Lipidomic analyses of *Mycobacterium tuberculosis* based on accurate mass measurements and the novel "*Mtb* LipidDB"

Mark J. Sartain,[1,*] Donald L. Dick,[†] Christopher D. Rithner,[†]

Dean C. Crick,[*] and John T. Belisle,[1,*]

Mycobacteria Research Laboratories,[*] Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO, 80523; Central Instrument Facility,[†] Department of Chemistry, Colorado State University, Fort Collins, CO, 80523

[1]To whom correspondence should be addressed:

Mark J. Sartain
1441 N 34th Street
Seattle, WA, 98109
Phone: (206) 732-2153
E-mail: msartain@systemsbiology.org

John T. Belisle
1690 Campus Delivery
Fort Collins, CO, 80523-1690
Phone: (970) 491-5384
Fax: (970) 491-8707
E-mail: jbelisle@colostate.edu

Running Title: LC/MS lipidomics strategy for *M. tuberculosis*

ABBREVIATIONS

Ac_2PIM_2, diacylated diacylglycerophosphoinositoldimannoside; Alpha-MA, alpha mycolic acid; amu, atomic mass units; CID, collision-induced-dissociation; DAT, diacyltrehalose; DIM A, phthiocerol dimycocerosate; DIM B, phthiodiolone dimycocerosate; GP, glycerophospholipid; GWE, glycosylated wax ester; MF, molecular feature; MFE, molecular feature extraction; *Mtb*, *Mycobacterium tuberculosis*; OD, optical density; PE, phosphatidylethanolamine, PI, phosphatidylinositol; PIM, phosphatidylinositolmannoside; SD, standard deviation; TG, triacylglycerol

ABSTRACT

The cellular envelope of *Mycobacterium tuberculosis* is highly distinctive and harbors a wealth of unique lipids possessing diverse structural and biological properties. However, the ability to conduct global analyses on the full complement of *M. tuberculosis* lipids has been missing from the repertoire of tools applied to the study of this important pathogen. We have established methods to detect and identify lipids from all major *M. tuberculosis* lipid classes through liquid-chromatography-mass-spectrometry (LC/MS) lipid profiling. This methodology is based on efficient chromatographic separation and automated ion identification through accurate mass determination and searching of a newly created database (*Mtb* LipidDB) that contains 2,512 lipid entities. We demonstrate the sensitive detection of molecules representing all known classes of *M. tuberculosis* lipids from a single crude extract. We also demonstrate the ability of this methodology to identify changes in lipid content in response to cellular growth phases. This work provides a customizable framework and resource to facilitate future studies on mycobacterial lipid biosynthesis and metabolism.

Supplementary key words: Mycobacterium, tuberculosis, lipidomics, lipid profiling, triacylglycerol, phthiocerol dimycocerosate

INTRODUCTION

Mycobacterium tuberculosis (*Mtb*), the causative agent of tuberculosis, possesses a cellular envelope rich with a complex milieu of lipids and carbohydrates that have been the subjects of intense research efforts for decades (1-3). Lipids comprise up to 60 percent of the *Mtb* cellular dry weight, and provide a hydrophobic barrier that enhances cell envelope impermeability. Furthermore, specific lipids of *Mtb* are shown to play critical roles in virulence and pathogenesis (4-6). The abundance and biological importance of the *Mtb* lipids has resulted in extensive and elegant studies to elucidate their structures and functions (1-3). In many cases the lipids of *Mtb* are unique to this pathogen or shared only with other members of this genus.

Earlier studies demonstrate variability in lipid profiles among different strains of *Mtb* (7-11) and that minor variations in the structure of individual lipids can occur with changes in the growth environment (12-20). However, targeted and nontargeted assays that monitor changes in *Mtb* lipid profiles are generally performed by traditional TLC-based methods (21) and global lipidomics analyses in *Mtb* have been restricted due to limits in the technology to detect and rapidly identify a large number of lipids in a single experiment. Two-dimensional NMR was recently applied to examine global mycobacterial lipid profiles and this approach allowed for the identification of key lipid differences in ^{13}C-enriched cellular extracts (22). Although this approach easily detects changes in lipid patterns it is limited by the complexity of the NMR spectra and the overlapping chemical properties of many lipids. Alternatively, MS-based lipidomic strategies allowing simultaneous detection, identification, and quantification of structurally diverse lipid components of *Mtb* also were evaluated. Leavell and Leary (23) developed an algorithm to analyze high-resolution Fourier transform-ion cyclotron resonance mass data obtained from direct infusions of complex *Mtb* lipid extracts. The algorithm isolates ions from spectra, and assigns identities from a "user-defined" lipid library based on exact mass. The methodology facilitated rapid comparisons of highly complex spectra and yielded significant findings on the metabolic control of virulence lipids (15). Shui et al. (24) took an alternative strategy and demonstrated the utility of C18 reversed-phase HPLC in combination with ESI-MS to efficiently separate and detect complex lipids of lower abundance. Several anionic *Mtb* lipid classes were characterized, and lipid profiling was able to identify mycolic acid profile shifts that occurred in response to different physiologic growth conditions. However, these previous MS based analyses were not supported by a complete and portable *Mtb* lipid database, and did not fully integrate database interrogation with the ability to resolve individual lipids by HPLC and ion-data extraction techniques.

Building upon the previous success of MS-based lipidomic approaches for *Mtb* and to provide tools for comprehensive lipid profiling we generated a novel database of *Mtb* lipids, "*Mtb* LipidDB" that allows for lipid identification using accurate-mass measurements obtained in either negative or positive ion modes. When coupled to molecular feature detection from LC/MS spectra the *Mtb* LipidDB allows for automated exact-mass based lipid identification. Applying this approach to crude *Mtb* lipid extracts we achieved identification of the majority of

known extractable lipids from this bacterium. Moreover, we demonstrated the utility of this methodology to rapidly identify quantitative changes that occur in the lipid profile of *Mtb* in response to growth conditions.

MATERIALS AND METHODS

Bacterial Growth and Lipid Extractions

Mtb strain H37Rv was propagated in glycerol-alanine salts medium (25) at 37 °C for 14 days as previously described (26). Cells were harvested by centrifugation (3,000 *g* for 10 min), washed three times with phosphate-buffered saline pH 7.4 (PBS), and inactivated by gamma irradiation (27). To prepare crude lipid extracts, a modified Bligh and Dyer (28) method was used. Briefly, 200 mg of wet cells were lyophilized in a 13 x 100 mm silanized borosilicate glass tube with a Teflon-screw cap. The dried cells were extracted with 6 mL of chloroform/methanol/water (10:10:3, v/v/v) overnight with constant stirring at room temperature. The sample was centrifuged at 3,000 *g* for 10 min to remove the delipidated cells. The monophasic lipid extract was transferred to a 13 x 100 mm glass tube, dried under a gentle stream of nitrogen, and stored at -20 °C until use. Prior to LC/MS analysis, the dried lipid extract was dissolved in chloroform/methanol/water (10:10:3 v/v/v) to approximate a 4 µg/µl concentration, centrifuged at 3,000 *g* for 10 min, and transferred to an autosampler vial. To prepare the dilution series, cells were homogenized by bead-beating and serially diluted two-fold with water in glass tubes. The diluted homogenates were lyophilized and lipids were extracted as described above. Based on the starting wet weight of the cell pellet, the lipid dry weights were estimated from routinely observed lipid extraction yields.

For *Mtb* H37Rv growth phase comparisons, three 500 mL Nephelo-sidearm flasks containing 50 mm stirbars and 200 mL Middlebrook 7H9 broth (Difco, Detroit, MI) supplemented with oleic acid-dextrose catalase (OADC, Difco) and 0.05% Tween 80 (Sigma-Aldrich, St. Louis, MO) were inoculated with 2 mL of a logarithmic culture to achieve a theoretical optical density at 580 nm (OD_{580}) of 0.004. The flasks were capped tightly and incubated at 37 °C under constant stirring at 150 rpm. The OD_{580} of each culture was measured at regular intervals for 11 days and aliquots (10 mL) were taken at 73, 108 and 265 hours, corresponding to logarithmic, transitionary and stationary growth phases, respectively. The cell pellets were harvested and washed 3 times with phosphate-buffered saline (pH 7.4). Crude lipid extracts were prepared as described above and randomized for LC/MS analyses.

Liquid Chromatography

An Agilent 1200 HPLC (Agilent Technologies, Palo Alto, CA) with a 2.1 (i.d.) x 150 mm, 3.5 µm XBridge C18 column (Waters Corp., Milford, MA) heated to 45 °C was used with a binary solvent system and a flow rate of 320 µL/min. A 2.1 (i.d.) x 10 mm, 3.5 µm XBridge C18 guard

column (Waters) was placed in series in front of the analytical column. The system was equilibrated with 100% of solvent A [5 mM ammonium acetate in methanol/water (99:1 v/v)], and an aliquot of the lipid extract (5 µL=20 µg dried extract) was applied to the column. Solvent A was maintained at 100% for 2.0 min, followed by a 30.0-min linear gradient to 100% solvent B [5 mM ammonium acetate in *n*-propanol/hexane/water (79:20:1 v/v/v)], and held at 100% solvent B for 3.0 min. All solvents and chemicals purchased were MS or HPLC grade.

