APC, Cy7APC
	2 (488+595)	FITC, PE, Cy5.5PE, Cy7PE, TR, APC, Cy5.5APC or Cy7APC
		FITC, PE, Cy5.5PE or Cy7PE, TR, APC, Cy5.5APC, Cy7APC

[a] Any of the listed combinations can be augmented with a 407 nm Krypton laser and the fluorochromes CasB and CasY (or A430) to increase the number of colors by two without significantly affecting compensation requirements.

[b] Number of lasers, followed by excitation lines (see footnote a).

[c] Combinations are listed in rough order of increasing quality of measurements. i.e., for any given set of combinations, the first ones will tend to have lower compensation requirements and higher brightness. There are many more possible combinations than can be listed in this table. The order does not take into account commercial availability or ease of conjugate preparation. Substitutions: PerCP can be substituted for Cy5PE on lower laser-power instruments; Cy5.5PerCp can be substituted for Cy5.5PE on any system; A595 can be substituted for TR on any system. TRPE could be added to most combinations not employing 595 lasers or used as a substitute for Cy5PE. Where two dyes are listed as 'A or B', either can be used in the combination listed, but 'A' is preferred for reasons of brightness or compensation.

light from the primary fluorochrome for high sensitivity, but as little as possible from other fluorochromes to reduce the compensation required. In general, these two criteria work against each other; therefore, for any given set of fluorochromes used in an experiment, the optimal filter set may be different. However, we have found that the filter set shown in Fig. 1 and Table 3 is close to optimal for many experimental setups that are currently in use.

A further important selection criterion is the ability to block scattered light from the laser lines. Modern filters are considerably better than the old filters, allowing for example to collect the FITC signals with a much wider bandpass filter than was previously possible (old, 530/30; new, 525/50) while still blocking scattered light from the 488-nm laser line.

For experiments using a small subset of these fluorochromes, we found that using the widest possible bandpass that still excludes as many other fluorescences as possible yields the best results. Nonetheless, switching to wide-open filters will generally only increase the measured signal by 20–100% compared to the bandpass filters shown in

Table 2
Lasers for flow cytometry[a]

Laser type	Cooling	Line (nm)	Approx. cost	Fluorochromes
Argon Ion	Air	488	$10K	FITC, PE, TRPE, PerCP, Cy5PE, Cy5.5PE, Cy5.5PerCP, Cy7PE, A488
Krypton	Water	568	$30K	PE, TRPE, PerCP, Cy5PE, Cy5.5PE, Cy5.5PerCP, Cy7PE, A568, TR, A595
		647		APC, Cy5APC, Cy5.5APC, Cy7APC (Cy5PE, Cy5.5PE, Cy7PE)[e]
Violet-enhanced Krypton ion[b]	Water	407+413	$50K	CasB, CasY, A430
Dye[c]	Water	595	$15K[d]	TR, A595, APC, Cy5, Cy5.5APC, Cy7APC (TRPE, Cy5PE, Cy5.5PE)[e]
Doubled Nd-YAG	Air	532	$12K	PE, TRPE, PerCP, Cy5PE, Cy5.5PE, Cy5.5PerCP, Cy7PE
HeNe	Air	632	$6K (50 mW)	APC, Cy5, Cy5.5APC,
Diode	Air	635	$1K (10 mW)	Cy7APC (Cy5PE, Cy5.5PE)[e]

[a] Except as noted, lasers are rated at 50 mW output or greater at the given line.
[b] 405-nm diode lasers with sufficient power for cytometry will be available soon; currently, 4-mW lasers cost about $4K.
[c] Currently, 5-mW HeNe lasers with a 594-nm line can be obtained for about $3K; this power level is probably too low except for bench-top (flow cell) cytometers.
[d] A dye laser also requires a 5-W, 488-nm pump laser, adding another $20K to the cost of this laser. However, the 488-nm line of the pump laser (running in 'all lines' mode) can be split off and used as the primary laser beam in the instrument, obviating the need to purchase the primary Argon laser.
[e] Dyes in parentheses will also be excited by this laser, although their primary excitation is by other laser lines.

Table 2. For flow cytometric experiments, 2-fold increases in signal-to-background levels are marginal compared to the optimizations that can be achieved in other parts of the experiment.

4. Compensation

4.1. What is compensation?

Compensation is the process by which the spectral overlap between different fluorochromes is mathematically eliminated. The algorithm of compensation is a straightforward application of linear algebra, and should not be thought of as a subtraction process. Compensation between detectors can be performed either by hardware, after signal detection but before logarithmic conversion and/or digitization, or post-collection by software.

While compensation is one of the most important processes required for proper data analysis in flow cytometry, it is also perhaps the least well understood. In order to properly design, implement, and analyze multicolor experiments, users must be aware of the effects of compensation, understand how to apply it correctly, and recognize when data are not properly compensated. While we will discuss some of these topics here in brief, the reader is strongly encouraged to read the texts (Roederer, 1999) and web pages (Roederer, 1997b) devoted to this topic for more in-depth information.

4.2. Compensation complications

Currently, most instruments provide the capability

Table 3
Optimal interference filters for fluorochromes[a]

Fluorochrome	Filter[b]
CasB	440/40
CasY, A430	545/90
FITC	525/50
PE	575/25
Cy5PE	665/30
PerCP	680/30
Cy5.5PE, Cy5.5PerCP	720/45[c]
Cy7PE	785/45[d]
TRPE, TR, A595	625/40
APC	660/40
Cy5.5APC	705/50[c]
Cy7APC	750LP[d]

[a] Optimization is a function of maximizing the amount of signal collected while minimizing the amount of spillover from other fluorochromes. Different filters may be optimal for particular applications; see text.

[b] Filters are given as '*x*/*y*', where *x* is the wavelength in nanometers of the center of the bandpass and *y* is the width in nanometers of the band. LP, long pass filter. Filters have an extra coating to block scattered light.

[c] Use of a 720/45 filter for Cy5.5PE is optimal when Cy5PE is also to be measured. For measuring Cy5.5PE (or Cy5.5PerCP) in the absence of Cy5PE, then a 705/50 is better. Because APC has a lower emission wavelength than Cy5PE, the 705/50 works well for Cy5.5APC in combination with APC. However, if Cy5.5APC is to be measured with Cy5, then the 720/45 filter should be selected.

[d] The 750LP and the 785/50 are roughly equivalent for collecting Cy7 emission; either can be used.

for compensating between several spectrally adjacent pairs of detectors: for example, between FITC and PE, and between PE and Cy5PE. For a two-color analysis, such pairwise compensation is sufficient and complete. However, the introduction of a third fluorochrome introduces the potential for complex interactions that cannot by fully corrected by such pair-wise compensations. For example, FITC has low but measurable emission into the Cy5PE detector. This overlap may not be completely corrected by the combination of the FITC→PE and PE→Cy5PE compensation settings, causing what appears to be artefactual Cy5PE fluorescence when using a very bright FITC stain. Obviously, as more than three colors are used, these problems increase.

