0026-895X/04/6504-813–821$20.00
MOLECULAR PHARMACOLOGY
Vol. 65, No. 4
2998/1136429
Mol Pharmacol 65:813–821, 2004
Printed in U.S.A.

MINIREVIEW

Computational Lipidomics: A Multiplexed Analysis of Dynamic Changes in Membrane Lipid Composition during Signal Transduction

Jeffrey S. Forrester, Stephen B. Milne, Pavlina T. Ivanova, and H. Alex Brown

Department of Pharmacology and the Vanderbilt Institute of Chemical Biology, Vanderbilt University School of Medicine, Nashville, Tennessee

Received September 25, 2003; accepted October 30, 2003

This article is available online at http://molpharm.aspetjournals.org

ABSTRACT

Recent successes in defining the roles of lipids in cell signaling have stimulated greater interest in these versatile biomolecules. Until recently, analysis of these molecules at the species level has required labor-intensive techniques. The development of electrospray ionization mass spectrometry (ESI-MS) has made possible the detection and identification of thermally labile biological molecules, such as phospholipids. The "soft" ionization does not cause extensive fragmentation, is highly sensitive, accurate, and reproducible. Thus, this method is well suited for analyzing a broad range of phospholipids without elaborate chromatographic separation. Evaluating the vast amounts of data resulting from these measurements is a rate-limiting step in the assessment of phospholipid composition, requiring the development and application of computational algorithms for mass spectrometry data. Here we describe computational lipidomics, a novel analytical technique, coupling mass spectrometry with statistical algorithms to facilitate the comprehensive analysis of hundreds of lipid species from cellular extracts. As a result, lipid arrays are generated to indicate qualitative changes that occur in lipid composition between experimental or disease states, similar to proteomic and genomic analyses. This review presents a methodological strategy for using ESI-MS combined with a high-power computational analysis to profile time-dependent changes in cellular phospholipids after the addition of an agonist or to evaluate changes promoted by pathophysiological processes. As an illustration, we describe the methods and approaches used to generate lipid arrays for The Alliance for Cellular Signaling (AfCS). These arrays are contributing to a more complete understanding of the participants of cellular signaling pathways after activation of cell surface receptors.

From the early 1960s through the 1980s, the pioneering research of Eugene Kennedy defined the major pathways of phospholipid synthesis. Yet an appreciation of the involvement of phospholipids in cellular processes has lagged well behind that of nucleic acids and proteins for much of the biological community. Long viewed as simply the building blocks of the semipermeable membrane whose primary functions were compartmentalization and ion content regulation, one of the first hints of the central importance of lipids in cellular signaling came when Hokin and Hokin (1953) showed that treating avian pancreatic slices with cholinergic drugs resulted in both the secretion of amylases and the concomitant turnover of membrane phospholipids. The significance of this observation went initially unrecognized because of a lack of appreciation of the diverse cellular roles of phospholipids and the inherent difficulties in their measurement. Elucidation of the inositol-mediated calcium gating opened up a new era in the study of these important biomolecules (reviewed in Berridge, 1993). It is now understood that lipids play important roles as second messengers in signal transduction processes as well as participating in membrane topology by creating specialized domains of specific lipid and protein complexes in both normal and diseased cells (Ramanadham et al., 1998; Bev-

This work was supported by the National Institutes of Health grant GM58516 and The Alliance for Cellular Signaling.

ABBREVIATIONS: ESI, electrospray ionization; MS, mass spectrometry; CID, collision induced dissociation; PC, phosphatidylcholine; PE, phosphatidylethanolamine; SM, sphingomyelin; PS, phosphatidylserine; PI, phosphatidylinositol; PA, phosphatidic acid; LPC, lysophosphatidylcholine; MS/MS, tandem mass spectrometry; BCR, B-cell receptor; AIG, Anti-IgM.

ers et al., 1999; Ivanova et al., 2001; Shen et al., 2001; Alb et al., 2003; Han and Gross, 2003).

The burgeoning interest in the role of phospholipids in cellular functions has spurred research in methodology for their identification and measurement. Until recently, lipid analysis has not benefited from paradigm shifts produced from new technologies with the same kind of impact as monoclonal antibodies for protein isolation and the various cDNA microarrays technologies available in molecular genetics. The ability of the researcher to delineate the role played by phospholipids in biological processes would be greatly enhanced if the technology to monitor dynamic changes in the cellular lipid composition during these events were available for a reasonable expenditure of effort and cost. The increasing sophistication of biological modeling schemata is intimately linked to the ability to acquire large, relevant sets of measurements from cellular phenomena. Moreover, the greatest promise of such schemata lies in their ability to form integrative models from readouts that previously seemed like disparate collections of data, making quantum leaps in the understanding of biological systems possible. These considerations make the ability to comprehensively measure changes in phospholipids a critical addition to the armamentarium of cellular biologists, complementing existing technologies such as gene arrays and proteomics.

Mass spectrometry has emerged as a forerunner in developmental technologies with the ability to analyze comprehensive changes in membrane lipid composition. Advances in ionization techniques have made routine the preparation and analysis of biological extracts with this highly versatile class of instruments. Lipid chemists have used mass spectrometry as a potent structural proof technique for many years. In this review, we describe how we built upon previous advances to develop methods for generating a qualitative map of cellular lipid changes at the species level from mass spectral data. The results of the analysis are displayed as a lipid array summarizing statistically significant changes in cellular lipids across an experimental time course after challenge with a biological agonist. Complete descriptions of the computationally intensive portion of the analysis are given, as well as an example of the technique from data collected in conjunction with our colleagues in The Alliance for Cellular Signaling (AfCS).

Importance of Phospholipids in Cell Function. Glycerophospholipids are the basic building blocks of cellular membranes, and their chemical structure and diversity are well suited for this important physiological role. These molecules consist of a glycerol backbone having a polar phosphate-head group attached to the third carbon, and acyl, alkyl, or alkenyl moieties attached at the *sn-1* and *sn-2* positions. The different head groups determine the glycerophospholipid classes, and the variety of fatty acyl chains contribute to the diversity of species within a class (Fig. 1). Although all glycerophospholipids contain a glycerol backbone, the diversity of head groups, acyl chains, and degree of unsaturation can produce hundreds of different lipid species existing within a given cell. This enormous number of structural combinations allows for a large variety of physical and chemical membrane properties, including membrane permeability, fluidity, and curvature. Local variations in the concentrations of particular lipid species mediate these properties through the formation of lipid rafts and the regulation of lipid-protein interactions. As residents of the plasma membrane, glycerophospholipids also serve as a substrate pool for the production of second messengers regulating cellular signaling events at the initial site of receptor activation.