Mass Spectrometry

An Agilent 6220 TOF equipped with an Agilent ESI/APCI multimode source was used for accurate mass analysis of the LC eluent. Positive (+) and negative (-) ion data were generated by operation of the mass spectrometer in a mixed ESI/APCI mode with a capillary voltage of 2000 V, nebulizer of 45 psig, drying gas of 8 L/min, gas temperature of 300°C, vaporizer temperature of 200°C, corona of 2 µA, fragmentor of 120 V, charging voltage of 2000 V, skimmer of 60 V, and OCT RF of 250 V. Mass spectra were acquired in 4 GHz High Resolution mode at a rate of 1.02 spectra/s and 9700 transients/spectrum, and data was collected as profiled spectra over a mass range of 250 to 3200 Da. Mass calibration was performed with an Agilent tune mix from 100 to 2700 Da, and an external reference sprayer introduced mass ions of *m/z* 922.009798 (+ ion) and *m/z* 980.016375 (- ion) to enable accurate mass determinations. Data was collected with the Agilent MassHunter WorkStation Data Acquisition software version B.02.00.

An Agilent 6520 qTOF was used for MS/MS analyses of the triacylglycerol (TG) lipid-group $[M+NH_4]^+$ ions. The instrument setup was the same as described above except the OCT RF was set at 750 V. Positive ion mass spectra were acquired in Auto MS/MS mode and collision energies with slope of 6.5V/100Da and offset 2.0V were used for fragmentation.

***Mtb* LipidDB Development**

To enable accurate-mass MS-based searching of lipid ions, the novel database "*Mtb* LipidDB" was created with Microsoft Excel 2007 Pro (Microsoft, Redmond, WA). The database was populated with molecular formulas, structures, and exact masses for the known lipids of *Mtb*. This information was obtained from a survey of the extensive *Mtb* biochemical literature describing lipid structures purified from *Mtb* H37Rv and *Mycobacterium bovis* BCG strains (see references in Supplementary Material). The collective data from these principal publications were used to calculate every theoretically possible fatty acyl combination for each lipid subclass and level 4 class. The reported structural variations were used to generate a representative molecular structure for each *Mtb* lipid subclass or level 4 class using ChemBioDraw Ultra 11.0 software (Cambridgesoft, Cambridge, MA). Lipids considered to be biosynthetic intermediates, such as phosphatidic acid, hydroxymycolates, and sugar-linked decaprenyl phosphates were not included in the *Mtb* LipidDB. Simple free fatty acids, including tuberculostearic acid, as well as lipid subclasses lacking sufficient detailed structural definition (i.e. definition of the fatty acid types), such as the triacylated trehaloses (29), were also not included in the database. Although

absent in the laboratory strain *Mtb* H37Rv used in this study, the glycosylated phthiodiolone dimycocerosates and glycosylated phthiocerol dimycocerosates (phenolic glycolipid) were included in the database due to their biological importance in clinical strains.

In addition to the parent *Mtb* LipidDB, two searchable database files named "+MH_MtbLipid.csv" and "-MH_MtbLipid.csv" were developed to interface with the Agilent MassHunter software and allow for searching of MS data against the *Mtb* LipidDB. The +MH_MtbLipid.csv and -MH_MtbLipid.csv files contained entries for each molecular ion that corresponded to a single lipid group in positive and negative ion modes, respectively. The MassHunter software provides the option to identify and collapse multiple molecular ions and adducts into a single feature. However, when only a single molecular ion was detected the software default was to subtract or add the mass of a proton (1.007276 Da) to the experimentally-derived molecular ion before interrogation of the +MH_Mtb_Lipid.csv and -MH_MtbLipid.csv database files, respectively. Thus, to allow for the matching of lipid groups that ionized as a single adducted molecular ion other than H^+ or H^-, (e.g. $[M+Na]^+$ and $[M+Ac]^-$ ions), each +MH_Mtb_Lipid.csv and -MH_MtbLipid.csv database file entry contained a mass value equal to the appropriate adduct but with 1.007276 Da subtracted from or added to the calculated molecular ion, respectively. The selection of molecular ions included in each searchable file was an iterative process based on empirically observed ionization properties of each lipid subclass. For instance, in positive ion mode $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, $[2M+Na]^+$, and $[M+2Na-H]^+$ ions were observed for phosphatidylethanolamines (PEs), while TGs were found to primarily generate $[M+NH_4]^+$, $[M+Na]^+$, and $[2M+Na]^+$ ions. Thus, with all 2,512 lipid groups represented, the +MH_MtbLipid.csv file contained 9,178 molecular ion entries and the -MH_MtbLipid.csv file contained 5,311 molecular ion entries. The *Mtb* LipidDB and the searchable database files are provided as Supplementary Material, and have been made freely available at http://www.mrl.colostate.edu.

Data Processing and Analyses

LC/MS data files were processed with the MassHunter Qualitative Analysis Software version B.02.00 (Agilent Technologies, Santa Clara, CA). Molecular features (MFs) were extracted from the raw data using the Molecular Feature Extraction (MFE) algorithm. This algorithm locates related co-variant ions (isotopes and charge states) from accurate-mass LC/MS data, and combines these ions into a single feature. The MF extraction parameters used were: Extraction Algorithm: small molecule, Peak Filters: ≥ 500 counts, Ion Species: +H and –H only, Peak Spacing Tolerance: 0.0025 *m/z* plus 7.0 ppm, Isotope Model: common organic molecules, Charge State: 1-2, Compound Filters: none, Mass Filters: none, Mass Defect: none. The resulting MFs were then identified with MassHunter by searching the +MH_MtbLipid.csv (+ ion data) or the -MH_MtbLipid.csv (- ion data) database file with the following search parameters: Values to Match: mass only, Match Tolerance: 5 ppm, Charge Carriers: +H and –H, Charge State

Range: 1-2. The lists of both identified and unidentified features were exported as an analysis report in Microsoft Excel (Microsoft, Redmond, WA) format. Further data comparisons were accomplished with Microsoft Excel. For quantitative purposes, lipid groups were compared using the most dominant molecular ion volumes observed for each lipid subclass as follows: TG: $[M+Na]^+$, phthiocerol dimycocerosate (DIM A): $[M+Na]^+$, phthiodiolone dimycocerosate (DIM B): $[M+Na]^+$, PE: $[M-H]^-$, phosphatidylinositol (PI): $[M-H]^-$, and cardiolipin (CL): $[M-H]^-$

RESULTS

Development of the novel *Mtb* LipidDB

The *Mtb* LipidDB (supplementary data and summarized in Table 1) was organized in a manner that adhered as closely as possible to the classification hierarchy and structural nomenclature set forth by the LIPID MAPS consortium (LIPID Metabolites and Pathways Strategy; http://www.lipidmaps.org) (30-31). Six of the eight lipid categories defined by LIPID MAPS are represented in the *Mtb* LipidDB. The *Mtb* LipidDB contains 15 lipid main classes, 46 lipid subclasses, and 16 level 4 lipid classes. In instances where *Mtb* lipids did not fit into the LIPID MAPs classification system, 30 novel lipid subclasses (e.g. diacyltrehaloses (DAT)) and 16 level 4 classes (e.g. alpha mycolic acids (Alpha-MA)) were created. These novel subclasses and level 4 classes were developed for *Mtb* LipidDB organization purposes only, and have not been accepted by LIPID MAPS as part of their classification hierarchy. A key organizational difference between the *Mtb* LipidDB and LIPID MAPs occurs at the species level of classification (Fig. 1). In lieu of "lipid species", the *Mtb* LipidDB uses "lipid groups", where each lipid group possessed a unique chemical formula corresponding to a unique exact mass. Unlike lipid species an individual lipid group was not distinguished by stereochemistry, unsaturated bond position, or the length and position of individual fatty acyl substituents. However, each lipid group allowed assignment of the head group composition and the sum composition (total fatty acyl carbon number and total degree of fatty acyl unsaturation) and in many cases represents multiple unique lipid species. The lipid group classification level was designed for the express purpose of simplifying exact mass-based MS identification, where a successful database query returns a single lipid group that could represent multiple isobaric lipid species. An additional justification for this classification level was the general paucity of sufficient detailed structural information at the lipid species level for many of the mycobacterial lipids reported in the current body of literature. In total the *Mtb* LipidDB contains 2,512 lipid groups, and when all potential ion-adducts were incorporated the searchable database files contain 14,489 mass entries.