Indeed, for a typical 11-color stain, one would potentially need to use up to 100 pair-wise compensation controls to fully compensate all colors (in reality only about 40 of these are non-zero). Furthermore, because of the multiple interactions between these controls, some settings may actually be less than zero; something that current hardware controls do not allow. Finally, since setting interacting pairwise compensation controls is an iterative process, it would take virtually forever to manually set them properly.

An additional complication to compensation is brought about by the use of tandem dyes (e.g., Cy5PE, Cy7APC, etc.). Tandem dyes are covalently linked combinations of 'donor' (e.g., PE) and 'acceptor' (e.g., Cy5) dyes. The ratio of acceptor to donor is usually in the range of 3:1 to 10:1, depending on the optimization strategy used by the manufacturer. In fact, different tandem-conjugated antibodies from the same manufacturer often have slightly different ratios of acceptor to donor. The problem is that the compensation required to correct tandem emissions in other fluorescence channels strongly depends on this ratio. Therefore, a different compensation setting may be needed for every different tandem conjugate used in a particular experiment. In essence, this means that the compensation setting needs to be adjusted for every different staining combination in use.

Given these complications, it is apparent that complete compensation on multicolor systems requires computer intervention. Currently no instrument is commercially available that provides users with the ability to automatically adjust compensation settings on a per-stain basis, nor even to provide complete compensation among all pairs of fluorescence detectors. Therefore, the only solution as yet is the use of software analysis programs that are applied after data collection is completed for compensation of the data. Examples of such software include WinList (for PC) and FlowJo (for Macintosh).

The use of software to set compensation simplifies the process considerably. Computers can set the compensation correctly (there is no subjective criterion applied) and can be instructed to appropriately vary the compensation matrix according to the specific fluorochromes used in any given sample. However, there are significant limitations to software compensation (see below). In the interim before complete automated hardware compensation is avail-

able, the optimal solution is to combine both hardware and software compensation to minimize the limitations of each.

4.3. Limitations of compensation

Both available hardware and software compensation have significant limitations that place restrictions on the utility of PFC. Hopefully, some of these limitations will soon be obviated by new technology developments, which would enable the easier set-up of new reagent combinations for multicolor flow cytometric measurements. Until then, it is important to understand these limitations and how they may be best addressed.

One limitation that is shared by both hardware and software compensation is due to the low precision of measurements in the far-red channels, Cy7APC, Cy5.5APC, and to a certain extent APC. In these channels, for dimly stained cells, the actual number of photoelectrons detected by the PMT can be very low, due to the inefficiency of the currently available PMT to detect light at this wavelength. Even the 'deep red-sensitive' PMTs now available do not overcome this problem. The unavoidable counting errors inherent in the measurement can lead to the introduction of significant artefacts. For example, if only five photoelectrons are collected, the counting error is $\sqrt{5}$, or 45%. This error propagates to other channels that require compensation from this detector. The result is that the 'broadening' of populations that occurs due to compensation (Roederer, 1997b, 1999) is significantly worse for the far-red emitting fluorochromes.

Hardware compensation operates on the analog signals collected from the PMT, before further signal processing is done. For detectors with sufficient signal, this value is both precise and accurate, rendering compensation close to exact. Problems arise when compensation is required between signals propagated from different lasers, since the signals must be electronically delayed so that they can be presented to the analog compensation circuitry at the same time. This electronic delay is difficult to perform precisely, without changing the shape of the electronic pulse generated by the PMT. Therefore, hardware-based interlaser compensation is typically much less accurate than intra-laser compensation. Examples for significant inter-laser compensations are shown in Table 4. Potential signal degradation due to inter-laser compensation can occur with dye combinations that are found at high levels on the same cell.

The principal limitation of software compensation is caused by the inaccurate conversion from linear to logarithmic signals by the electronics of the flow cytometer. Since this conversion is typically performed by analog circuitry, it is merely an approximate conversion. The software, however, treats the data values as if they were an exact conversion, which can result in compensated values that are significantly inaccurate. The major inaccuracy in the electronic linear to log conversion is due to the fact that the dynamic range of the output is not exactly four decades, as the software (and the user) is told to expect, but can range from 3.5 to 4.3 decades. The effects of this error are insidious, in that data can be properly compensated at a particular intensity, under-

Table 4
Significant intralaser compensations[a]

Fluorochrome	Detector			
	TR	APC	Cy5.5APC	Cy7APC
First→second laser compensation				
PE	1.4	0.68	0.17	
Cy5PE	0.32	9.0	6.8	1.0
Cy5.5PE	0.11	1.3	13.3	2.8
Cy7PE				1.1
Second→first laser compensation[b]				
	Cy5PE	Cy5.5PE	Cy7PE	
TR	2.2	1.1		
APC	4.6	1.3	0.38	
Cy5.5APC	0.91	5.1	2.9	
Cy7APC	1.4	0.9	14	

[a] Significant spillovers are those above 0.1%. Values are given as the percentage of the signal level in the primary detector for a given fluorochrome that is found in the detector listed in each column. Only interlaser spillovers are shown; intralaser spillovers are significantly higher but can be more easily compensated by hardware. Values are given for a 488-nm first laser and a 595-nm second laser. Different laser lines can give qualitatively similar but quantitatively different results. Spillover values are system dependent and determined by optics, laser power, and gain of detector (PMT type, voltage, amplifier) on each channel pair.

[b] There are no significant spillovers into the FITC or PE channels from these dyes.

compensated at lower (or higher) intensities, and overcompensated at higher (or lower) intensities. The less compensation that needs to be done by the software the smaller the extent of this problem is. This again points to the importance of appropriate reagent/color choices, as described below.

Given the respective limitations of hardware and software compensation, our approach has been to use hardware compensation for as many pairs of signals as possible (Bigos et al., 1999). Currently we use 14 hardware controls, dedicated to signals that require the largest amount of intralaser compensation. Software compensation is used following data acquisition to complete the compensation process. It is critical to set the pair-wise settings conservatively, i.e., to under-compensate considerably where necessary. There are two reasons for this: inadvertent overcompensation cannot be corrected in software and thus causes irretrievable data error; and, given the multiple interactions, proper compensation between one pair of fluorochromes can lead to overcompensation of a different pair. For example, proper compensation of TR into APC often results in (unrecoverable) overcompensation of Cy5PE into APC.

5. Choosing the right fluorochrome

As outlined above, fluorochromes differ considerably in relative 'brightness'. Conjugation of the various fluorochromes to the same antibody and staining of the same population of cells can therefore result in large differences in resolution of positive and negative events (Fig. 2). For single-color staining for a highly expressed marker such as CD8 on T cells (Fig. 2), the differences are merely cosmetic and no alterations in the frequency of negative and positive events are found after staining with different conjugates. However, when staining for surface markers that exhibit high and low levels of expression, for example the expression of CD28 or FAS on T cells, a relatively 'dull' fluorochrome can result in an underestimation of the size of the positive population (false-negatives).