Recent reports advocate the involvement of lipids in signaling pathways modulating cell survival, proliferation, and migration, as well as in pathophysiological disease states including inflammation, angiogenesis, and cancer (Pyne and Pyne, 2000; Mills and Moolenaar, 2003). For example, lysophosphatidic acid, produced by the enzymatic activity of a variety of phospholipases, including phospholipases A1, A2, and the lysophospholipase D autotaxin, plays a vital role in a variety of cellular and biological actions that increase motility and invasiveness of cells (Mills and Moolenaar, 2003). In addition, evidence supports the involvement of lysophosphatidic acid and the lysophospholipid sphingosine-1-phosphate in hematopoiesis or stem cell trafficking, a process crucial in stem cell biology and bone marrow transplantation (Whetton et al., 2003). Sphingosine-1-phosphate has also been implicated in angiogenesis, a critical process of cancer progression (Wang et al., 1999). Second messengers derived from precursor phospholipids also act as mediators in inflammation and neurodegeneration. Studies report that prostanoid production from arachidonic acid, the starting molecule of the proinflammatory eicosanoids, is important in regulating vital aspects of the inflammatory response seen in arthritis and asthma (Heller et al., 1998). In addition, arachidonic acid regulates neural membrane biology, including protein-lipid interactions and *trans*-synaptic signaling, together with abnormalities in these pathways that have been described as contributing to the pathophysiology of Alzheimer's disease (Bazan et al., 2002). The diverse roles of lipids in cell functions and disease processes have stimulated renewed interest in phospholipids and encouraged development of improved methods to determine comprehensive changes in membrane lipid composition.

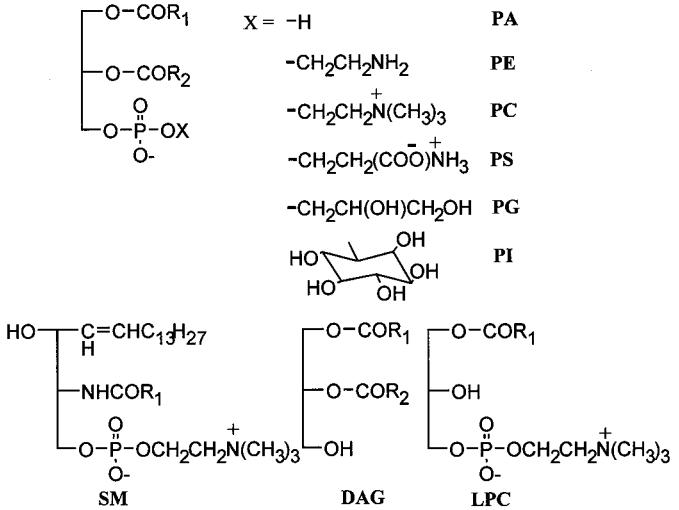


Fig. 1. Structures of glycerophospholipid classes. In the upper left segment is shown the structure of compounds commonly referred to as phosphoglycerides. The R_1 and R_2 represent long-chain carboxylic acids usually connected via an ester bond to the primary and secondary alcohol residues of glycerol. The –X moiety refers to the headgroups shown in the upper right of the figure. Joined in ester linkage to the phosphoric acid residue they yield PA, PE, PC, PS, phosphatidylglycerol (PG), and PI, respectively. In the lower figure is illustrated the structure for SM, diacylglycerol (DAG), and one configuration of LPC.

Resolution and Identification of Phospholipids

Methods for Phospholipid Analysis. The number of distinct species of lipids in any given cell type is unknown. The need for mammalian cells to generate such divergent species (reviewed by Dowhan, 1997) has been the subject of debate in recent years. Traditionally, phospholipid species and lipid classes have been characterized by a combination of various analytical techniques, including thin-layer chromatography, gas chromatography, and high-performance liquid chromatography (Christie, 2003). In most applications, these methods require intricate multistep preparations and relatively large amounts of sample. Lipid analysis has also been hampered by the complexities involved in resolving lipid extracts at the individual species level, obscuring the amount of intraclass specificity involved with their in situ biological reactions. Moreover, many bioactive lipids occur in extremely low cellular concentrations and are difficult to resolve and measure. Advances in mass spectrometry have created a vehicle for the measurement of global changes in the cellular lipome through their high sensitivity and ability to comprehensively examine complex biological extracts.

Identification of Phospholipids by Mass Spectrometry. Until the 1980s, the primary ionization source for mass analysis was electron impact or chemical ionization. The challenges arising from sample desorption and ion formation associated with these ionization methods limited researchers to small molecules and excluded many of the larger thermally labile molecules found in biological systems. The introduction of the "soft" ionization techniques including fast atom bombardment, matrix-assisted laser desorption/ionization, and electrospray ionization (ESI) revolutionized the ionization capabilities for mass analysis. The use of electrospray mass spectrometry on phospholipid extracts has made the routine analysis of these complex biological samples possible (Han and Gross, 1994, 1995, 1996; Kerwin et al., 1994; Kim et al., 1994; Smith et al., 1995; Brügger et al., 1997; Fridriksson et al., 1999; Khaselev and Murphy, 2000; Ivanova et al., 2001; Hsu and Turk, 2003). Comprehensive reviews on the development and application of mass spectrometric methods for analysis of lipids have recently been published (Murphy et al., 2001; Pulfer and Murphy, 2003). Similar efforts are underway in the analysis of other classes of signaling lipids, including sphingolipids (Sullards and Merrill, 2001) and eicosanoids (Dickinson and Murphy, 2002), and as such will not be detailed here.

ESI is a convenient ionization technique developed by Fenn et al. (1989) that is used to produce gaseous ions from highly polar, mostly nonvolatile biomolecules such as lipids. The sample is injected as a liquid at low flow rates (1–10 μl/min) through a capillary tube to which a strong electric field is applied. The field generates additional charges to the liquid at the end of the capillary and produces a fine spray of highly charged droplets that are electrostatically attracted to the mass spectrometer inlet. The evaporation of the solvent from the surface of a droplet as it travels through the desolvation chamber increases its charge density substantially. When this increase exceeds the Rayleigh stability limit (Blakeley and Vestal, 1983; Taflin et al., 1989), ions are ejected and ready for MS analysis. This is a relatively mild ionization process causing minimal fragmentation or thermal decomposition. The technique is directly applied to solutions without derivatization, with high accuracy, sensitivity, and reproducibility.

The generation of lipid arrays requires the identification of individual species represented by peaks in the mass spectrum. This identification process begins with a full scan in both positive and negative mode to determine the *m/z* values referring to a molecular ion's monoisotopic molecular weight. The detection and resolution of phospholipid classes is based on the ability of representative molecules to acquire positive or negative charges under the electrospray high energy source. The primary type of ionization for each molecule is based on its chemical structure, which leads to differences in the ability to detect particular classes in the two ionization modes. Zwitterionic phospholipids such as phosphatidylcholine (PC), lysophosphatidylcholine (LPC), phosphatidylethanolamine (PE), and sphingomyelin (SM) can be detected in either positive or negative ionization mode. They are more efficiently detected in positive mode with the exception of PE, which is detected equally well in either mode. By contrast, the anionic phospholipids phosphatidylserine (PS), phosphatidylinositol (PI), phosphatidic acid (PA), and phosphatidylglycerol are negatively charged at neutral pH and produce molecular ion peaks detected only in negative mode. The structural analysis of the individual molecular ion peaks is accomplished by tandem mass spectrometry. This involves subjecting an ion of interest to collision-induced dissociation (CID), where the molecule is fragmented because of the interaction with a collision gas. The resulting fragments can then be used to generate a product or "daughter" ion spectrum where headgroup and/or fatty acyl chain compositions are disclosed (Fig. 2). The results of this fragmentation are dependent on the instrument used and the chemistry of the molecule. Glycerophospholipids detected in positive ionization mode reveal mainly headgroup information upon CID either as a headgroup fragment peak (as in PC, LPC, or SM) or as a fragment peak formed by neutral loss of the headgroup (as in PE and lyso-PE). Fragmentation in negative mode yields *sn-1* and *sn-2* fatty acid residues, thus providing structural information on the acyl chain composition. There are also ions characteristic to the headgroup or its fragments, as well as common fragments for all glycerophospholipids (e.g., *m/z* 79 for PO_3^-, *m/z* 97 for $H_2PO_4^-$, *m/z* 153 for [glycerophosphate-H_2O-H]$^-$).