LC/MS and automated ion identification allows lipidomic analysis of *Mtb*

A crude lipid extract from *Mtb* strain H37Rv grown in standard liquid medium was subjected to C18 RP-chromatography and the effluent directly analyzed by accurate-mass (ESI/APCI)-MS. Data collected in the positive- and negative-ion modes were processed using the MFE algorithm of the MassHunter software package, resulting in 1,916 and 744 MFs for the positive- and negative-ion spectra, respectively. The MF lists were searched against the +MH_MtbLipid.csv and -MH_MtbLipid.csv database files with a mass error tolerance of ± 5 ppm, yielding database matches for 672 of the 1,916 positive-ion MFs (35.1%) and 248 of the 744 negative-ion MFs (21.8%). Supplementary table S1 provides detailed attributes for each identified molecular feature. The identified MFs accounted for 74.8% and 58.4% of the total positive- and negative-ion MF peak abundances (ion volumes), respectively. This indicated that the unassigned MFs were likely low abundant products representing undefined ion-adducts, minor uncharacterized lipid structures, and/or background contaminating molecules. The absolute average mass accuracies were 1.46 (±1.60 SD) ppms for positive-ion MFs and 1.87 (±1.96 SD) ppms for negative-ion MFs. The combined positive- and negative-ion data resulted in the identification of lipid groups from all 6 lipid categories, all 15 lipid main classes, 31 of the 45 lipid subclasses, and 12 of the 14 level 4 classes represented in the *Mtb* LipidDB. It should be noted that the glycosylated wax ester (GWE) subclass is not produced by *Mtb* strain H37Rv and was not included in the number of potential lipid subclasses and level 4 classes. The assigned MFs, grouped by lipid subclass or level 4 class were plotted by molecular ion *m/z* and LC retention time (Figs. 2 and 3). This demonstrated the clustering of the lipid subclasses or level 4 classes based on retention time and mass. Moreover, the large numbers of plotted MFs revealed the complexity of the ion patterns that must be assigned for each *Mtb* lipid subclass or level 4 class. To further illustrate the sample complexity and the LC separation, extracted chromatograms for all identified molecular features are shown (Fig. S1).

Important observations to consider in MF assignment

As expected, it was possible to observe multiple MFs with nearly identical retention times that matched to different molecular ion entries from the same lipid group due to the formation of multiple ion adducts. For instance there was a preponderance of both $[M+Na]^+$ and $[M+NH_4]^+$ ions observed for TGs, while there was bias towards both $[M+H]^+$ and $[M+Na]^+$ ions for PEs. However, we also observed instances of multiple MFs with different chromatographic behaviors that matched to the same database entry (i.e. the same molecular ion). Specifically, there were 109 (+ ion mode) and 24 (- ion mode) MFs assigned to the same database entry as that of at least one other MF. This likely resulted from the detection of multiple isobaric lipid species that shared the same sum chemical composition, but disparate elution times caused by slightly different fatty-acyl repertoires. Thus, subtracting instances where multiple MFs matched to the same lipid group, a total of 415 unique lipid groups out of the 2,512 present in the database were identified in the *Mtb* lipid extract. Positive-ion and negative-ion data identified 314 and 193 unique lipid groups, respectively, with an overlap of 92 lipid groups.

The exact mass search window (in this case ± 5 ppm) could also impact MF assignment. Specifically, in the *Mtb* LipidDB were 29 instances where a lipid group mass was within ± 5 ppm of another lipid group mass. For example, the mass of the sulfolipid III (SL-III) (C90) lipid group is 0.8 ppm less than the mass of CL (81:1). The issue of mass window overlap was further complicated by the inclusion of multiple molecular ions for each lipid group. For instance, the database mass entry for the diacylated diacylglycerophosphoinositoldimannoside (Ac_2PIM_2) (70:0) $[M+Na-2H]^-$ molecular ion is 1.4 ppm less than the database mass entry for the Ac_2PIM_2 (72:3) $[M-H]^-$ molecular ion. In total there were 154 out of 9,177 possible instances of mass window overlap in the +MH_MtbLipid.csv file, and 244 out of 5,310 possible instances in the -MH_MtbLipid.csv file. Analyses of the *Mtb* crude lipid extract resulted in only 29 positive-ion and 28 negative-ion MFs that matched to two lipid entries. However, the MassHunter software provided a scoring algorithm for each database match based on a combination of ppm accuracy and isotopic abundance pattern fit. This scoring mechanism generally appeared to select the correct lipid group as the top candidate displayed a retention-time in agreement with the respective parent lipid subclass.

Quantitative and detection limit capabilities.

In addition to providing mass and retention time information, the MassHunter MFE algorithm provided a measurement of abundance, termed "volume" for each MF, where volume = [retention time window] x [isotope ion peak heights]. To assess the linearity of this quantitative feature for identified lipid groups, a dilution series of *Mtb* cells was extracted and the total lipids were analyzed by LC/MS as described above. The MF volumes for representative lipid groups yielded a linear relationship to the quantity of lipid over a dynamic range of approximately three logs. Further, the most abundant lipid subclasses (DIM A, DIM B, PI, TG) could be detected in extracts that were equivalent to $<10^6$ *Mtb* cells (Fig. 4).

Assessment of *Mtb* lipid profiles from different growth phases

Triplicate liquid cultures of *Mtb* strain H37Rv were sampled at three time points corresponding to logarithmic phase, stationary phase, and an intermediate growth phase between these two termed "transitionary" phase (Fig. S2). Crude lipid extracts were prepared from the collected cells and subjected to LC/MS analyses. MF volumes were used to establish relative quantities for each identified lipid group, and comparisons revealed growth-phase-dependent differences in lipid class abundance and composition.

Striking differences in both total TG abundance and TG fatty acyl composition were observed across each of the three growth phases. When expressed as a percentage of all identified lipid groups (+ ion data), TG lipid groups accounted for 35.8%, 4.1%, and 77.3% of the ion volumes

in logarithmic, transitionary, and stationary phases, respectively, indicating significant changes in cellular concentrations of TG as the cultures aged (Fig. 5A inset). Large differences in TG group abundances were also observed from the LC/MS chromatograms (Fig. S2). The TG fatty acid composition also varied between growth phases with a shift from predominantly lower mass TG lipid groups in logarithmic phase to higher mass TG groups in stationary phase (Fig. 5A, Fig. S3). The sum compositions indicated that the smaller TGs were likely esterified with C16, C18, and C19 fatty acids, while longer C22, C24, and C26 acyl functions likely esterified the larger, stationary-phase specific TGs. The latter observation fits well with previous structural reports of mycobacterial TG (13, 32-33). To further support this structural information the most abundant logarithmic- and stationary-phase TG group $[M+NH_4]^+$ ions were targeted for collision-induced-dissociation (CID) MS/MS (Fig. 5B). Major diacyl product ions were observed that were consistent with C16:1 and C18:1 being the primary fatty acids of the major logarithmic TG (52:3) group, and C16:0, C18:0, and C26:0 fatty acids esterified to the major stationary-phase specific TG (60:0) group.

Growth-phase-dependent changes in glycerophospholipid (GP) sum compositions were also observed (Fig. S4 A-C). For instance, PE (34:1) was the most abundant PE lipid group in logarithmic phase, while the PE (34:0) lipid group dominated in stationary phase. Interestingly, the sum composition of the PI and PE lipid groups differed significantly from each other even in the same growth phase, indicating that fatty acid distributions were specific to each GP subclass. Regardless of GP subclass, however, each of the PE, PI, and CL subclasses showed a significant increase in the saturated bond content of their fatty acyl moieties as the cultures aged (Fig. 6).

Finally, notable changes were observed within the wax diester subclass. Comparisons of the relative abundances of DIM A lipid groups revealed significant increases in higher-mass DIM A forms as the cultures entered stationary phase (Fig. 7). The average DIM A mass was 1367.2 (±0.6 SD) atomic mass units (amu) in logarithmic phase, 1377.4 (±1.5 SD) amu in transitionary phase, and 1385.9 (±1.8 SD) amu in stationary phase. Comparable levels of mass increases were also observed for DIM B lipid groups (data not shown). In addition to the observed mass increases, the ratio of total DIM A to total DIM B lipid group ion volumes increased as the cultures aged. The total DIM A/B content was 1.31 (±0.08 SD) in logarithmic phase, 1.85 (±0.16 SD) in transitionary phase, and 3.43 (±0.40 SD) in stationary phase.

DISCUSSION

The global characterization of the *Mtb* lipid composition presents a significant technical challenge due to the structural diversity within this group of macromolecules and the number of complex structures that are unique to *Mtb* and related species. However, based on the efforts of earlier investigators to fully elucidate the structure of a large majority of these products, it was possible to derive a database of *Mtb* lipids that possessed 2,512 structures and that could be

interrogated with mass spectrometry data based on the 14,489 mass entries that accounted for multiple ion adducts of both positive and negative-ion data. The interrogation of this database with mass spectrometry data of whole lipid extracts from *Mtb* allowed for the automated detection and identification of 415 lipid structures encompassing all 15 of the known classes of *Mtb* lipids. This work represents the largest single survey of *Mtb* lipid structures from a single study and provides the ground work for lipidomics studies that are on par with those now being performed with eukaryotic organisms (34).

The *Mtb* LipidDB was developed to allow lipidomic investigations in an unbiased manner. Thus, lipid group structures and masses were obtained by calculating all potential combinations of the fatty acids reported in the literature for a specific lipid. This approach resulted in database inclusion of lipid structures containing fatty acyl combinations not previously reported, but that are possible based on the biochemical literature. This expansion of the database was most noted with the mycobacterial lipids commonly referred to as phosphatidylinositolmannosides (PIMs), where 1,037 of the 2,512 lipid structures in the *Mtb* LipidDB were assigned to a PIM structure. This reflects the variable composition of the PIM structures with one to six mannose residues and one to four acyl moieties (35). Of the 1,037 potential PIM structures included in the *Mtb* LipidDB, 77 were identified in our analyses. This can be directly compared to other MS-based studies of PIMs isolated from *M. bovis* BCG (36-37), where 73 of the possible 1,037 PIM groups included in our database were identified. The inclusion of theoretical lipid structures based on sound biochemical evidence, and more general descriptions in the literature of acyl composition is not uncommon. In fact in eukaryotes it is estimated that there are at least 25,000 TG structures, including those with ether linked fatty acids (38).