False-positive results are seen particularly with dyes and conjugates that have high 'background staining'. High background staining can occur when excessive concentrations of the conjugated antibody are used, or when conjugates are used that were prepared with a non-optimized ratio of fluorochrome to antibody. Titration of reagents and optimization of reagent conjugations are therefore particularly important when attempting to scale up a one- or two-color stain to a stain that contains three or more colors. Some reagents, such as Cy5PE, also exhibit a certain degree of 'stickiness'; in the case of Cy5PE this is seen particularly with B cells. Finally, the amount of cellular autofluorescence varies significantly in different areas of the spectrum. In particular, autofluorescence overlaps greatly with FITC, CasB, and CasY. Therefore, highly autofluorescent cells such as macrophages might appear as 'dull positive' in these detectors, despite the fact that they might not stain specifically with any of the reagents used.

The choice of fluorochrome/conjugate used for identifying a given marker will depend on the levels of expression and the cell types on which the marker is expressed. If this information is not at hand, it is advisable to use one of the 'bright' fluorochromes such as PE, Cy5PE or APC, and use the duller stains for identification of surface markers that are known, such as the lineage markers for lymphocyte subset staining. As an example, we show in Fig. 3 various dual-color staining combinations of splenic mouse cells for CD19, a lineage marker for B cells, and CD5. CD5 is expressed at high levels on all T cells and at low levels on a small subset of B cells in the spleen (Kantor and Herzenberg, 1993). $CD5^+$ B cells express somewhat higher levels of CD19. As apparent from the figure, resolution of the dully stained $CD5^+$ B cells is achieved only when the biotinylated CD5 antibody is revealed with streptavidin (SA) conjugated to PE, Cy5PE or APC (see also Kantor and Roederer, 1997). Staining with other SA-fluorochrome conjugates does not reveal a distinct $CD5^+$ B cell population. It is also useful to note that staining with SA-CasB resulted in an apparent staining of a small population of CD19-dull cells. However, this 'staining' is in fact due to the autofluorescence of macrophages which appears as a dull positive stain in many channels (data not shown).

Hence, simply the choice of fluorochrome used will determine whether this small B cell population

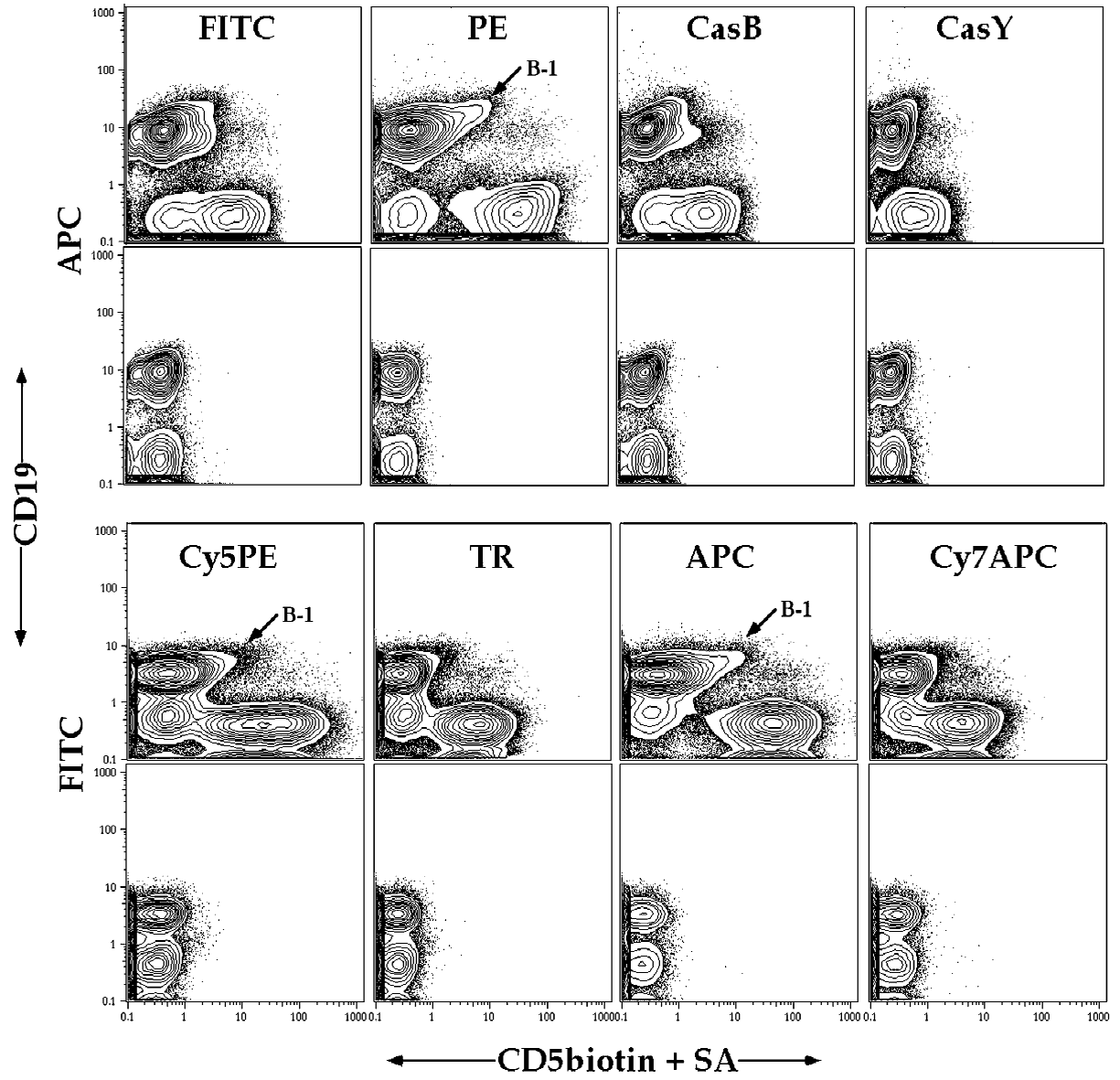


Fig. 3. Shown are 5% probability contour outlier plots of mouse spleen cells stained with a monoclonal antibody to the pan-B cell marker CD19, conjugated as indicated to either FITC or APC and with (upper panel) or without (lower panel) biotinylated anti-CD5 monoclonal antibody. CD5 recognizes all T cells and a small population of $CD19^+$ B cells. CD5 was revealed by adding streptavidin (SA) conjugated to the various fluorochromes indicated to the samples. Note that only staining with SA-PE, -Cy5PE or -APC resulted in a clear staining of the $CD19^+$ $CD5^+$ B cell population (indicated by the arrow).

can be identified at all. Importantly, each of the reagents used was able to clearly separate the $CD5^+$ $CD19^-$ T cell populations from the $CD5^-$ cells. The general rule of thumb is therefore: the lower the expression level of a given marker, the more likely that the choice of a bright fluorochrome will help to reveal its expression. Fig. 2 might help to gauge the relative brightness of fluorochromes.