Based on the fragmentation patterns observed for each set of cellular extracts, a "fragmentation library" can be created allowing easy consultation and determination of phospholipid species within that cell type (Milne et al., 2003). Lipid species are denoted with the abbreviation for the phospholipid headgroup preceded by *xx*:*y*, where *xx* represents the total carbon number in the fatty acid side chain(s) and *y* represents the number of double bonds in them. The method described above makes possible the detection of more than 300 phospholipid species (Milne et al., 2003). The immense amount of data resulting from these techniques became the rate-limiting step. The need to efficiently handle this magnitude of information led to the development of computational analysis software.

Mathematical Analysis of Mass Spectrometry Data

Large Data Sets and Comprehensive Analysis. Modern computing power is rapidly changing the way experimental data are viewed, analyzed, and conceptualized. Traditional experimentation often involves the effort to verify a hypothesis through apparent differences observed in a select few output variables as a system evolves through time and possibly under multiple starting conditions. Because of both experimental and modeling constraints, few examples exist of an attempt to decipher the behavior of a large collection of variables changing with respect to time in a complicated dynamic system. This represents, however, the necessary paradigm for achieving the next level of understanding in the complex inner workings of biological models such as cellular signaling.

In this section, we address the computationally intensive portion of functional lipidomics, which has emerged as a result of the ability to identify and measure hundreds of lipids simultaneously from biological extracts. The analysis combines statistical and computational techniques that can be applied to high-throughput in vivo screening of compounds to assay their global effect on the cellular lipome. The assaying of agonist-induced perturbations in cellular lipid concentrations can provide insight into the metabolic and signaling pathways, yielding a hypothesis-generating technique for the creation of highly specific investigations. The interpretations of findings must be validated by conventional hypothesis-driven research, but the comprehensive nature of the findings typically suggests starting points for verification. The data analysis was approached from the unbiased vantage point of pattern analysis (i.e., Are there repeatable

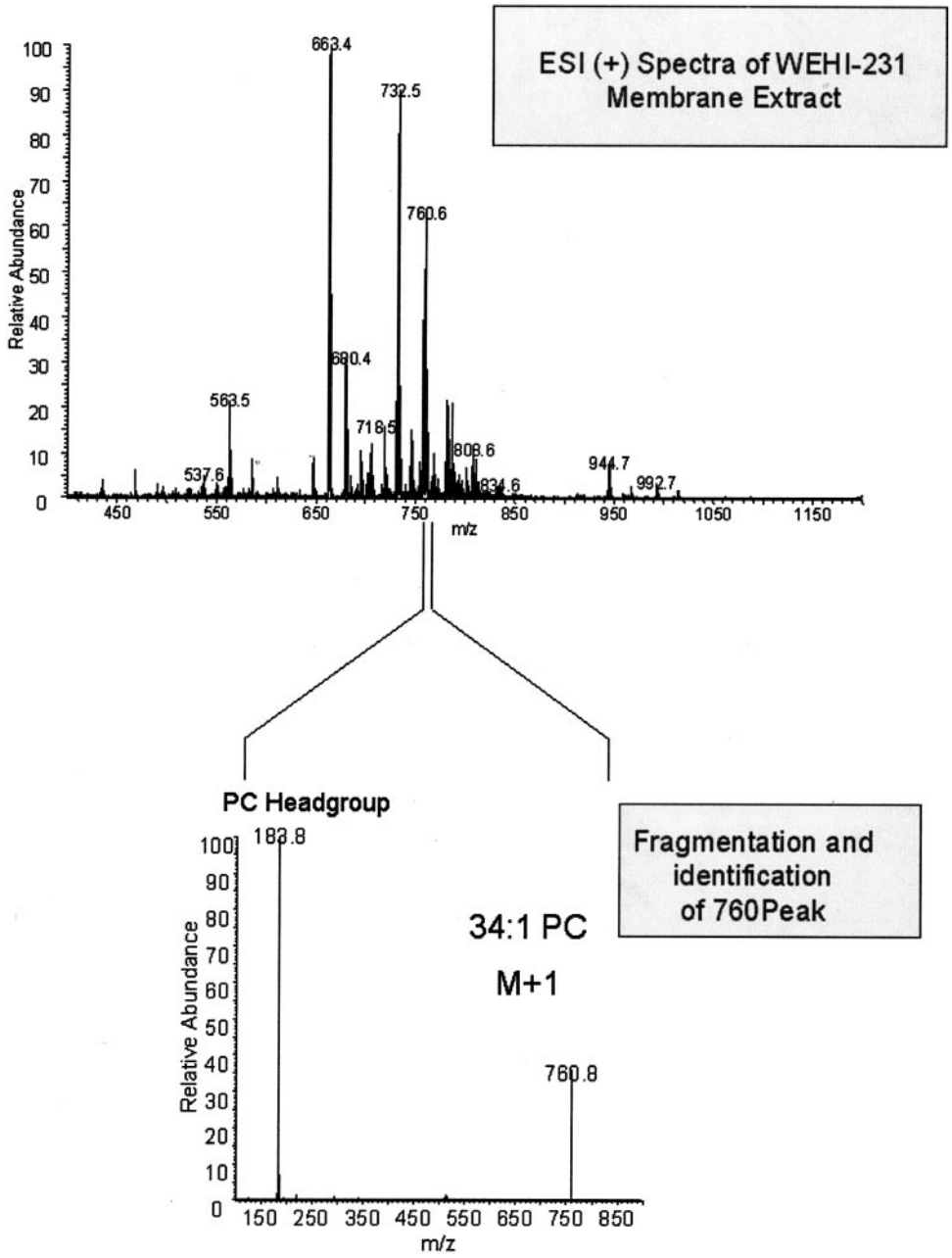


Fig. 2. Fragmentation and identification of WEHI-231 phospholipids. Individual lipid species from the total cell extract were subjected to CID. In the example shown, the 760-*m/z* peak was fragmented. The resulting fragmentation pattern identified the peak as 34:1 PC. The *m/z* values shown are experimentally derived averages of 60 scans (full scan) and 15 scans (fragmentation scan). Slight differences from exact masses may be caused by instrument calibration and measurement variability.

pattern changes? If so, identify.). Significant and reproducible patterns in lipid species changes provide a wealth of information for further study and possibly contain findings of predictive or diagnostic value.