Beyond the ability to identify individual lipid groups or species, mass spectrometry offers the potential to perform relative quantitative analyses based on the total ion abundance of individual molecular features (39). This approach was used to monitor lipid changes over different phases of *Mtb* growth in batch culture. TG abundance waned and waxed dramatically as the cultures aged, and TG fatty acid sum compositions also shifted markedly, with the appearance of higher molecular weight TG groups in stationary phase. For the purposes of this study, the stationary phase was the point at which the total bacterial growth rate ceased as measured by optical density. Such tightly-capped batch cultures grown to high cellular densities have been suggested to be depleted of oxygen (40), thus the observed accumulation of TG in stationary phase would agree with previous reports demonstrating TG production increases under hypoxic and/or stressful conditions (12-14, 16, 33). Structural analyses suggested C26:0 fatty acids were major components of these TG groups, also in agreement with previous studies on hypoxia-induced TG molecules (13-14, 33). Alterations in fatty acid composition were not limited to the TGs, as the PE, PI, and CL lipid classes also demonstrated similar changes that were class specific. However, common to each class was an observed increase in overall saturation levels as the cultures aged. It should be noted that the presence of polysorbate (Tween 80) in the media may have influenced fatty acyl compositions, as this detergent supplies oleic acid esters for growth.

The average size of DIM A and DIM B lipids was also shown to increase as the cultures aged. DIM size increases were previously observed in *Mtb* cells cultivated on odd-chain carbon sources and cholesterol, and also in *Mtb* cells from infected mice lungs (15, 20). These mass increases were shown to result from addition of methylene units to the esterified mycocerosic acids. These data build on evidence that suggest *Mtb* actively modulates its lipid composition in response to the changing microenvironment. Thus, this lipidomic approach provides a working platform for the analyses of more refined *in vitro* growth culture models that incorporate defined stress factors (41-48) (19), as well as comparative analyses of lipid profiles from *Mtb* cells isolated from infected tissues. Such *in vivo* studies would provide valuable insight into the metabolism and physiology of *Mtb*. Additionally, combining the 2D HSQC NMR approach applied by Lee and colleagues (22) would potentially provide synergistic methods to produce complementary levels of information. A current limitation to the quantitative lipid profiling is that the true stoichiometric relationships between lipid classes cannot be determined. Even differences between lipid groups belonging to the same lipid subclass could not be used to determine true abundance, as acyl chain length and degree of unsaturation are known to influence ionization responses (49). Absolute quantification could be achieved with the incorporation of multiple internal lipid standards to which individual lipid ionization could be normalized. Such approaches are becoming more commonplace for eukaryotic-based lipidomics studies (34). However, appropriate isotopically labeled standards specific for many of the unique *Mtb* lipids are lacking, and require development to achieve this level of analysis.

A central element to our MS based lipidomics strategy was automated product identification based on accurate-mass matching. This approach has become more commonplace for metabolomics-based experiments that utilize high-resolution mass spectrometers (50-51). For global lipid characterization, however, this methodology can be complicated by differential acylation of the same lipid resulting in isobaric lipid species. Thus, the *Mtb* LipidDB was designed such that the output returned only fatty acid sum composition, thus limiting potential mass overlap between database entries. In this study we demonstrated a relatively small number of instances where mass overlap (± 5 ppm) occurs between lipid groups in the *Mtb* LipidDB. Indeed, it may be advantageous that *Mtb* does not produce appreciable levels of phosphatidic acid, phosphatidylserine or phosphatidylcholine, as these would overlap in mass with other *Mtb* phospholipids as observed in eukaryotes (52). Furthermore, we note that the unique large, apolar molecular characteristics of many *Mtb* lipids provide distinguishing mass features for these molecules in a eukaryotic lipid background (unpublished observations). We also view the automated identification process described in this study as a screening method, and believe that database identifications based on accurate mass alone should be regarded as tentative. Lipid identifications of biological interest should be subjected to MS/MS characterization and structural confirmation, as was done with select TG groups in this study. Interpretation of the resulting data can be aided by well-documented *Mtb* lipid collision-induced-dissociation fragmentation patterns. Indeed we have obtained confirmatory MS/MS product ion spectra for representatives of many of the lipid subclasses identified in this study (data not shown).

Identification confidence has also been shown to be enhanced with the incorporation of retention time criteria into the database searching process in the presence of internal standards that allow for normalization of retention times (53-56). Thus, performing MS/MS analyses in a spot-check fashion would allow for further retention time curation of the searchable database files, and this could easily be developed through an iterative process within a laboratory. Finally, the accuracy of the database identifications will also improve with the increasing resolution of newly developed mass spectrometers.

The limitations of the current *Mtb* LipidDB also need to be recognized. Firstly, the database is constrained by the reported lists of fatty acids considered for the various lipid group repertoires. For example, only the fatty acids observed by Walker et al. (32) were used to calculate GP lipid group masses, yet other atypical fatty acids may be esterified to mycobacterial GPs and would thus escape identification from the database search. The *Mtb* LipidDB does not take into consideration mass shifts due to oxidation or degradation of lipids, which may be keys to a complete understanding of *Mtb* lipid metabolism. Only well defined lipids were included in the database, a novel and recently described wax ester lipid identified from *Mtb* cultivated under iron-limiting conditions was not included because of its partial structural elucidation (48). Biosynthetic lipid intermediates, such as phosphatidic acid, acyl-coenzyme A, and sugar-linked decaprenol phosphates, were also not included in the database. Finally, the database in general is restricted to the lipid composition of the *Mtb* H37Rv strain used in this study, and does not contain lipids found in other mycobacterial species. Keeping these limitations in mind, the flexible data format of the database should allow users to easily amend and add lipid entries according to their specific research interests.

This work was funded by the Bill & Melinda Gates Foundation TB Drug Accelerator Grant # 42755

REFERENCES

1. Brennan, P. J., and H. Nikaido. 1995. The envelope of mycobacteria. *Annu Rev Biochem* **64**: 29-63.
2. Daffe, M., and P. Draper. 1998. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39**: 131-203.
3. Crick, D. C., P. J. Brennan, and M. R. Mcneil. 2004. The Cell Wall of *Mycobacterium Tuberculosis*. *In* Tuberculosis, Second Edition. W. N. Rom and S. M. Garay, editors. Lippincott Williams & Wilkins, Philadelphia. 115-134.
4. Cox, J. S., B. Chen, M. McNeil, and W. R. Jacobs, Jr. 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402**: 79-83.
5. Reed, M. B., P. Domenech, C. Manca, H. Su, A. K. Barczak, B. N. Kreiswirth, G. Kaplan, and C. E. Barry, 3rd. 2004. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431**: 84-87.
6. Rao, V., N. Fujiwara, S. A. Porcelli, and M. S. Glickman. 2005. *Mycobacterium tuberculosis* controls host innate immune activation through cyclopropane modification of a glycolipid effector molecule. *J Exp Med* **201**: 535-543.
7. Goren, M. B., O. Brokl, and W. B. Schaefer. 1974. Lipids of putative relevance to virulence in *Mycobacterium tuberculosis*: correlation of virulence with elaboration of sulfatides and strongly acidic lipids. *Infect Immun* **9**: 142-149.
8. Daffe, M., and M. A. Laneelle. 1988. Distribution of phthiocerol diester, phenolic mycosides and related compounds in mycobacteria. *J Gen Microbiol* **134**: 2049-2055.
9. Daffe, M., M. McNeil, and P. J. Brennan. 1991. Novel type-specific lipooligosaccharides from *Mycobacterium tuberculosis*. *Biochemistry* **30**: 378-388.
10. Watanabe, M., Y. Aoyagi, M. Ridell, and D. E. Minnikin. 2001. Separation and characterization of individual mycolic acids in representative mycobacteria. *Microbiology* **147**: 1825-1837.
11. Watanabe, M., Y. Aoyagi, H. Mitome, T. Fujita, H. Naoki, M. Ridell, and D. E. Minnikin. 2002. Location of functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. *Microbiology* **148**: 1881-1902.
12. Garton, N. J., H. Christensen, D. E. Minnikin, R. A. Adegbola, and M. R. Barer. 2002. Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum. *Microbiology* **148**: 2951-2958.
13. Daniel, J., C. Deb, V. S. Dubey, T. D. Sirakova, B. Abomoelak, H. R. Morbidoni, and P. E. Kolattukudy. 2004. Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in *Mycobacterium tuberculosis* as it goes into a dormancy-like state in culture. *J Bacteriol* **186**: 5017-5030.
14. Sirakova, T. D., V. S. Dubey, C. Deb, J. Daniel, T. A. Korotkova, B. Abomoelak, and P. E. Kolattukudy. 2006. Identification of a diacylglycerol acyltransferase gene involved in accumulation of triacylglycerol in *Mycobacterium tuberculosis* under stress. *Microbiology* **152**: 2717-2725.
15. Jain, M., C. J. Petzold, M. W. Schelle, M. D. Leavell, J. D. Mougous, C. R. Bertozzi, J. A. Leary, and J. S. Cox. 2007. Lipidomics reveals control of *Mycobacterium tuberculosis* virulence lipids via metabolic coupling. *Proc Natl Acad Sci U S A* **104**: 5133-5138.
16. Garton, N. J., S. J. Waddell, A. L. Sherratt, S. M. Lee, R. J. Smith, C. Senner, J. Hinds, K. Rajakumar, R. A. Adegbola, G. S. Besra, P. D. Butcher, and M. R. Barer. 2008. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med* **5**: e75.
17. Matsunaga, I., T. Naka, R. S. Talekar, M. J. McConnell, K. Katoh, H. Nakao, A. Otsuka, S. M. Behar, I. Yano, D. B. Moody, and M. Sugita. 2008. Mycolyltransferase-mediated glycolipid exchange in Mycobacteria. *J Biol Chem* **283**: 28835-28841.