6. Choosing the right combinations of fluorochromes

What are the criteria for choosing the right combination of fluorochromes? As outlined above, one major criterion is the expected levels of expression for the markers of interest. Markers such as CD5 on B cells can only be identified when a bright conjugate is used; but identifying CD5 on T cells can be accomplished with any fluorochrome. An unknown surface molecule is initially best studied with a conjugate that uses one of the 'brightest' fluorochromes, PE, Cy5PE, or APC. Other considerations include the spectral overlaps that exist between the various dyes in the staining combinations. Although hardware or software compensation can eliminate some of the associated problems, the careful choice of reagents can minimize the need for compensation.

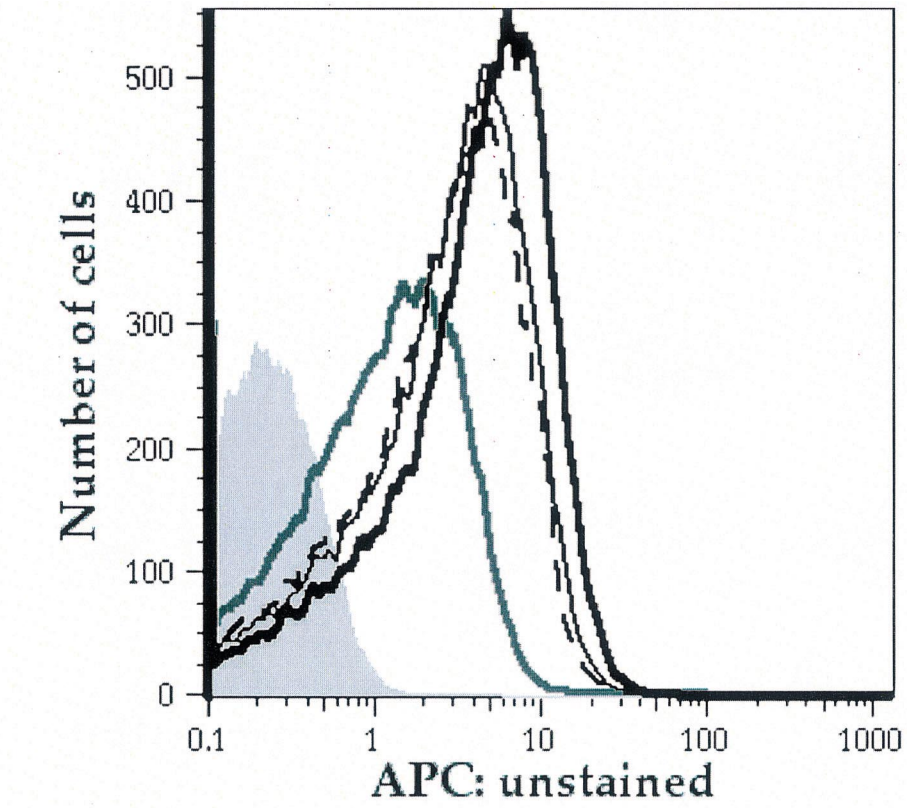
Fig. 4 shows an example of how the APC signal is affected by using one to four very brightly staining markers that all have significant spillover into the this channel. This figure therefore shows an example of how best *not* to design an experiment. In this example we chose to add antibodies to B cell markers conjugated to Cy5PE, TR, Cy5.5APC and Cy7APC to the staining mix and then looked at the background APC levels. Remember that APC is one of the dyes that are in general 'bright' (see Figs. 2 and 3) and is therefore a good choice for identifying dimly expressed markers such as CD5 on B cells. As we add staining reagents to the cells, the background levels of the APC signal increase significantly, from a mean fluorescence intensity of 0.2 to 4.1 (Fig. 4). This increase in mean fluorescence intensity at the APC detector is observed despite the fact that software compensation was applied to the samples and each compensation pair seemed to have been compensated appropriately (not shown; see Fig. 5 for explanation). Therefore, after addition of the Cy5PE, TR, Cy5.5APC and Cy7APC stains, a dim positive staining with a reagent conjugated to APC could no longer be distinguished from the increased background staining; the effective 'brightness' of APC is reduced 20-fold. Since a number of colors that are currently used for PFC show large overlaps into the APC detector, the usefulness of this detector is substantially compromised unless reagents are chosen for the other conjugates that do not stain the cells identified by the APC stain.

The other crucial lesson from this experiment is that isotype controls are only appropriate when used in the context of every other antibody stain. That is, a sample that was stained with isotype controls on all channels would show a background of only 0.2 on the APC channel, considerably less than the background observed on this channel given a full staining set minus the APC stain (4.1). This is why false positives will be identified without the use of appropriate controls.

In the example shown in Fig. 3, i.e., measuring the low levels of CD5 on $CD19^+$ cells, one should choose to label CD19 (well-expressed on all B cells) with a color that does not have a large compensation component with the fluorochrome used for CD5. On the other hand, CD8 (part of the 'dump' channel) could be assigned to such a color even though CD8 is extremely highly expressed. This is because CD8 is not expressed on the cells of interest ($CD19^+$ cells) and thus compensating CD8 fluorescence would have no impact on the quantitation of CD5 on B cells.

The dyes excited by the violet-enhanced krypton laser are particularly useful for multicolor staining combinations. Although they appear rather 'dull' mainly due to the large autofluorescence background in these channels (Fig. 2), they have minimal overlap with any of the other fluorochromes and are easily compensated against each other. Therefore any combination with a 407-nm line-excited dye should work without problems, as long as the marker chosen for the CasB or CasY conjugation is expressed at sufficiently high levels to separate the stained from the unstained cell population. In many cases, significant compensation can also be avoided by using dyes excited by different lasers (however, see Table 4 for dyes that *do* require significant interlaser compensation).

In summary, the appropriate choice of reagents and reagent combinations requires certain knowledge of the characteristics of the individual fluorochromes, such as 'brightness' and 'spillover' into other channels, to avoid the use of inappropriate or useless staining combinations. Typically, multiple different combinations of fluorochrome/antibody pairings



	Cy5-PE	TR	Cy5.5-APC	Cy7-APC	MFI (APC)
	-	-	-	-	0.2
	+	-	-	-	0.8
	+	+	-	-	2.5
	+	+	+	-	2.9
	+	+	+	+	4.1

Fig. 4. Shown in solid gray is the histogram of background APC signals (i.e., no APC reagent was added) of splenic mouse cells stained with FITC-anti-Ig-κ and gated to identify all B cells (data not shown). Stepwise addition of other reagents (identified by '+' in the table) that have considerable spectral spillover into the APC channel increased the background level of the APC signal. This is shown in the shift of the histogram plots to the right and the increase in mean fluorescence intensity level (*but not median*!) indicated in the table. This increase in the mean was observed despite appropriate software compensation and is a direct result of the broadening 'artefact' of compensation (Roederer, 1997a,b, 1999); see also Fig. 5. Reagents used: Cy5PE-anti-B220, TR-anti-IgD, Cy5.5APC-anti-CD19, Cy7APC-anti-IgM.