The goal of the mathematical formulation presented here is the comprehensive analysis of qualitative changes in cellular lipid concentrations, as measured by mass spectrometry, between two experimental conditions as a biological system evolves through time. Computer algorithms for the processing of the data have been developed using the *S*-Plus Version 3.3 for Windows programming suite. The analysis programs were constructed around two central concepts, which are detailed below.

Data Normalization. Using peak intensities for assaying relative amounts of components in an observed mass spectrum proves troublesome, unless the compounds have been carefully prepared and calibrated against an internal standard (Siuzdak, 1996). Sample-to-sample fluctuation observed in these intensities for a particular *m/z* ratio can be large, even in apparent exact replicates. Moreover, ionization efficiencies for compounds vary, even within species of the same class, and do not correlate directly in a known way a priori to their molar concentration. This generated the need for the development of a normalization method involving the creation of a unitless number that would prove to be a more robust measure of the signal strength at a particular *m/z* value with respect to the overall pattern observed. The goal of this normalization is to create conditions in which the experiments are comparable with one another in a statistical sense, thereby allowing the construction of a qualitative map of lipid changes at the cellular level. The results of this analysis could be expanded, for example, through the inclusion of internal standards for changing compounds to determine the magnitude of the changes. Two choices for this unit-less number were obvious candidates: signal strength in standard statistical units and signal ranks.

In the first method, the mean and standard deviation of the intensities observed at all *m/z* values in the scanning range are computed. These statistics are then used to transform the signal intensity at each of the *m/z* values as $I^* = (I - \text{mean})/SD$, with the transformed intensity represented as the number of standard deviations the signal occurs above or below the mean signal strength. Thus, a signal with intensity equal to the mean intensity of the data set would receive a score of zero, and any signal with intensity below the mean would receive a negative score. This normalization scheme has the effect of equating the first two statistical moments in each data set, (i.e., the transformed data has a mean of zero and a variance of 1). The second method involves using the rank of the signal compared with all the other points in the data set as the transformed intensity measure. Thus, for a scanning range of 1000 *m/z* with a gradation of 0.10, the observed intensities would be mapped in a one-to-one correspondence with the integers 1 to 10,000. This method has proven to be highly robust against the wide changes in signal magnitude observed. After completion of the normalization, the transformed intensity signals are then carried into the second part of the analysis.

Use of Shewhart Control Charts. Shewhart Control charts are statistical devices used to detect process changes in complex systems as they evolve through time (Shewhart, 1931). Their primary function is to sort out random variation (noise) from special cause variation (signal) as a process evolves along a time axis. The basic procedure involves drawing samples of size *n* from the process under study at various time points and computing a statistic of interest such as the sample mean. These values are then plotted along the time axis and a set of *Control Limits* is calculated for the statistic computed. These limits represent the expected variability in the statistic and are computed from the process output, assuming the underlying distribution remains stable. As a result, a kind of running hypothesis test is constructed. Figure 3 is a control chart for the sample mean constructed from five data sets each containing four measurements taken with respect to time. The means of the sets are connected with a solid line. The chart also shows the grand mean as well as the lower and upper control limits. These limits are constructed as the three σ limits for the variance in the sample mean as estimated from the average sample standard deviation. For a statistical grounding in the construction of these limits, recall that even if the distribution of the individual measurements is not Gaussian, as long as it is essentially unimodal, the means of these observations should be reasonably approximated with a Gaussian distribution as a consequence of the Central Limit Theorem. The area between these limits represents the expected variability in the *mean* of four observations of the process, not the individual measurements.

A process is said to be "in-control" if it exhibits only random variation; i.e., all points (means in this case) are within the control limits and no nonrandom patterns are present. To aid in the detection of nonrandom patterns, the areas between the control limits and the grand mean are divided into three zones labeled A, B, and C as they proceed toward the center of the chart. These zones represent one-σ distances from the grand mean and can be used to provide additional statistical tests for signal drift, in that the sample means should fall within the two C zones with probability 0.68 and within the C and B zones taken together with probability 0.95. For example, if two means within a cluster of three adjacent time points occur in zone A, this would be taken as an indication

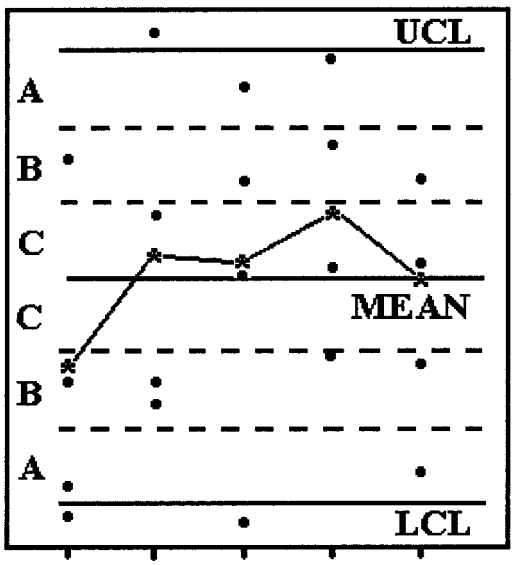


Fig. 3. A Shewhart Control Chart for the group means of five data sets each with four repetitions is shown. The individual measurements are indicated by ●, whereas the intragroup averages are indicated by * and connected by a solid line.

that the signal mean has undergone a shift in its underlying distribution. If a process is found to be in control, it is concluded to be stable over the given time course, and the control limits generate a profile for the variation in the measured statistic. The signal represented in Fig. 3 is in control over the time course studied. Although the group means are within the control limits, some of the individual measurements are not. From a biological standpoint, *m/z* transformed signal values shown to be in-control over the time course in the basal condition represent molecules in which metabolic cellular events are negligible, as measured by mass spectrometry.

Computational Analysis

The analysis programs used in the construction of the lipid arrays use the concepts developed above. Data are collected from the mass spectrometer as Raw files in the Xcalibur software package from Thermo Finnigan (San Jose, CA). These files are translated into text files for reading into S-Plus. In the first part of the analysis, the data from each spectrum are smoothed using a Nadaraya-Watson kernel regression estimate (Nadaraya, 1964; Watson, 1964) to remove the shoulder effect, which produces extraneous peaks. An example is given in Fig. 4. After smoothing, each data set is normalized using one of the transformations described above. The computer then parses the data set, looking for "Low-High-Low" patterns and flagging each "High" point as a peak, collecting them as a data table. During this concatenation, the locations (*m/z* ratio) of the peaks are averaged to compensate for measurement error in the mass spectrometer.