18. Ojha, A. K., A. D. Baughn, D. Sambandan, T. Hsu, X. Trivelli, Y. Guerardel, A. Alahari, L. Kremer, W. R. Jacobs, Jr., and G. F. Hatfull. 2008. Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* **69**: 164-174.
19. Deb, C., C. M. Lee, V. S. Dubey, J. Daniel, B. Abomoelak, T. D. Sirakova, S. Pawar, L. Rogers, and P. E. Kolattukudy. 2009. A novel in vitro multiple-stress dormancy model for *Mycobacterium tuberculosis* generates a lipid-loaded, drug-tolerant, dormant pathogen. *PLoS One* **4**: e6077.
20. Yang, X., N. M. Nesbitt, E. Dubnau, I. Smith, and N. S. Sampson. 2009. Cholesterol metabolism increases the metabolic pool of propionate in *Mycobacterium tuberculosis*. *Biochemistry* **48**: 3819-3821.
21. Dobson, G., D. E. Minnikin, S. M. Minnikin, M. Parlett, M. Goodfellow, M. Ridell, and M. Magnusson. 1985. Systematic analysis of complex mycobacterial lipids. *In* Chemical methods in bacterial systematics. M. Goodfellow and D. E. Minnikin, editors. Academic Press, London. 237-265.
22. Mahrous, E. A., R. B. Lee, and R. E. Lee. 2008. A rapid approach to lipid profiling of mycobacteria using 2D HSQC NMR maps. *J Lipid Res* **49**: 455-463.
23. Leavell, M. D., and J. A. Leary. 2006. Fatty acid analysis tool (FAAT): An FT-ICR MS lipid analysis algorithm. *Anal Chem* **78**: 5497-5503.
24. Shui, G., A. K. Bendt, K. Pethe, T. Dick, and M. R. Wenk. 2007. Sensitive profiling of chemically diverse bioactive lipids. *J Lipid Res* **48**: 1976-1984.
25. Takayama, K., H. K. Schnoes, E. L. Armstrong, and R. W. Boyle. 1975. Site of inhibitory action of isoniazid in the synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J Lipid Res* **16**: 308-317.
26. Sonnenberg, M. G., and J. T. Belisle. 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect Immun* **65**: 4515-4524.
27. Covert, B. A., J. S. Spencer, I. M. Orme, and J. T. Belisle. 2001. The application of proteomics in defining the T cell antigens of *Mycobacterium tuberculosis*. *Proteomics* **1**: 574-586.
28. Bligh, E. G., and W. J. Dyer. 1959. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**: 911-917.
29. Munoz, M., M. A. Laneelle, M. Luquin, J. Torrelles, E. Julian, V. Ausina, and M. Daffe. 1997. Occurrence of an antigenic triacyl trehalose in clinical isolates and reference strains of *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* **157**: 251-259.
30. Fahy, E., S. Subramaniam, H. A. Brown, C. K. Glass, A. H. Merrill, Jr., R. C. Murphy, C. R. Raetz, D. W. Russell, Y. Seyama, W. Shaw, T. Shimizu, F. Spener, G. van Meer, M. S. VanNieuwenhze, S. H. White, J. L. Witztum, and E. A. Dennis. 2005. A comprehensive classification system for lipids. *J Lipid Res* **46**: 839-861.
31. Fahy, E., S. Subramaniam, R. C. Murphy, M. Nishijima, C. R. Raetz, T. Shimizu, F. Spener, G. van Meer, M. J. Wakelam, and E. A. Dennis. 2009. Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res* **50 Suppl**: S9-14.
32. Walker, R. W., H. Barakat, and J. G. Hung. 1970. The positional distribution of fatty acids in the phospholipids and triglycerides of *Mycobacterium smegmatis* and *M. bovis* BCG. *Lipids* **5**: 684-691.
33. Low, K. L., P. S. Rao, G. Shui, A. K. Bendt, K. Pethe, T. Dick, and M. R. Wenk. 2009. Triacylglycerol utilization is required for regrowth of in vitro hypoxic nonreplicating *Mycobacterium bovis* bacillus Calmette-Guerin. *J Bacteriol* **191**: 5037-5043.
34. Ejsing, C. S., J. L. Sampaio, V. Surendranath, E. Duchoslav, K. Ekroos, R. W. Klemm, K. Simons, and A. Shevchenko. 2009. Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. *Proc Natl Acad Sci U S A* **106**: 2136-2141.

35. Nigou, J., M. Gilleron, T. Brando, and G. Puzo. 2004. Structural analysis of mycobacterial lipoglycans. *Appl Biochem Biotechnol* **118**: 253-267.
36. Hsu, F. F., J. Turk, R. M. Owens, E. R. Rhoades, and D. G. Russell. 2007. Structural characterization of phosphatidyl-myo-inositol mannosides from *Mycobacterium bovis* Bacillus Calmette Guerin by multiple-stage quadrupole ion-trap mass spectrometry with electrospray ionization. I. PIMs and lyso-PIMs. *J Am Soc Mass Spectrom* **18**: 466-478.
37. Hsu, F. F., J. Turk, R. M. Owens, E. R. Rhoades, and D. G. Russell. 2007. Structural characterization of phosphatidyl-myo-inositol mannosides from *Mycobacterium bovis* Bacillus Calmette Guerin by multiple-stage quadrupole ion-trap mass spectrometry with electrospray ionization. II. Monoacyl- and diacyl-PIMs. *J Am Soc Mass Spectrom* **18**: 479-492.
38. McAnoy, A. M., C. C. Wu, and R. C. Murphy. 2005. Direct qualitative analysis of triacylglycerols by electrospray mass spectrometry using a linear ion trap. *J Am Soc Mass Spectrom* **16**: 1498-1509.
39. Sandra, K., S. Pereira Ados, G. Vanhoenacker, F. David, and P. Sandra. 2010. Comprehensive blood plasma lipidomics by liquid chromatography/quadrupole time-of-flight mass spectrometry. *J Chromatogr A* **1217**: 4087-4099.
40. Wayne, L. G. 2001. In Vitro Model of Hypoxically Induced Nonreplicating Persistence of *Mycobacterium tuberculosis*. *In Mycobacterium tuberculosis* Protocols. T. Parish and N. G. Stoker, editors. Humana Press, Totowa, NJ. 247-268.
41. Wayne, L. G., and L. G. Hayes. 1996. An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun* **64**: 2062-2069.
42. Betts, J. C., P. T. Lukey, L. C. Robb, R. A. McAdam, and K. Duncan. 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43**: 717-731.
43. Fisher, M. A., B. B. Plikaytis, and T. M. Shinnick. 2002. Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J Bacteriol* **184**: 4025-4032.
44. Hampshire, T., S. Soneji, J. Bacon, B. W. James, J. Hinds, K. Laing, R. A. Stabler, P. D. Marsh, and P. D. Butcher. 2004. Stationary phase gene expression of *Mycobacterium tuberculosis* following a progressive nutrient depletion: a model for persistent organisms? *Tuberculosis (Edinb)* **84**: 228-238.
45. Muttucumaru, D. G., G. Roberts, J. Hinds, R. A. Stabler, and T. Parish. 2004. Gene expression profile of *Mycobacterium tuberculosis* in a non-replicating state. *Tuberculosis (Edinb)* **84**: 239-246.
46. Voskuil, M. I. 2004. *Mycobacterium tuberculosis* gene expression during environmental conditions associated with latency. *Tuberculosis (Edinb)* **84**: 138-143.
47. Voskuil, M. I., K. C. Visconti, and G. K. Schoolnik. 2004. *Mycobacterium tuberculosis* gene expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis (Edinb)* **84**: 218-227.
48. Bacon, J., L. G. Dover, K. A. Hatch, Y. Zhang, J. M. Gomes, S. Kendall, L. Wernisch, N. G. Stoker, P. D. Butcher, G. S. Besra, and P. D. Marsh. 2007. Lipid composition and transcriptional response of *Mycobacterium tuberculosis* grown under iron-limitation in continuous culture: identification of a novel wax ester. *Microbiology* **153**: 1435-1444.
49. Koivusalo, M., P. Haimi, L. Heikinheimo, R. Kostiainen, and P. Somerharju. 2001. Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response. *J Lipid Res* **42**: 663-672.
50. Ding, J., C. M. Sorensen, Q. Zhang, H. Jiang, N. Jaitly, E. A. Livesay, Y. Shen, R. D. Smith, and T. O. Metz. 2007. Capillary LC coupled with high-mass measurement accuracy mass spectrometry for metabolic profiling. *Anal Chem* **79**: 6081-6093.