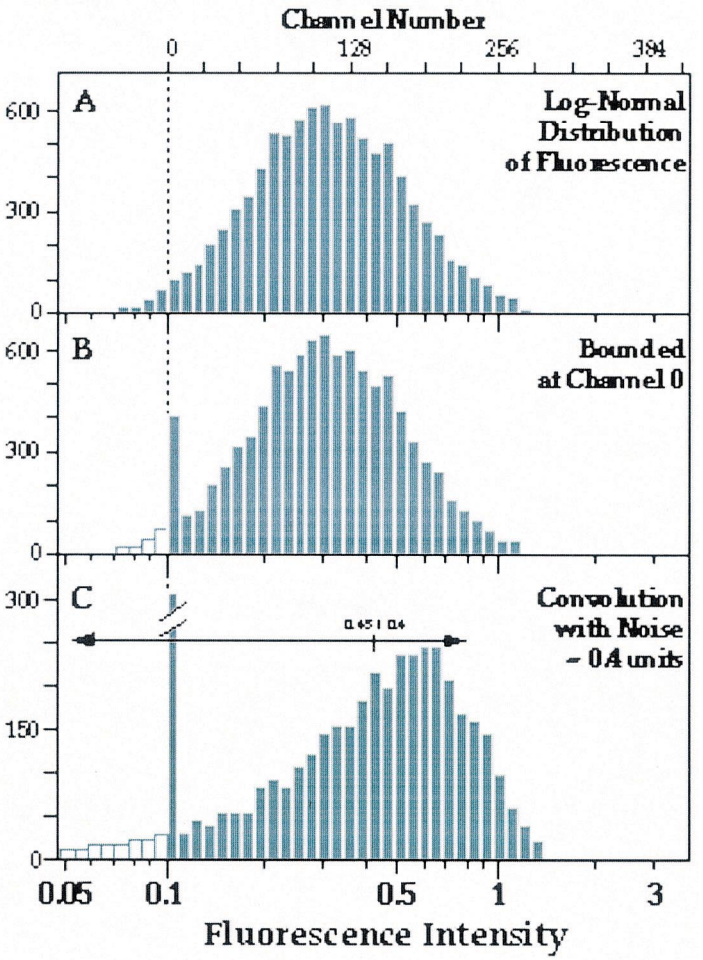


Fig. 5. The effect of random noise added to an autofluorescence distribution. The process of fluorescence compensation introduces additional error into the distribution of measured fluorescence. Histograms in this figure represent a mathematical modeling of this process. (A) A typical distribution of cellular autofluorescence. The fluorescence distribution is log-normal, i.e., it is a Gaussian when viewed on a logarithmic axis. The distribution is shown only for the first 400 (of 1024) channels of a typical measurement parameter; this represents the first 1.5 decades of fluorescence. (B) The distribution in (A), after instrumentation boundary conditions are imposed. Since there is no 'zero' on a logarithmic axis, the low end of the fluorescence scale is arbitrary. Any events with fluorescence below the lowest value (0.1 here) are automatically assigned to channel zero and given a fluorescence value of 0.1. This results in an artificial peak at channel zero, the sum of all events with fluorescence less than this (shown in white bars). This same artefact occurs in two-dimensional histograms used for contour plots or density plots, causing the appearance of an artificial 'peak' (multiple contour lines) on the axis. This peak is an artefact of the measurement and visualization processes, and importantly, it does not represent a population of events distinct from those adjacent to it. (C) Effects of compensation. To model the error introduced by compensation, a random amount of fluorescence (averaging 0 ± 0.4 fluorescence units) was added to each event in the distribution shown in (B). The fluorescence of about 25% of all events is now shifted to below 0.1 units (open bins), consequently these events are forced into the single point at channel zero. Note the scale change of the vertical axis. The reason for the asymmetric effect on the distribution (i.e., the right edge of the histogram increased very little, but the left tail increased greatly, and the mode of the distribution has increased) is a consequence of the logarithmic scaling: a noise distribution that is linearly distributed was added to the log-normal distribution. Thus, as shown by the arrows, adding 0.4 units of fluorescence to an event at 0.45 (to 0.85) moves it much less 'distance' (in the log domain) than subtracting the equivalent amount (to 0.05). Compensation introduces errors that are symmetrically distributed in the linear domain. Thus, one of the effects of compensation is to alter the autofluorescence distribution from log-normal to that shown in (C). It is important to note that the *median* fluorescence intensity in (C) is exactly the same as it is in (B), but the *mean* fluorescence has increased, and the *mode* (above channel 0) has increased even more. Because of our internal bias to use the mode (peak) to estimate a population's center rather than the median or mean, which is virtually impossible to estimate visually, the distribution appears to have moved substantially. Finally, note that placement of a 'positive' gate at, for example, 1 fluorescence unit would be appropriate for the 'uncompensated' parameter (B), but not for the compensated parameter (C), even though both represent unstained distributions.

must be evaluated in order to select the panel that provides the most information.

7. Enhanced accuracy of measurement with multicolor flow cytometry

Despite the relatively good resolution that the use of SA-PE, Cy5PE and APC afforded for the staining of CD5 on a small B cell subpopulation in the spleen (Fig. 3), it would be difficult to accurately determine the frequency of these cells using this two-color stain. In Fig. 6 we outline how the simultaneous use of four (Fig. 6B) and six (Fig. 6C) colors enhances the accuracy of the measurement and the amount of information that can be obtained from one stain (Fig. 6A). In this figure we compare the staining pattern of spleen cells isolated from wild-type mice and gene