In the next stage of the analysis, the program constructs a Shewhart control chart for the mean of the transformed signal at each peak identified in the basal condition, and tests to determine whether the signal is "in control". This includes parsing the data for means that occur beyond the control limits as well as using the control chart zones to look for nonrandom patterns. The analysis then uses the control limits obtained from the basal condition to examine the output from the stimulated case at all *m/z* values where the signal is found to be "in control". In these instances, the signal has been stable over the time course in the basal condition and, using the basal control limits for comparison, it allows for a pooling of the information contained there. Thus, we plot the data from the stimulated condition on a control chart generated from the basal data (Fig. 5).

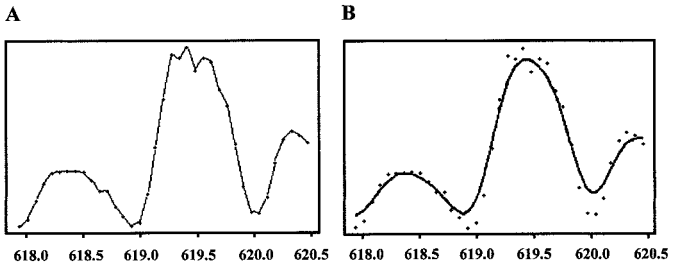


Fig. 4. Data collected from the mass spectrometer in the 618.0- to 620.5-*m/z* range before and after smoothing. A, unsmoothed data connected with a solid line indicating 3 separate peaks in the 619.0-to-620.0 range with a total separation of 0.28. B, the result of the Kernel Density Estimate shows a smoothed version of this data with a single representative peak in this range occurring near 619.5.

The analysis uses the rules described above to examine the stimulated data for nonrandom variation compared with the profile generated in the basal condition. This includes parsing for time points beyond the control limits as well as searching for nonrandom patterns that can be deduced from the control chart zones. As seen in the right of Fig. 5, the third and fourth time points are indicated as "out of control" because they consecutively appear in the A zone of the extended control chart. Thus, this *m/z* value would be scored as having increased at both the third and fourth time points in the stimulated data set.

Two other possibilities also require explanation here. In the first case, the basal data behaves in an "out-of-control" manner, (i.e., contains some nonrandom time-related variation). The second possibility involves peaks that appear in different frequencies within the basal and stimulated conditions. When the basal condition exhibits "out-of-control" variation, extending the control limits would be inappropriate. In this instance, a Welch modified two-sample *t* test is performed at each of the time points to determine whether differences exist in the means between the two conditions at the given time. In the second case, a binomial test is performed, with the null hypothesis that a peak has an equal chance of appearing in either the basal or stimulated case, to determine whether the observed difference in the number of occurrences in the two conditions is significant.

Generation of Lipid Arrays. After testing for statistically significant differences between experimental conditions at each time point-peak combination, the results are grouped into a comprehensive array containing the *m/z* values observed as peaks and the time course, on the vertical and horizontal axes, respectively. An excerpt of a lipid array is shown in Fig. 6. Peaks that have been identified by tandem mass spectrometry are documented as specific lipid species. Each *m/z* and time point combination found to be increasing is scored with a positive one (+1), whereas those decreasing are assigned a negative one (−1). Statistically stable combinations are scored with a zero. These arrays can be color-coded to enhance readability and in many cases provide a striking display of cellular lipid changes through time after challenge with a biological agonist.

When examining a system that includes a large number of time point-peak combinations, a significant opportunity for false-positives is created. This is illustrated by considering

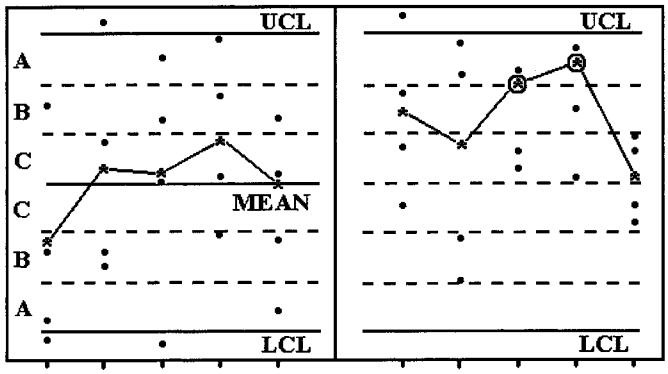


Fig. 5. Control chart for data in two experimental conditions. The left contains the control chart generated from the data in Fig. 2. The right extends the limits found in the basal condition and plots the data sets and their means for the stimulated condition. Means meeting the defined criteria for out of control are indicated by a circled *.

that if 1000 different peaks are analyzed over five time points, it generates 5000 chances for a false-positive. If the α value is set at 0.05, one would anticipate that 250 false indicators on the lipid array occur by chance alone. Methods involving the reduction of the false positive rate by effectively decreasing the α value (e.g., the Bonferroni inequality) have been found to be extremely detrimental to the sensitivity of the detection process. As an alternative, repeating the experiment multiple times and summing the cells from the resulting arrays can counter this effect. Because random errors are unlikely to occur in the same position, after several repetitions, the result is seen to converge as a stable map of lipid changes.

Proof-of-Concept. To demonstrate the ability of a mathematical formulation such as the one described above to identify subtle differences between similar biological extracts, a proof-of-concept experiment was preformed. This experiment was designed to determine the efficacy of the mathematical algorithm in locating the components of a chemically defined cocktail of lipids added to cellular extracts and to assess the resulting false alarm rate. The admixture was constructed from commercially available phospholipid preparations obtained from Avanti Polar Lipids (Birmingham, AL). The chosen lipid standards were supplemented at the indicated concentrations: 34:1 PA (75 μg/ml), 16:0 LPC (250 μg/ml), 34:2 PI (100 μg/ml), 16:0 PE (200 μg/ml), 28:0 PE (200 μg/ml), and 32:0 PE (200 μg/ml). The 34:2 PI standard was a complex mixture isolated from soy plant extracts composed of a major 34:2 PI species and minor amounts of 34:3, 36:4, 36:5, and 36:6 PI as well.

HL-60 pellets containing $\sim 10 \times 10^6$ cells were extracted using a modified Bligh and Dyer procedure and dried in a Centrivap Concentrator (Labconco Corporation, Kansas City, MO). Samples were resuspended in a chloroform/methanol solution into which the chemically defined standards were added. The samples were analyzed via direct injection in the negative ionization mode. Five sets of data were generated, each consisting of three repetitions that included HL-60 extracts in the absence and presence of the external standards. The data were analyzed using a modified version of the procedure described above where the means of the transformed signals were compared using a Welch modified two-sample t test (lack of a time course precludes the use of control charts) with the α value for the individual tests set at 0.005. Lipid arrays were constructed for the five experiments and the results were combined to form a summary array of the observed lipid changes. A total of 887 peaks were identified and analyzed by the software. Peaks that were found to be changing in the spiked sample condition were subjected to CID MS/MS to determine their molecular species. Of the 887 peaks analyzed, 87 (9.8%) exhibited a unidirectional change in four or five of the experiments performed, which was defined as highly significant. Of these 87 peaks, 59 were indicated as increasing, whereas 28 were found to be decreasing. All six standards were located in the spectra, including the complete collection of PI lipids, each receiving a significance score of 5, except for 32:0 PE and 36:6 PI, which were scored with a 4. Many of the remaining increasing peaks were subsequently identified using CID MS/MS, and several were found to be gas-phase dimers of the various standards added. Of the total collection of peaks found to be increasing, 20 were found to be consistent with the addition of the standards. This yields a false alarm rate of 67/887 = 7.6%. The results are summarized in Fig. 7.