51. Sana, T. R., J. C. Roark, X. Li, K. Waddell, and S. M. Fischer. 2008. Molecular formula and METLIN Personal Metabolite Database matching applied to the identification of compounds generated by LC/TOF-MS. *J Biomol Tech* **19**: 258-266.
52. Schwudke, D., J. T. Hannich, V. Surendranath, V. Grimard, T. Moehring, L. Burton, T. Kurzchalia, and A. Shevchenko. 2007. Top-down lipidomic screens by multivariate analysis of high-resolution survey mass spectra. *Anal Chem* **79**: 4083-4093.
53. Kurvinen, J. P., J. Aaltonen, A. Kuksis, and H. Kallio. 2002. Software algorithm for automatic interpretation of mass spectra of glycerolipids. *Rapid Commun Mass Spectrom* **16**: 1812-1820.
54. Haimi, P., A. Uphoff, M. Hermansson, and P. Somerharju. 2006. Software tools for analysis of mass spectrometric lipidome data. *Anal Chem* **78**: 8324-8331.
55. Yetukuri, L., M. Katajamaa, G. Medina-Gomez, T. Seppanen-Laakso, A. Vidal-Puig, and M. Oresic. 2007. Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *BMC Syst Biol* **1**: 12.
56. Ding, J., C. M. Sorensen, N. Jaitly, H. Jiang, D. J. Orton, M. E. Monroe, R. J. Moore, R. D. Smith, and T. O. Metz. 2008. Application of the accurate mass and time tag approach in studies of the human blood lipidome. *J Chromatogr B Analyt Technol Biomed Life Sci* **871**: 243-252.

FIGURE LEGENDS

Fig 1. Organizational summary of the *Mtb* LipidDB and LIPID MAPS classification systems.

Fig 2. Identified molecular features from positive ion LC/MS analyses of an *Mtb* total lipid extract. 672 molecular features ions were matched to lipid groups representing 25 lipid subclasses from the *Mtb* LipidDB. ◆ hydroxyphthioceranic acid, ◆ Alpha MA, ▲ Methoxy MA, ✕ Keto MA, ✱ GMM, + DIMA, + DIMB, − PE, ◆ Lyso-PE, ■ GP, ▲ PI, ✕ Lyso-PI, ✱ CL, - Ac_1PIM_1, ■ PIM_2, ■ Ac_1PIM_2, ▲ Ac_2PIM_2, ✕ Ac_1PIM_3, ✱ Ac_2PIM_3, + Ac_1PIM_6, − Ac_2PIM_6, − MG, ◆ DG, ◆ TG, ▲ MK, ✕ DP, ✱ DAT_1, ● DAT_2, ✕ PAT, ■ Mbt w/ Fe, ▲ Mbt –Fe. See Supplementary Material for lipid group abbreviations.

Fig 3. Identified molecular features from negative ion LC/MS analyses of an *Mtb* total lipid extract. 244 molecular features ions were matched to lipid groups representing 24 lipid subclasses from the *Mtb* LipidDB. ◆ hydroxyphthioceranic acid, ◆ Alpha MA, ▲ Methoxy MA, ✕ Keto MA, ✱ GMM, ● TMM, − PE, ■ GP, ▲ PI, ✕ Lyso-PI, ✱ CL, ● Lyso-PIM_1, + PIM_1, - Ac_1PIM_1, − Lyso-PIM_2, ■ PIM_2, ■ Ac_1PIM_2, ▲ Ac_2PIM_2, ✕ Ac_1PIM_3, ✱ Ac_2PIM_3, ● Ac_1PIM_4, + Ac_1PIM_6, − Ac_2PIM_6, ▲ MK, ✱ DAT_1, ● DAT_2, ✕ PAT, + SL-II or SL-II', − SL-III, − Ac_2SGL, ◆ MPM. See Supplementary Material for lipid group abbreviations.

Fig 4. Ion response scatter plots for an *Mtb* total lipid extract dilution series. The observed ion volumes were selected from representative lipid group ions detected across a wide dynamic range. Values were plotted against the amount of lipid extract injected (bottom x-axis) and the equivalent number of extracted *Mtb* cells (top x-axis).

Fig 5. Comparison of TG profiles from *Mtb* growth phases. A: TG lipid group ion volumes with the same carbon number were combined and normalized to total TG ion volume (for individual TG lipid group profiles see Fig. S4). Error bars indicate ± SD (n = 3). The inset shows relative TG abundance as % of total identified lipid group ion volumes. B: MS/MS spectra of the $[M+NH_4]^+$ molecular ions for representative TG groups. The major logarithmic-phase TG (52:3) ion produced two major diacyl product ions at *m/z* 575.5010 and *m/z* 603.5308 that resulted from neutral losses of C18:1 and C16:1 fatty acids, respectively. The major stationary-phase specific TG (60:0) ion produced three major diacyl product ions at *m/z* 579.5345, *m/z* 691.6605, and *m/z* 719.6682 which indicated neutral losses of C26:0, C18:0, and C16:0 fatty acids, respectively.

Fig 6. Comparison of GP subclass unsaturated bond content from *Mtb* growth phases. Percent of each lipid subclass containing lipid groups with fatty acyl sum composition possessing 0-4 unsaturated bonds. (LOG: logarithmic phase; TRANS: transitionary phase; STAT: stationary phase).

Fig 7. Growth-phase dependent changes in DIM A molecular size. DIM A group ion volumes were normalized to total DIM A ion volumes. Error bars indicate ± SD (n = 3).

Table I. Summary of the Mtb Lipid Database

- Fatty Acyls (FA) [FA] {745}
  - Fatty Acids and Conjugates [FA01] {641}
    - Branched Fatty Acids [FA0102] {41}
    - Mycolic Acids (MA) [FA0116] {600}
      - Alpha Mycolic Acids (Alpha-MA) [FA0116z] {21}
      - Methoxy Mycolic Acids (Methoxy-MA) [FA0116z] {21}
      - Keto Mycolic Acids (Keto-MA) [FA0116z] {21}
      - Glucose Monomycolates (GMM) [FA0116z] {84}
      - Trehalose Monomycolates (TMM) [FA0116z] {84}
      - Trehalose Dimycolates (TDM) [FA0116z] {369}
  - Fatty Esters [FA07] {104}
    - Wax diesters [FA0707] {52}
      - Phthiocerol Dimycocerosates (DIMA) [FA0707z] {26}
      - Phthiodiolone Dimycocerosates (DIMB) [FA0707z] {26}
    - Glycosylated wax diesters (GWE) [FA07y] {52}
      - Glycosylated Phthiocerol Dimycocerosates (PGL-tb) [FA07yz] {26}
      - Glycosylated Phthiodiolone Dimycocerosates [FA07yz] {26}
- Glycerophospholipids (GP) [GP] {1295}
  - Glycerophosphoethanolamines (PE) [GP02] {36}
    - Diacylglycerolphosphoethanolamines (PE) [GP0201] {27}
    - Monoacylglycerolphosphoethanolamines (Lyso-PE) [GP0205] {9}
  - Glycerophosphoglycerols (PG) [GP04] {50}
    - Diacylglycerophosphoglycerols (PG) [GP0401] {40}
    - Monoacylglycerophosphoglycerols (Lyso-PG) [GP0405] {10}
  - Glycerophosphoinositols (PI) [GP06] {34}
    - Diacylglycerophosphoinositols (PI) [GP0601] {26}
    - Monoacylglycerophosphoinositols (Lyso-PI) [GP0605] {8}
  - Glycerophosphoglycerophosphoglycerols (CL) [GP12] {138}
    - Diacylglycerophosphoglycerophosphodiradylglycerols (CL) [GP1201] {138}
  - Glycerophosphoinositolglycans [GP15] {1037}
    - Monoacylglycerophosphoinositolmonomannosides (Lyso-PIM1) [GP15y] {8}
    - Diacylglycerophosphoinositolmonomannosides (PIM1) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositolmonomannosides (Ac1PIM1) [GP15y] {63}
    - Monoacylglycerophosphoinositoldimannosides (Lyso-PIM2) [GP15y] {8}
    - Diacylglycerophosphoinositoldimannosides (PIM2) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositoldimannosides (Ac1PIM2) [GP15y] {63}
    - Diacylated diacylglycerophosphoinositoldimannosides (Ac2PIM2) [GP15y] {91}
    - Monoacylglycerophosphoinositoltrimannosides (Lyso-PIM3) [GP15y] {8}
    - Diacylglycerophosphoinositoltrimannosides (PIM3) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositoltrimannosides (Ac1PIM3) [GP15y] {63}
    - Diacylated diacylglycerophosphoinositoltrimannosides (Ac2PIM3) [GP15y] {91}
    - Monoacylglycerophosphoinositoltetramannosides (Lyso-PIM4) [GP15y] {8}
    - Diacylglycerophosphoinositoltetramannosides (PIM4) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositoltetramannosides (Ac1PIM4) [GP15y] {63}
    - Diacylated diacylglycerophosphoinositoltetramannosides (Ac2PIM4) [GP15y] {91}
    - Monoacylglycerophosphoinositolpentamannosides (Lyso-PIM5) [GP15y] {8}
    - Diacylglycerophosphoinositolpentamannosides (PIM5) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositolpentamannosides (Ac1PIM5) [GP15y] {63}
    - Diacylated diacylglycerophosphoinositolpentamannosides (Ac2PIM5) [GP15y] {91}
    - Monoacylglycerophosphoinositolhexamannosides (Lyso-PIM6) [GP15y] {8}
    - Diacylglycerophosphoinositolhexamannosides (PIM6) [GP15y] {26}
    - Monoacylated diacylglycerophosphoinositolhexamannosides (Ac1PIM6) [GP15y] {63}
    - Diacylated diacylglycerophosphoinositolhexamannosides (Ac2PIM6) [GP15y] {91}
- Glycerolipids (GL) [GL] {218}
  - Monoradylglycerols [GL01] {18}
    - Monoacylglycerols (MG) [GL0101] {18}
  - Diradylglycerols [GL02] {61}
    - Diacylglycerols (DG) [GL0201] {61}
  - Triradylglycerols [GL03] {139}
    - Triacylglycerols (TG) [GL0301] {139}
- Prenol Lipids (PR) [PR] {7}
  - Quinones and Hydroquinones [PR02] {4}
    - Ubiquinones [PR0201] {4}
  - Polyprenols [PR03] {3}
    - Bactoprenols [PR0301] {1}
    - Bactoprenol monophosphates [PR0302] {1}
    - Bactoprenol diphosphates [PR0303] {1}
- Saccharolipids (SL) [SL] {226}
  - Acyltrehaloses [SL03] {226}
    - Diacyltrehaloses (DAT) [SL03y] {13}
      - 2,3-di-O-acyltrehaloses (DAT1) [SL03yz] {6}
      - 2,3-di-O-acyltrehaloses (DAT2) [SL03yz] {7}
    - Polyacyltrehaloses (PAT) [SL03y] {50}
    - Sulfolipids (SL) [SL03y] {163}
      - Sulfolipid I (SL-I) [SL03yz] {52}
      - Sulfolipid II or II' (SL-II or SL-II') [SL03yz] {52}
      - Sulfolipid III (SL-III) [SL03yz] {35}
      - Diacylated Sulfolipid (Ac2SGL) [SL03yz] {24}
- Polyketides (PK) [PK] {21}
  - Linear polyketides [PK01] {5}
    - Mannosyl-b1-phosphomycoketides (MPM) [PK01y] {5}
  - Non-ribosomal peptides/polyketide hybrids [PK14] {16}
    - Mycobactins (Mbt) [PK14y] {16}