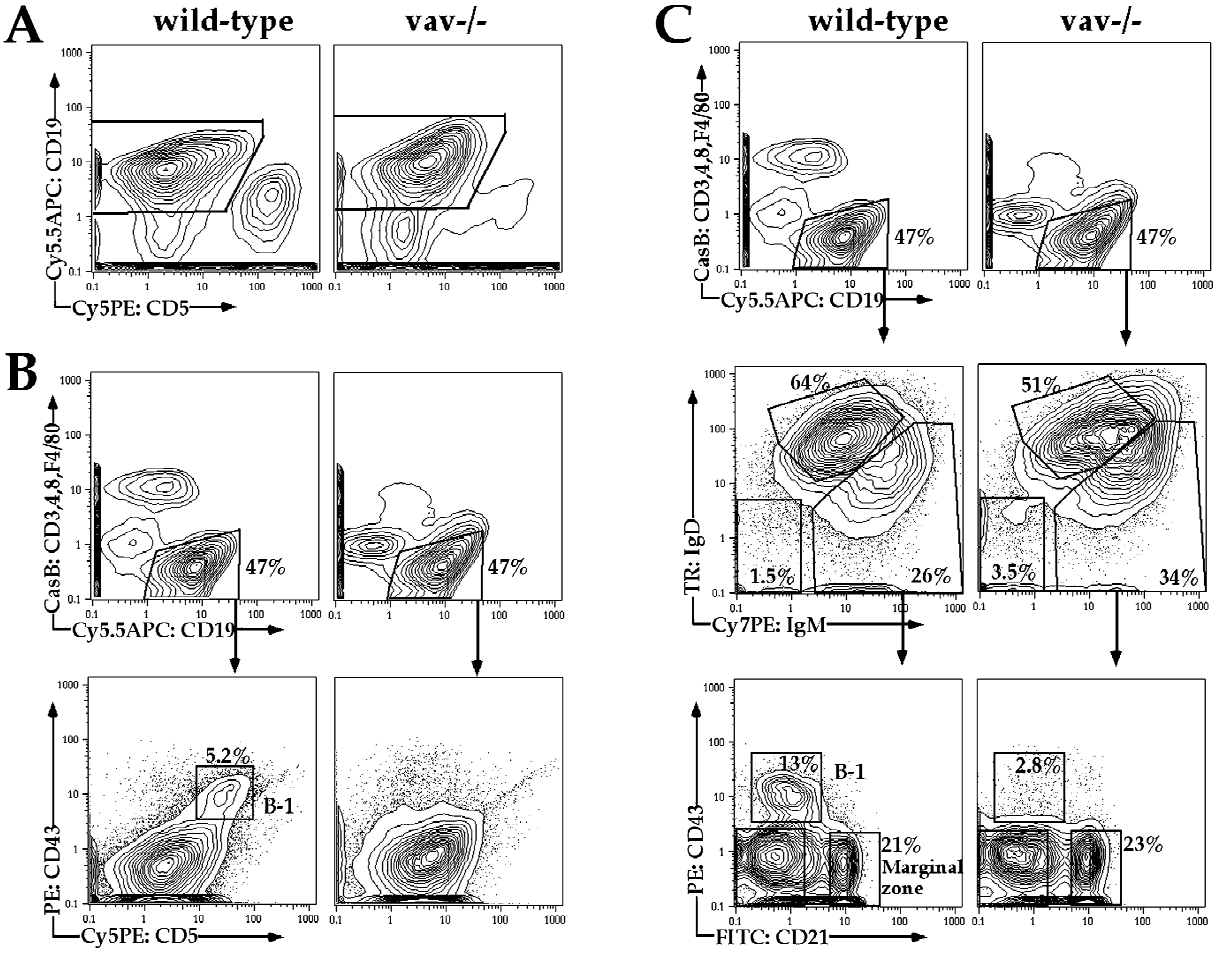


Fig. 6. Shown are 5% probability contour outlier plots of (A) two-, (B) four- and (C) six-color stains of single cell suspensions prepared from mouse spleens derived from normal wild-type C57BL/6 mice and mice deficient in the signaling molecule vav (vav$-/-$). (A) CD19 and CD5 staining of splenic B cells shows a reduction in the number of $CD5^+$ $CD19^-$ T cells and $CD19^+$ $CD5^+$ B cells. (B) Addition of a dump channel (reagents to CD3, CD4, CD8 and F4/80) and a second marker (CD43) expressed on mouse 'B-1' cells identifies the lack of B-1 cells among the overall normal frequency of $CD19^+$ B cells in vav$-/-$ mice compared to their wild-type controls. (C) Addition of anti-IgM, anti-IgD and anti-CD21 in the staining cocktail reveals alterations in IgM and IgD expression of follicular B cells and the appearance of normal frequencies of marginal zone ($CD21^{hi}$ $CD43^{neg}$) and immature ($CD21^{lo}$ $CD43^{neg}$) B cells and the absence of B-1 cells ($CD21^{lo}$ $CD43^+$) among IgM^{hi} IgD^{lo} cells.

targeted mice, deficient in vav, a signaling molecule important in appropriate signaling through the antigen-receptor expressed by the B cell. The presence of vav has been shown to be crucial for normal B cell development (Zhang et al., 1995). Vav−/− mice have splenic B cells that differ considerably in phenotype (Fig. 6) and function from their wild-type controls. Comparison of a two-color stain of CD19 and CD5 expression using Cy5PE for staining of CD5 (Fig. 6A) identifies a subpopulation of $CD5^+$ (>14 units) $CD19^+$ (>0.9 units) cells in the spleens of the wild-type mice. In the gene-targeted mice this population seems absent. However, it is not clear from this stain whether the levels of CD5 expression are reduced, or whether this population of B cells is missing in these mice. In fact, a somewhat more inexperienced investigator might not even see a difference in the CD5-staining patterns of the $CD19^+$ cells for these two different mice. However, a clear difference in the frequency of the $CD5^+$ $CD19^-$ T cell population is easily demonstrated.

Much better quantitation of the $CD5^+$ $CD19^+$ B cell population is achieved with a four-color stain, as shown in Fig. 6B. This four-color stain includes a so-called 'dump channel', a combination of reagents conjugated to the same fluorochrome that are expressed on cells that are *not* of interest. The use of a dump channel is a particularly good way for enhancing the fidelity of a given stain, since not only stained cells, but also 'sticky' and highly autofluorescent cells can be clearly separated from the cells of interest. Another reason for including a dump channel in a staining combination is the identification of small cell subsets, for example to identify the lineage marker negative stem cells found in bone marrow or to identify dendritic cells in peripheral lymphoid tissues. In Fig. 6B, we show how B cells are identified first by their expression of CD19 and the lack of expression of various T cell markers (CD3, CD4 and CD8) and a macrophage marker (F4/80). Simultaneous staining for CD5 and CD43 now allows identification and accurate enumeration of a population of $CD5^+$ and $CD43^+$ B cells that express $CD19^+$, since CD43 and CD5 co-expression is a hallmark of ($CD5^+$) 'B-1 cells' in the mouse (Wells et al., 1994). Note that these markers can only be used in combination with CD19, as both are also expressed on virtually all T cells. This four-color stain therefore clearly reveals that vav−/− mice completely lack B-1 cells in the spleen, a conclusion that could not have been drawn from any two-color analysis.

The use of six or seven colors (Fig. 6C) on the same cell samples shown in Fig. 6B reveals a much more detailed analysis of the effects of vav-deletion on the other peripheral B cell subpopulations found in the spleen of mice. As shown in Fig. 6C, the relative frequency of $CD19^+$ cells is identical in the two mouse strains. However, staining of $CD19^+$ B cells for expression of surface IgM and IgD revealed strong differences in the expression pattern of these molecules on cells from vav−/− and control mice, respectively. Normal splenic mouse B cells contain a main population of cells expressing high levels of IgD and low levels of IgM, which are the relatively homogeneous, recirculating, follicular B cells. As seen in Fig. 6C, although the frequency of the cells does not differ greatly in the vav−/− mice, the levels of IgM expression are increased, indicating a maturation defect in the cells from these mice. Furthermore, the population of surface IgM- and IgD-negative cells, presumably cells that have undergone isotype switching, is increased 2–3-fold (1.5–3.5%). Finally, the population of IgM^{hi} IgD^{lo} cells contains at least three cell populations, as shown in the lower panel of Fig. 6C. These can be differentiated by their expression of CD21 and CD43 into immature B cells ($CD21^{neg}$ $CD43^{neg}$), marginal zone B cells ($CD21^{hi}$ $CD43^{neg}$) (Takahashi et al., 1997) and B-1 cells ($CD21^{lo}$ $CD43^+$) (Wells et al., 1994). As shown in the four-color stain, B-1 cells are completely absent in the vav−/− mice, whereas the two other populations, marginal zone and immature B, seem unaffected in these mice. With the help of this one six-color stain we can therefore identify all major known B cell populations present in the spleen of mice, allowing us to clearly identify and quantitate the effects of vav or any other gene of interest on these cells.