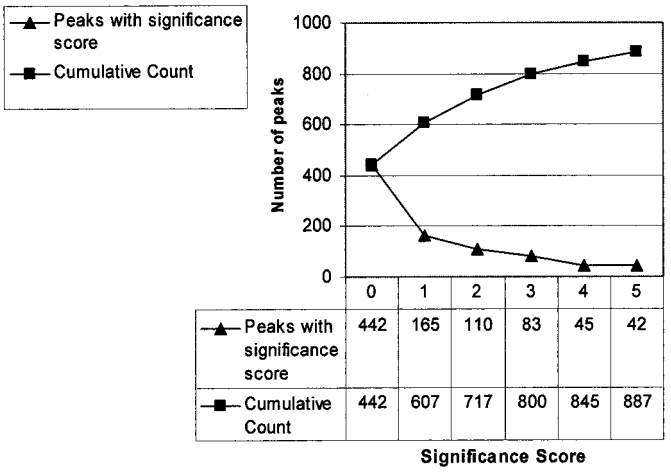
Lipid	m/z	Time (Minutes)							
		1	3	6	9	12	15	20	30
18:3 LPC	518.4	0	0	1	0	1	0	0	1
	519.4	0	0	0	0	0	0	0	0
18:2 LPC	520.4	0	0	0	0	0	1	1	0
	521.4	0	0	0	0	0	0	0	0
18:1 LPC	522.4	0	0	0	0	0	1	1	1
	523.4	0	0	0	0	0	0	0	0
18:0 LPC	524.3	0	0	1	1	1	1	1	1

Fig. 6. An example of one round of analysis is shown in this excerpt from the lipid array of RBL-2H3 mastocytoma cells challenged with antigen over the indicated time course (1, 3, 6, 9, 12, 15, 20, 30 min). This excerpt is from the positive ion scanning mode for the m/z range of 518.4 to 525.3. These were shown by CID MS/MS to correspond with several species of LPC. Array cells scored with a 1 (in red) indicate a time point where the stimulated data showed a statistically significant increase over the basal condition.

Example of Lipidomics Approach

MS Analysis of WEHI-231 Cells. As a demonstration of multiplexed lipidomic analysis, we present results from WEHI-231 cells, challenged with the B-cell receptor (BCR) agonist anti-IgM (AIG, 0.13 μM). These data were collected in conjunction with The Alliance for Cellular Signaling (AfCS) as one of several ligand-induced cellular response assays. The AfCS consists of seven experimental laboratories coupled with a bioinformatics core, focused on the overall goal of understanding the relationships between the inputs and outputs of signaling cells in a context-dependent manner. To achieve this goal, the AfCS laboratories have been applying existing and developmental technologies to acquire data from a large collection of cellular events. This data are to be reduced into a set of theoretical models to aid in the understanding of ligand response pathways. For additional information on the AfCS project, visit the web site at http://www.signaling-gateway.org.

The WEHI-231 samples were extracted using a modified



	0	1	2	3	4	5
Peaks with significance score	442	165	110	83	45	42
Cumulative Count	442	607	717	800	845	887

Fig. 7. Cumulative count of peaks versus absolute significance score. The number of peaks observed (▲) at a particular absolute significance value is seen to decrease rapidly as the score increases, with 87 being highly significant. This number is computed for each peak as the absolute value of the sum for the five individual experiments. A peak indicated as increasing or decreasing in four of the five tests would be plotted here as a four. Cumulative count (■) displays total number of peaks with that significance score or less.

Bligh and Dyer procedure and analyzed as described elsewhere (Milne et al., 2003). Identification of the individual glycerophospholipids present in the total lipid extracts (both basal and AIG-stimulated) was accomplished by tandem mass spectrometry (ESI-MS/MS). CID of the peaks of interest yielded fragmentation patterns that were used to identify the lipid(s) present at a particular *m/z* value. More than 300 glycerophospholipids have been simultaneously detected and identified in WEHI-231 total lipid extracts.

Instrumentation. Mass spectra were acquired on a Finnigan TSQ Quantum triple quadrupole mass spectrometer (Thermo Finnigan) equipped with a Harvard Apparatus syringe pump and an electrospray ionization source. Samples were analyzed at an infusion rate of 10 μl/min in both negative and positive modes over the range of *m/z* 400 to 1200. The data were collected at a resolution of 0.07 *m/z* units, producing 11,428 intensity measurements per run. Mass spectrometer parameters were optimized with 1,2-dioctanoyl-*sn*-glycero-3-phosphoethanolamine (16:0PE). Mass spectral data were collected using the Xcalibur software package (Thermo Finnigan).

Glycerophospholipid Changes in Basal versus AIG-Stimulated WEHI-231 Cells. Stimulation of the AfCS WEHI-231 BCR with AIG (0.13 μM) resulted in robust changes in glycerophospholipid concentrations. Lipid arrays were constructed for both the positive and negative ESI modes from 10 sets of samples. Each data set contained four replicates of paired samples that included a control (basal) and matched ligand (AIG) stimulated sample at each of five time points: 1.5, 3, 6, 15, and 240 min. Thus, each summary array was constructed from the summation of the 10 individual arrays and contains 400 samples. Excerpts from the positive (A) and negative (B) mode arrays are shown in Fig. 8. A total of 935 peaks were identified in the positive mode and 827 in the negative mode. Of 935 peaks in the positive mode, 810 (86.6%) remained statistically stable at all time points. Similarly, 722 (87.3%) of the negative peaks remained stable at all time points.

From a temporal perspective, lipids identified in both the positive and negative scanning modes showed little change after stimulation at the early time points (1.5 and 3 min), modest changes at 6 min, and the greatest differences were observed at 15 min. The majority of these lipid changes were seen to return to their basal state 4 h after stimulation. In the array for the positive ion mode, highly significant decreases were observed for many PC and/or PE species at the 6- and 15-min time points, with concomitant increases in several LPC compounds (Fig. 8A). In the negative ionization mode, clusters of PIs and PSs showed a highly significant decrease at the 15-min time point. Corresponding increases in lyso-PI, lyso-PS, and glycerophosphatidic acid species were also observed. An excerpt demonstrating the PS fraction of this result is shown in Fig. 8B. For a complete version of the lipid arrays, refer to AfCS web site (http://www.signaling-gateway.org/reports/v1/DA0011/WEHIIgMPos.pdf and http://www.signaling-gateway.org/reports/v1/DA0011/WEHIIgMNeg.pdf).