This example illustrates how adding more and more measurements on the same cell populations can significantly *simplify* the interpretation of the data. Rather than trying to draw conclusions from a series of two- (or even four)-color measurements that can only show changes in bulk populations of B cells, or from changes in relative expression levels of certain

markers on any of several indistinguishable cell types, the use of PFC allows us to pinpoint precisely the defects in B cell development of these mice.

8. Staining controls

Control stains are important for all experiments using flow cytometry, but they become particularly critical for complex multicolor stains. The higher the number of fluorochromes and antibodies used in each stain the greater the risk for artifacts introduced by compensation errors and/or reagent interactions. In general, two types of controls should be included and data collected with every experiment: compensation controls and staining controls.

It will become apparent from the discussion below that the number of control tubes might often exceed the number of sample tubes in a given experiment. It is necessary to invest this significant effort in the controls in order to ensure that the valuable sample that is stained with so many different markers does not suffer from experimental artefacts, or become uninterpretable.

8.1. Compensation controls

For each fluorochrome used in an experiment, one should include a 'compensation control', i.e., a single color stain, for which data are collected. It is *not* appropriate to skimp on the number of tubes collected by trying to design stains that can control for multiple compensation settings in a single tube. Ideally, the reagent used for the compensation sample should be the same as that used in the staining cocktail. For the non-tandem dyes (FITC, PE, TR, APC), however, a reagent that stains brightly on a good number of cells is sufficient and can be used instead of a reagent that might only stain a very small subpopulation of cells. In general the best result is obtained by using the brightest possible reagent as the compensation control. In contrast, and as outlined above, the tandem-dyes (Cy5PE, Cy7PE, Cy7APC, etc.) may exhibit great lot-to-lot variations. As lot-numbers of the fluorochromes are not usually supplied with the conjugate if bought from a commercial source, each different staining combination should have its own compensation control. For tandem dyes that are prepared in the laboratory, stains that use tandem dyes from the same lot should suffice. Thus, careful inventory of conjugates in a laboratory that includes the lot number of a tandem dye for each antibody conjugate can greatly reduce the number of compensation control samples for each experiment.

8.2. Isotype or unstained controls

The use of isotype control antibodies is usually of little value, since each antibody and antibody conjugate has very different characteristics in terms of background staining, 'stickiness', etc. If nonspecific staining is suspected, the use of an 'Fc-block', such as anti-Fc receptor (CD16/CD35) antibodies (for staining mouse cells) might be useful. These antibodies are usually applied for 10 or 15 min before the stains are added.

Since completely unstained cells have very different levels of background than cells that have been stained with multiple reagents (see Fig. 4 and its discussion), the best control for any given marker of interest in a multicolor staining combination is a stain that contains all reagents but the one of interest. For example, a five-color combination control would include one or more four-color combinations, each leaving out one of the reagents in the complete stain. In this way, positive versus negative gating can be more rigorously determined and applied. In some cases, this gate may not be a one-dimensional (histogram gate), but may require a sloping line in two dimensions because of the ever-increasing broadening imposed by compensation as signal intensity increases.

9. Advantages of PFC

Most of the advantages of increasing the number of colors in flow cytometry experiments are readily apparent. First, the information content increases geometrically with the number of parameters simultaneously analyzed. Second, and just as important, information can be obtained from multicolor experiments that is not available in any other way. For example, no combination of one-color stains can accurately enumerate or be used to isolate

$CD3^+CD4^+CD8-$ T cells (excluding, for example, $CD3^+CD4^+CD8^+$ T cells and small $CD4^+$ monocytes). Third, detailed immunophenotyping can be combined with functional assays (DNA/Cell cycle, metabolic activity, cytokine assays, apoptosis assays, etc.) to provide a distribution of activities across highly defined cell subsets. Fourth, by staining for multiple markers in one cocktail, far fewer samples need to be prepared for each experiment. This is particularly important when analyzing precious samples (e.g., pediatric samples, leukocytes isolated from biopsies, rare antigen-specific lymphocytes, mouse tissues that yield small numbers of cells). And fifth, less total reagents are used, since less duplication of the same reagent among multiple tubes is required.

A number of technical developments have also allowed the increasing use of flow cytometry for extensive functional studies on cells. Increasing the number of simultaneously measured parameters is particularly useful for these applications, since one can combine detailed immunophenotyping with functional studies. One example is the use of cytoplasmic cytokine staining for the identification of cytokine profiles among bulk populations of lymphocytes. By combining antibodies against a number of cell surface markers together with antibodies specific for cytokines, we recently demonstrated that populations of human peripheral blood T cells classified by their differential expression of various 'activation markers' differ in their relative frequencies of IL-4 and IFN-γ producers. Importantly, the relative frequencies of these phenotypically and functionally distinct T cell subpopulations in the blood differed among people with various disease states (Mitra et al., 1999). In the future, the use of additional antibodies against chemokine- and homing-receptors will likely lead to the identification of further T cell subsets. When combined with antibodies against a number of cytokines, this is likely to yield further important and novel information regarding the relationship between T cell phenotype, function and the migration pattern of these cells. Eventually, this might lead to the identification of correlates of disease pathogenesis that will provide important information on which to base therapeutic interventions.

The increased quality of the information obtained with PFC compared with two- or three-color flow cytometry can result in the unambiguous identification of rare cell subsets. For example, using eight-color PFC we were able to identify a small population of follicular B cells in the mouse, with an as yet unknown function, that expresses high levels of the nonclassical MHC molecule CD1 (Amano et al., 1998).

Other examples include the use of 10-color PFC for the identification of antigen-specific lymphocytes. Multimers of soluble MHC-class I molecules loaded with a peptide of known origin are now often used to identify antigen-specific $CD8^+$ T cells in human blood or various tissues in the mouse (Altman et al., 1996). As the number of antigen-specific T cells is often very small, even a small amount of contamination can lead to considerable errors in the quantitation of these cells which makes further analysis inaccurate. We recently showed that the use of PFC allows the identification and functional characterization of γδ T cells specific for a nonclassical MHC molecule T22. γδ T cells constitute roughly 1% of the splenic cell population of which about 0.5%, or 0.005% of all spleen cells, bind to the T22 tetramer (Crowley et al., 1999). Using multicolor cytometry we could show not only that those cells bind to the MHC multimer, but also that they alter their surface expression of activation markers such as CD69 and CD62L after binding the MHC multimer (Crowley et al., 1999).