Activation of the BCR by AIG in mature B cells results in a wide variety of cellular changes that include cell proliferation, differentiation, increased metabolic rate, and changes in cellular adhesion properties. In addition, these cells exhibit germinal center reactions, immunoglobulin isotype class switch DNA recombination, and somatic hypermutation of immunoglobulin V regions. Interestingly, although stimulation of the BCR results in activation and proliferation of mature B cells, it results in apoptosis in immature B cells. A detailed characterization of other changes induced is provided on the AfCS website. Many of the lipid changes observed during the stimulation of the BCR seem consistent with the associated cellular behavior. For example, highly significant decreases in the various PI species were observed and probably occur in response to the phospholipase C activation (i.e., hydrolysis and subsequent biosynthesis of PI 4,5-bisphosphate). The decreases in the PS lipids with the increase in lyso-PS may be functionally linked to initiation of the apoptotic pathway in the B-cell. The interpretation of the lipidomic profile generated, and its relationship to the overall signaling pathway, will provide exciting new challenges for the signaling community but is likely to implicate unanticipated participants in familiar biological processes.

Conclusions and Future Prospects. The importance of lipids as bioactive molecules in cellular signaling events is a driving force in the generation of improved technologies for their measurement. Advances in instrumentation, such as Q-TOF mass spectrometry, are improving the state of the art for lipid analysis immensely. The methodology is more accurate, sensitive, and becoming more amenable to high-throughput approaches. Precise identification of head groups and fatty acyl compositions has made it possible to identify individual species, rather than grouping together heterogeneous mixtures containing similar numbers of carbon atoms and unsaturated bonds. Computational lipidomics has the potential to create time-based comprehensive descriptions of

A

Lipid	m/z	Time (Minutes)				
		1.5	3	6	15	240
34:3 PC	756.7	0	-2	-3	-6	-3
	757.7	0	-2	-3	-7	-1
34:2 PC	758.7	0	-1	-5	-6	-2
	759.7	-3	-2	-6	-5	-4
34:1 PC 34:2 PS	760.7	-1	-2	-4	-5	-2
	761.7	-3	-2	-6	-6	-2
34:0 PC 38:0 PEe	762.7	-2	-2	-3	-6	0
	779.7	0	0	-2	-2	-1
36:5 PC	780.7	-1	0	-3	-5	-4
	781.7	0	-2	-3	-4	-6
36:4 PC	782.7	1	-1	-7	-6	-1
	783.7	1	-2	-6	-7	-2
36:3 PC	784.7	-1	-3	-5	-5	-3
	785.7	1	-2	-2	-4	-1
	545.5	1	2	-1	4	2
20:3 LPC	546.5	0	0	-1	1	1
	547.5	3	4	3	5	2
20:2 LPC	548.5	2	2	1	2	1
	549.5	1	1	2	4	1
20:1 LPC	550.6	5	3	3	6	2
	551.6	4	3	6	9	6

B

Lipid	m/z	Time (Minutes)				
		1.5	3	6	15	240
32:1 PS	732.4	-1	-2	-1	-3	0
	733.4	0	-1	-1	-3	1
32:0 PS	734.5	0	-2	-2	-5	3
	735.5	3	-5	-2	-7	1
36:4 PS	782.5	0	-1	-3	-7	0
	783.5	0	0	-1	-3	1
36:3 PS	784.5	-1	-1	-3	-6	-2
	785.5	-1	0	-2	-5	-2
36:2 PS	786.5	0	-2	-1	-4	0
	787.5	-1	-1	-1	-3	-1
38:4 PS	810.5	-2	-4	-1	-7	2
	811.5	0	-4	0	-8	1
38:3 PS	812.5	0	-1	1	-3	0
	813.5	-1	-2	-1	-4	-1
38:2 PS	814.5	1	-2	-1	-5	-3
	815.5	1	0	0	-2	0
20:3 LPS	546.3	2	1	3	4	-2
	547.3	3	2	3	3	1
	548.3	1	3	4	6	0
	549.2	0	2	0	2	0
20:1 LPS	550.3	1	2	4	6	0
	551.3	1	2	3	5	-1

Fig. 8. Excerpts from the WEHI-231 lipid arrays. A, negative ion mode. B, positive ion mode. The second column of each table indicates the mass-to-charge ratio (*m/z*), and the first column contains lipid species identified by CID MS/MS. The remaining five columns show the significance score, which is the number of experiments (of 10) in which the indicated peak (*m/z*) was shown to be statistically distinct from the basal at that respective time point, with positive numbers representing increasing and negative numbers indicating a decreasing signal. An array cell containing the number −7 is interpreted to mean the indicated species was observed to decrease in 7 of the 10 trials at that time point. These cells are color coded by signal frequency, with deep blue or red indicating 6 or more occurrences of 10 (shown highly significant through simulation) and light blue or red indicating 5 occurrences (shown significant through simulation). Green cells, representing a −4 to 4 significance score, indicate statistical stability between basal and stimulated conditions.

fluctuations in the cellular lipome through the combination of new and existing technologies. In the coming years, qualitative lipid analysis will be replaced by improved analytical approaches that permit precise quantitative determinations. Data of this type will prove highly useful in hierarchical clustering schemes as another method in differentiating receptor mediated cellular events involving lipid second messengers as well as membrane compositional remodeling. Identified pattern changes or lipid fingerprints can be used to identify and characterize cell-specific ligand combinations that are components of signal transduction pathways. Detailed identification of cellular phospholipid species, analysis of pattern changes, and how these relate to other signaling components is a central focus of the AfCS. We anticipate the comprehensive pattern changes in lipid composition will eventually be used systematically with proteomic and gene arrays to further advance the understanding of signaling pathways and contribute toward advances in molecular diagnostics. Recent reports show exciting advances in proteome-based diagnostics (e.g., Yanagisawa et al., 2003) and provide much encouragement for the new field of lipidomics. The ability to comprehensively measure real-time lipid expression provides an opportunity for the system to delineate further directions of investigation with less dependence on a priori knowledge of the processes at work. In addition, this analysis highlights the increasing necessity of interdisciplinary relationships between biology, chemistry, computer science, mathematics, and instrumentation for the solving of complex biological problems. Indeed, we can anticipate that the continued integration of these disciplines will usher in a new era of molecular diagnostics.

Acknowledgments

We gratefully acknowledge Robert Hsueh and our colleagues in the AfCS Cell Preparation and Analysis Lab for the high quality WEHI-231 cells used in some illustrations of lipid arrays. The complete set of arrays referred to in this review may be found on the AfCS web site, http://www.signaling-gateway.org. We thank Michelle Armstrong for expert assistance in mass spectrometry and Johnna Goodman for editorial and administrative support.

References

Alb JG, Cortese JD, Phillips SE, Albin RL, Nagy TR, Hamilton BA, and Bankaitis VA (2003) Mice lacking phosphatidylinositol transfer protein-α exhibit spinocerebellar degeneration, intestinal and hepatic steatosis and hypoglycemia. *J Biol Chem* **278:**33501–33518.

Bazan NG, Colangelo V, and Lukiw WJ (2002) Prostaglandins and other lipid mediators in Alzheimer's disease. *Prostaglandins Other Lipid Mediat* **68–69:**197–210.

Berridge MJ (1993) Inositol triphosphate and calcium signaling. *Nature (Lond)* **361:**315–325.

Bevers EM, Comfurius P, Dekkers DW, and Zwaal RF (1999) Lipid translocation across plasma membrane of mammalian cells. *Biochim Biophys Acta* **1439:**317–330.

Blakeley CR and Vestal ML (1983) Thermospray interface for liquid chromatography/mass spectrometry. *Anal Chem* **55:**750–754.

Brügger B, Erben G, Sandhoff R, Wieland FT, and Lehmann WD (1997) Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. *Proc Natl Acad Sci USA* **94:**2339–2344.

Christie WW (2003) *Lipid Analysis: Isolation, Separation, Identification and Structural Analysis of Lipids.* Oily Press, Bridgwater, England.

Dickinson JS and Murphy RC (2002) Mass spectrometric analysis of leukotriene A4 and other chemically reactive metabolites of arachidonic acid. *J Am Soc Mass Spectrom* **13:**1227–1234.

Dowhan W (1997) Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu Rev Biochem* **66:**199–232.

Fenn JB, Mann M, Meng CK, Wong SF, and Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science (Wash DC)* **246:**64–71.

Fridriksson EK, Shipkova PA, Sheets ED, Holowka D, Baird B, and McLafferty FW (1999) Quantitative analysis of phospholipids in functionally important membrane domains from RBL-2H3 mast cells using tandem high-resolution mass spectrometry. *Biochemistry* **38:**8056–8063.

Han X and Gross RW (1994) Electrospray ionization mass spectroscopic analysis of human erythrocyte plasma membrane phospholipids. *Proc Natl Acad Sci USA* **91:**10635–10639.

Han X and Gross RW (1995) Structural determination of picomole amounts of phospholipids via electrospray ionization tandem mass spectrometry. *J Am Soc Mass Spectrom* **6:**1202–1210.

Han X and Gross RW (1996) Structural determination of lysophospholipid regioisomers by electrospray ionization tandem mass spectrometry. *J Am Chem Soc* **118:**451–467.

Han X and Gross RW (2003) Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: a bridge to lipidomics. *J Lipid Res* **44:**1071–1079.

Heller A, Koch T, Schmeck J, and Van Ackern K (1998) Lipid mediators in inflammatory disorders. *Drugs* **55:**487–496.

Hokin MR and Hokin LE (1953) Enzyme secretion and the incorporation of P^{32} into phospholipids of pancreas slices. *J Biol Chem* **203:**967–977.

Hsu F-F and Turk J (2003) Electrospray ionization/tandem quadrupole mass spectrometric studies on phosphatidylcholines: the fragmentation processes. *J Am Soc Mass Spectrom* **14:**352–363.

Ivanova PT, Cerda BA, Horn DM, Cohen JS, McLafferty FW, and Brown HA (2001) Electrospray ionization mass spectrometry analysis of changes in phospholipids in RBL-2H3 mastocytoma cells during degranulation. *Proc Natl Acad Sci USA* **98:**7152–7157.

Kerwin JL, Tuininga AR, and Ericsson LH (1994) Identification of molecular species of glycerophospholipids and sphingomyelin using electrospray mass spectrometry. *J Lipid Res* **35:**1102–1114.

Khaselev N and Murphy RC (2000) Electrospray ionization mass spectrometry of lysoglycerophosphocholine lipid subclasses. *J Am Soc Mass Spectrom* **11:**283–291.

Kim HY, Wang TCL, and Ma YC (1994) Liquid chromatography/mass spectrometry of phospholipids using electrospray ionization. *Anal Chem* **66:**3977–3982.

Mills GB and Moolenaar WH (2003) The emerging role of lysophosphatidic acid in cancer. *Nat Rev Cancer* **3:**582–591.

Milne SB, Forrester JS, Ivanova PT, Armstrong MD, and Brown HA (2003) Multiplexed lipid arrays of Anti-Immunoglobulin M-induced changes in the glycerophospholipid composition of WEHI-231 cells. AfCS Research Report. Available at: http://www.signaling-gateway.org/reports/v1/DA0011/DA0011.htm.

Murphy RC, Fiedler J, and Hevko J (2001) Analysis of nonvolatile lipids by mass spectrometry. *Chem Rev* **101:**479–526.

Nadaraya E (1964) On estimating regression. *Theory Prob Appl* **9:**141–142.

Pulfer M and Murphy RC (2003) Electrospray mass spectrometry of phospholipids. *Mass Spectrom Rev* **22:**332–364.

Pyne S and Pyne N (2000) Sphingosine 1-phosphate signaling via the endothelial differentiation gene family of G-protein-coupled receptors. *Pharmacol Ther* **88:**115–131.

Ramanadham S, Hsu F-F, Bohrer A, Nowatzke W, Ma Z and Turk J (1998) Electrospray ionization mass spectrometric analyses of phospholipids from rat and human pancreatic islets and subcellular membranes: comparison to other tissues and implications for membrane fusion in insulin secretion. *Biochemistry* **37:**4553–4567.

Shen Z, Wu M, Elson P, Kennedy AW, Belinson J, Casey G, and Xu Y (2001) Fatty acid composition of lysophosphatidic acid and lysophosphatidylinositol in plasma from patients with ovarian cancer and other gynecological diseases. *Gynecol Oncol* **83:**25–30.

Shewhart WA (1931) *Economic Control of Quality of Manufactured Product.* D. van Nostrand Company, Inc., Toronto, Canada.

Siuzdak G (1996) *Mass Spectrometry for Biotechnology.* Academic Press, San Diego.

Smith PBW, Snyder AP, and Harden CS (1995) Characterization of bacterial phospholipids by electrospray ionization tandem mass spectrometry. *Anal Chem* **67:**1824–1830.

Sullards MC and Merrill AH Jr. (2001) Analysis of sphingosine 1-phosphate, ceramides, and other bioactive sphingolipids by high-performance liquid chromatography-tandem mass spectrometry. *Sci STKE* (67):PL1.

Taflin DC, Ward TL, and Davis EJ (1989) Electrified droplet fission and the Rayleigh limit. *Langmuir* **5:**376–384.

Wang F, Van Brocklyn JR, Hobson JP, Movafagh S, Zukowska-Grojec Z, Milstien S, and Spiegel S (1999) Sphingosine 1-phosphate stimulates cell migration through a G_i-coupled cell surface receptor. Potential involvement in angiogenesis. *J Biol Chem* **274:**35343–35350.

Watson G (1964) Smooth regression analysis. *Sankhya Series A* **26:**359–372.

Whetton AD, Lu Y, Pierce A, Carney L, and Spooncer E (2003) Lysophospholipids synergistically promote primitive hematopoeitic cell chemotaxis via a mechanism involving Vav 1. *Blood* **102:**2798–2802.

Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, Edgerton M, Gonzalez A, Nadaf S, et al. (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* **362:**433–439.

Address correspondence to: H. Alex Brown, Department of Pharmacology, Vanderbilt University Medical Center, 412B Preston Research Building, 23rd Ave South and Pierce, Nashville, TN 37232-6600. E-mail: alex.brown@vanderbilt.edu