In a separate study, we also used the MHC multimer approach to identify potentially tumor-reactive cytotoxic T cells in patients with metastatic melanoma (Lee et al., 1999). Because of the rarity of these cells (10^{-5} to 10^{-3} of PBMC), standard techniques are unable to derive much information about these cells. We were able to quantitate the expression of nearly 40 different cell surface antigens (to accurately pinpoint the T cell subset to which the antigen-specific T cells belong), as well as determine their cytokine profile, showing that these cells are functionally anergic (whereas viral antigen-specific T cells from the same patient samples were normally functional). We were even able to perform serological T cell receptor repertoire measurement to demonstrate that, in one patient, these cells (constituting 2% of all CD8 T cells) expressed only a single Vβ gene.

Although in routine use for only 2 years, 11-color

PFC has proven to be a most valuable tool for the study of different aspects of cellular immunology. These successes are almost certainly a harbinger of the new age of cytometry that we will enter as instrumentation to support true multicolor flow cytometric measurements become commercially available. This new age will encompass considerably more complex experiments, data analysis, and hurdles. In the end, however, these complex analyses will yield a more complete understanding of biology, especially immunobiology. Ultimately, PFC is a tool that will *simplify* complex biological processes so that we may better comprehend their interactions, their functions, and their role in biology.

We hope that with this brief outline on the advantages and pitfalls of multicolor flow cytometry we can encourage and provide some guidance to an increasing number of users for whom flow cytometry presents a new technology that allows them to address previously unanswerable questions.

Acknowledgements

The development of multicolor flow cytometry spanned more than half a decade and represents the efforts of a large number of individuals. We thank in particular Drs. Leonard and Leonore Herzenberg in whose laboratory both fluorescence-activated flow cytometry and Polychromatic Flow Cytometry were born. The work, continuous drive and implementation of technical improvements by Dr. David Parks, Marty Bigos, Richard Stovel, Tom Nozaki, Wayne Moore, and Adam Treister have led to the development of both the hardware and software tools necessary to make this technology work. We thank Drs. Alan Waggoner, Aaron Kantor, Rachel Gerstein, Michael Anderson, and Richard Haugland for their collaborative efforts in the development of many of the new fluorochromes we employ, as well as Drs. Randy Hardy and Alan Stall, who developed the conjugation protocols that we routinely use. We also like to acknowledge Drs. Stephen De Rosa, Dipendra Mitra, and Nobukazu Watanabe, who spearheaded many of the early research projects that helped focus the development efforts, and Gina Jaeger, Ometa Herman, and Iwan Tjioe for expert technical assistance. We are especially grateful to Dr. Ernie Huang and Susan Wormsley, to PharMingen and to Becton Dickinson for significantly supporting our reagent development program; and to both Becton Dickinson and Cytomation for instrumentation support. The vav−/− mice were provided by the Crabtree laboratory at Stanford. Finally, we wish to thank the Herzenberg laboratory and the flow cytometry community at large for their support, enthusiasm, and valuable ideas.

References

Altman, J.D., Moss, P.A.H., Goulder, P.J.R., Barouch, D.H., McHeyzer-Williams, M.G., Bell, J.I., McMichael, A.J., Davis, M.M., 1996. Phenotypic analysis of antigen-specific T lymphocytes. Science 274, 94–96.

Amano, M., Baumgarth, N., Dck, M.D., Brossay, L., Kronenberg, M., Herzenberg, L.A., Strober, S., 1998. CD1 expression defines subsets of follicular and marginal zone B cells in the spleen: b2-microglobulin-dependent and independent forms. J. Immunol. 161, 1710–1717.

Bigos, M., Baumgarth, N., Jager, G.C., Herman, O.C., Nozaki, T., Stovel, R.T., Parks, D.R., Herzenberg, L.A., 1999. Nine color eleven parameter immunophenotyping using three laser flow cytometry. Cytometry 36, 36–45.

Crowley, M.P., Fahrer, A.M., Baumgarth, N., Hampl, J., Gutgemann, I., Teyton, L., Chien, L.-h., 1999. A population of murine gd T cells that recognize a nonclassical MHC class I molecule. Science 287, 314.

Kantor, A., Roederer, M., 1997. FACS analysis of leukocytes. In: Herzenberg, L.A., Weir, D.M., Herzenberg, L.A., Blackwell, C. (Eds.), Handbook of Experimental Immunology, 5th Edition. Blackwell Science, Cambridge, pp. 49.1–49.13.

Kantor, A.B., Herzenberg, L.A., 1993. Origin of murine B cell lineages. Annu. Rev. Immunol. 11, 501–538.

Lee, P.P., Yee, C., Savage, P.A., Fong, L., Brockstedt, D., Weber, J.S., Johnson, D., Swetter, S., Thompson, J., Greenberg, P.D., Roederer, M., Davis, M.M., 1999. Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. Nat. Med. 5, 677–685.

Mitra, D.K., Rosa, S.C.D., Luk, A., Balamurugan, A., Khaitan, B.K., Tung, J., Mehra, N.K., Terr, A.I., O'Garra, A., Herzenberg, L.A., Herzenberg, L.A., Roederer, M., 1999. Differential representations of memory T cell subsets are characteristic of polarized immunity in Leprosy and atopic diseases. Int. Immunol. 11, 1801–1810.

Roederer, M., 1997a. Methods for fluorescent conjugation of monoclonal antibodies. http://www.drmr.com/abcon.

Roederer, M., 1997b. Compensation in flow cytometry: a perspective. http://www.drmr.com/compensation.

Roederer, M., 1999. Compensation. In: Robinson, J.P., Darzynkiewicz, Z., Dean, P.N., Dressler, L.G., Rabinovitch, P.S., Stewart, C.C., Tanke, H.J., Wheeless, L.L. (Eds.), Current Protocols in Cytometry. Wiley, New York, in press.

Roederer, M., DeRosa, S., Gerstein, R., Anderson, M., Bigos, M., Stovel, R., Nozaki, T., Parks, D., Herzenberg, L., Herzenberg, L., 1997. 8-Color, 10-parameter flow cytometry to elucidate complex leukocyte heterogeneity. Cytometry 29, 1–12.

Shapiro, H.M., 1994. In: Practical Flow Cytometry. Wiley-Liss, New York.

Takahashi, K., Kozono, Y., Waldschmidt, T.J., Berthiaume, D., Quigg, R.J., Baron, A., Holers, V.M., 1997. Mouse complement receptors type 1 (CR1;CD35) and type 2 (CR2;CD21). Expression on normal B cell subopulations and decreased levels during the development of autoimmunity in MRL/lpr mice. J. Immunol. 159, 1557–1569.

Wells, S.M., Kantor, A.B., Stall, A.M., 1994. CD43 (S7) expression identifies peripheral B cell subsets. J. Immunol. 153, 5503–5515.

Zhang, R., Alt, F.W., Davidson, L., Orkin, S.H., Swat, W., 1995. Defective signaling through the T- and B-cell antigen receptors in lymphoid cells lacking the vav proto-oncogene. Nature 374, 470–473.