LEGEND

- Category (Abbreviation) [LIPID MAPs Identifier] {# of lipid groups in category}
  - Main Class (Abbreviation) [LIPID MAPs Identifier] {# of lipid groups in Main Class}
    - Subclass (Abbreviation) [LIPID MAPs Identifier, y denotes new proposed Subclass] {# of lipid groups in Subclass}
      - Level 4 class (Abbreviation) [LIPID MAPs Identifier, z denotes new proposed Level 4 class] {# of lipid groups in Level 4 class}

Figure 1

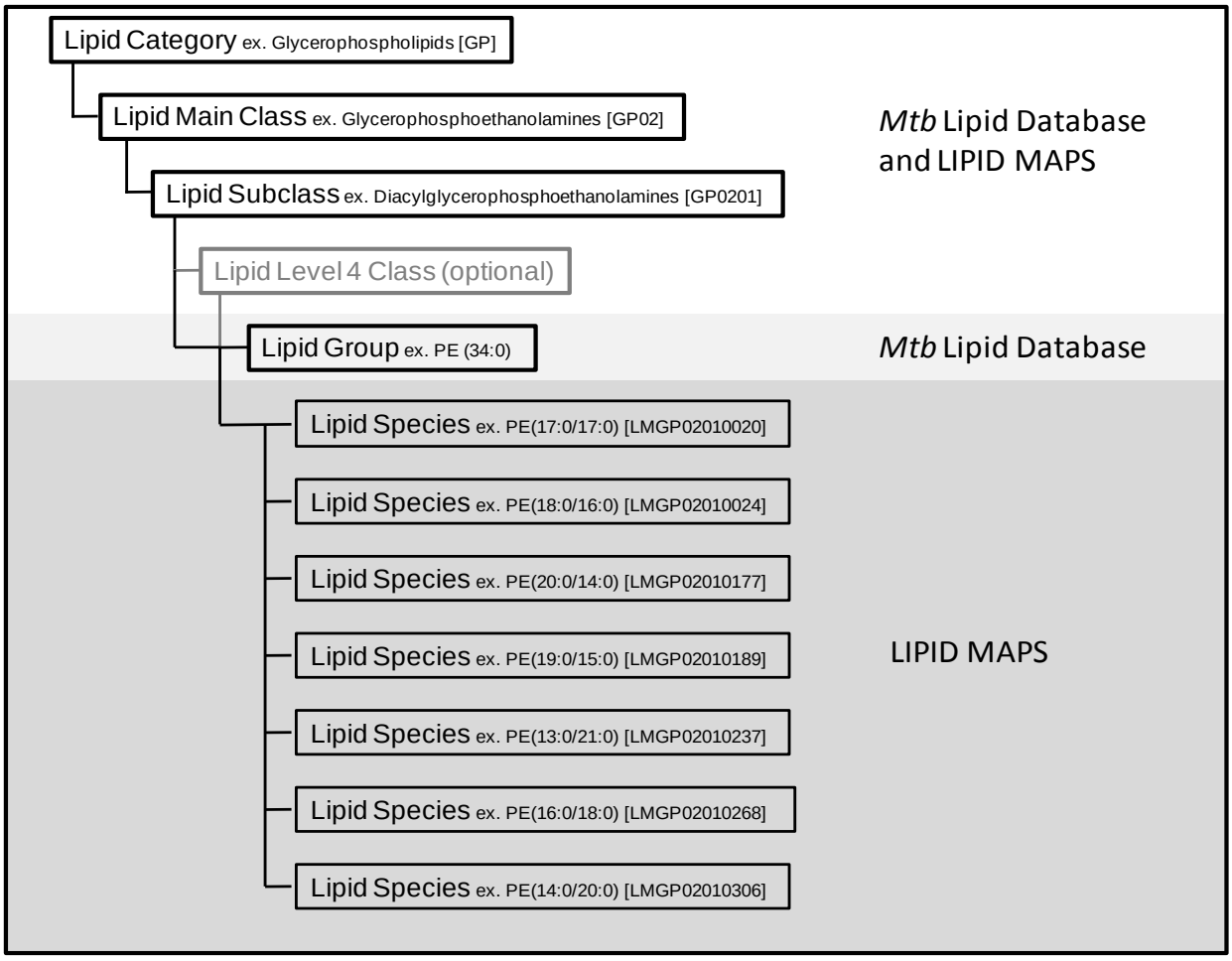

Figure 2

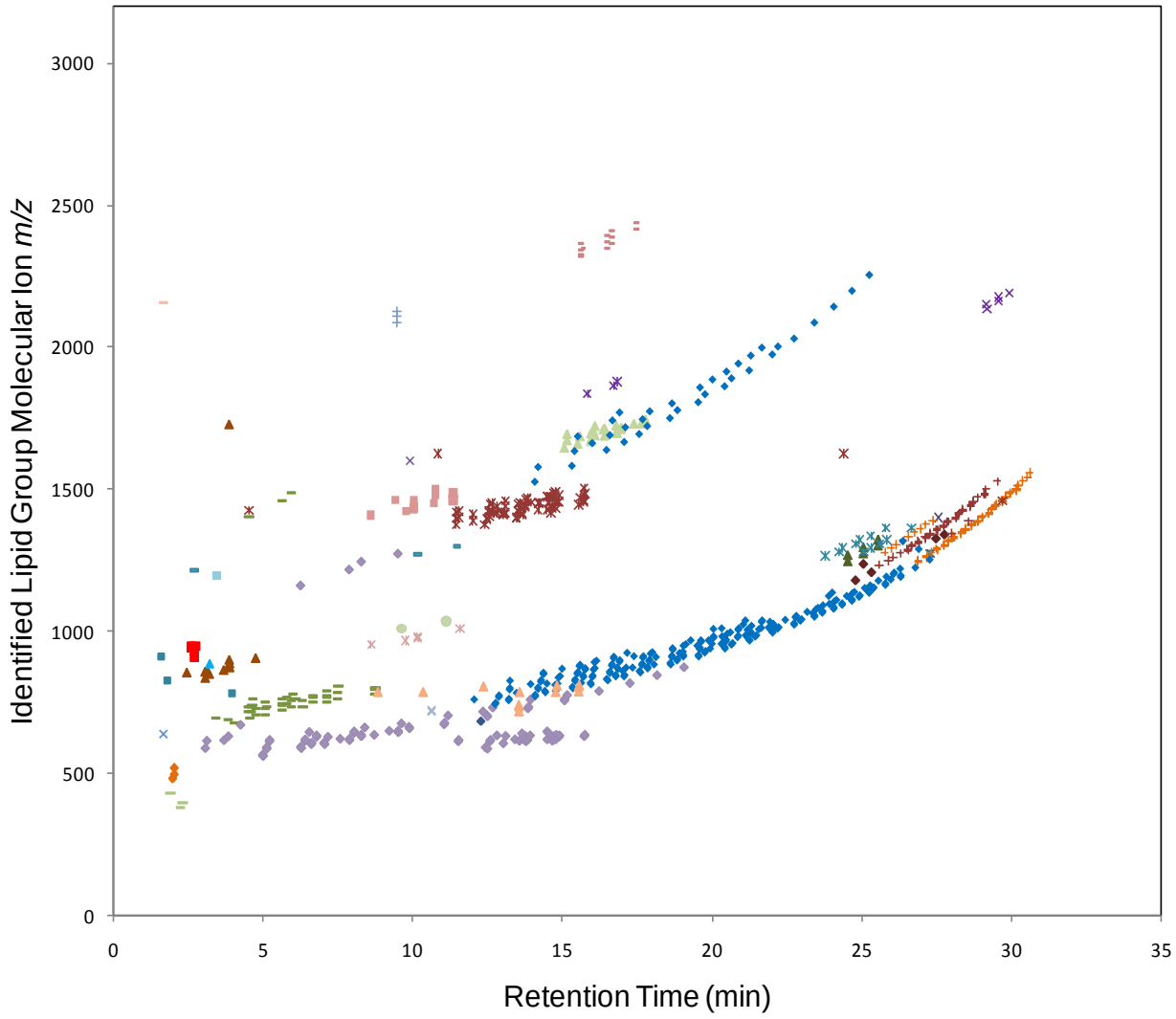

Figure 3

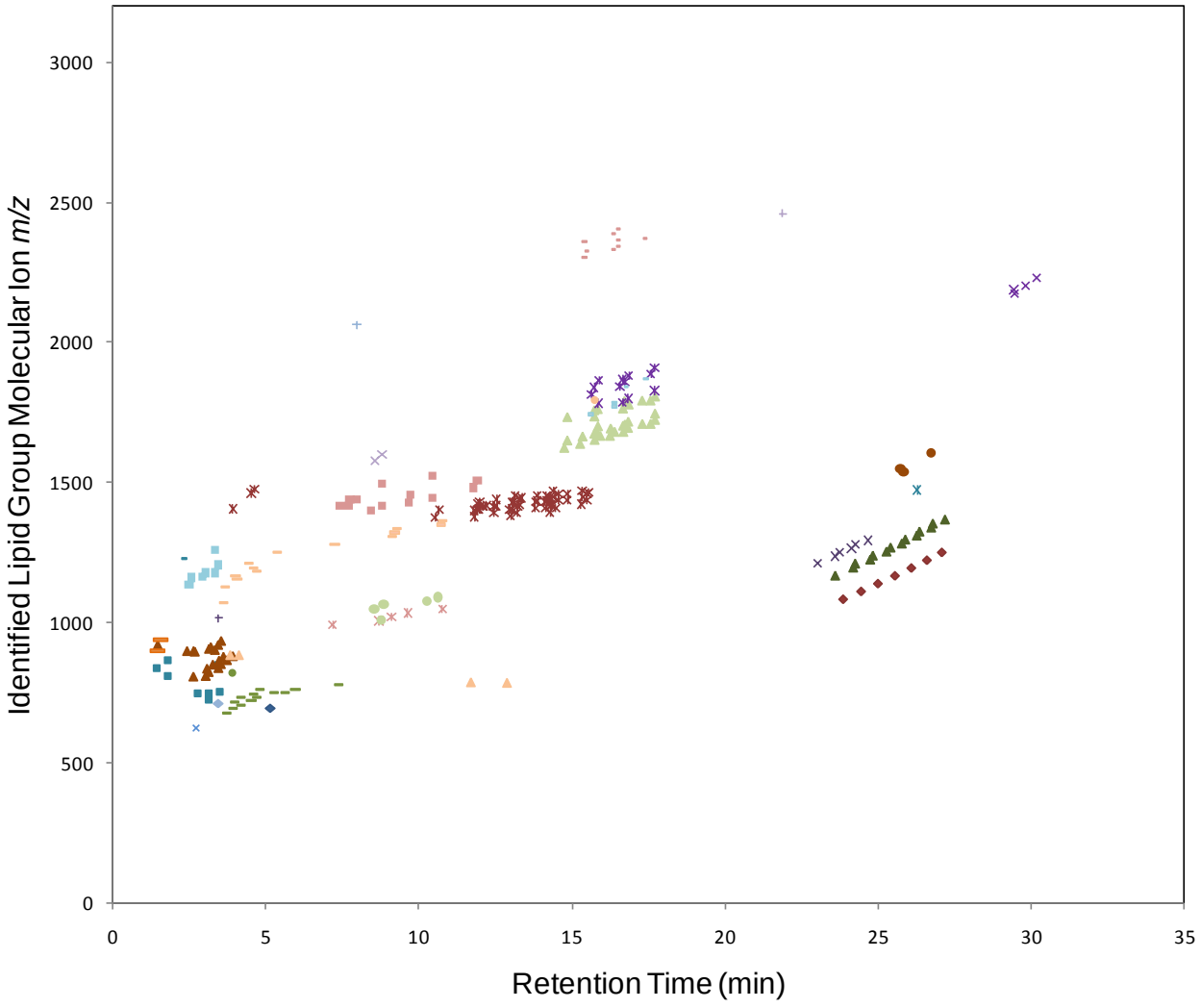

Figure 4

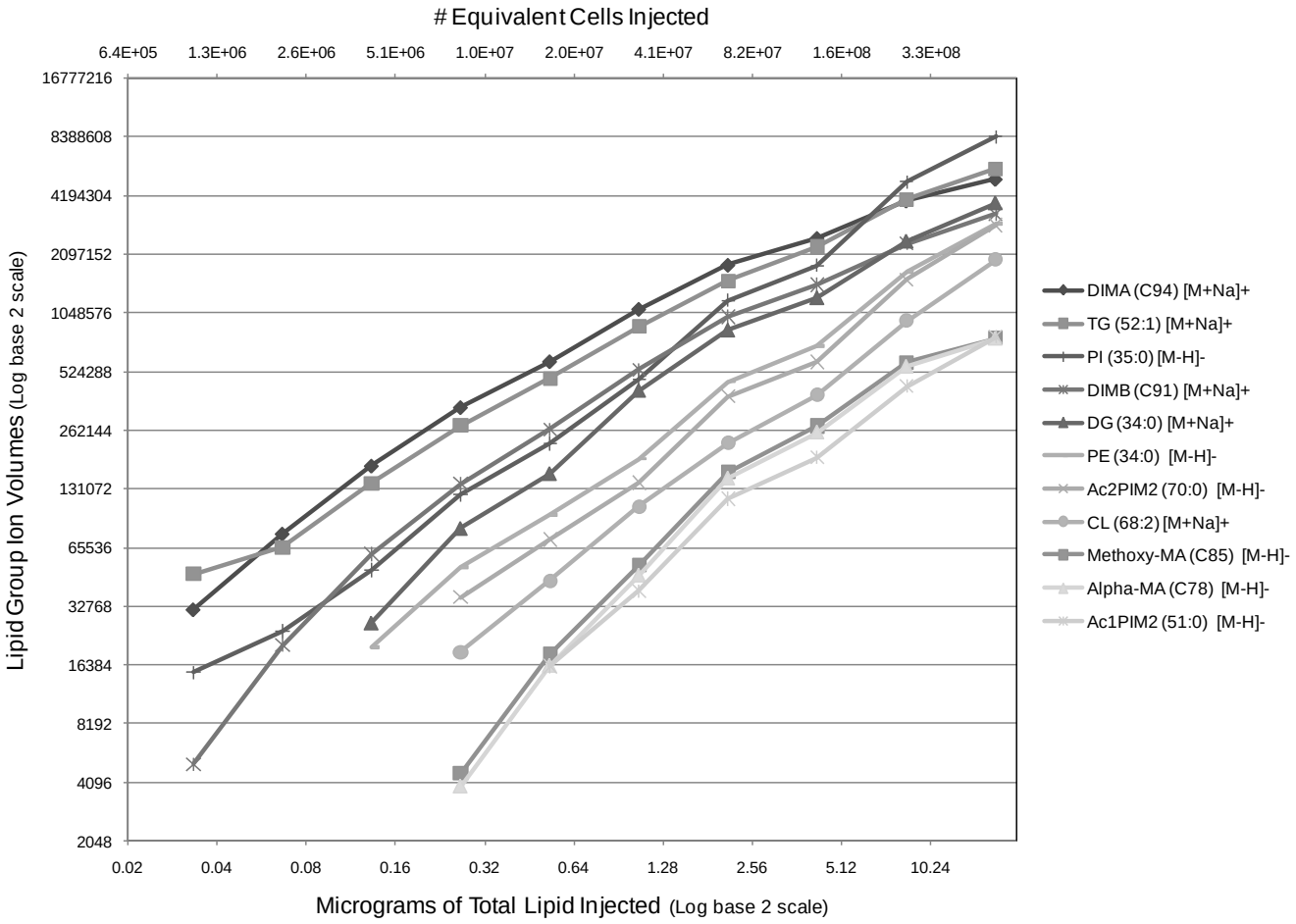

Figure 5

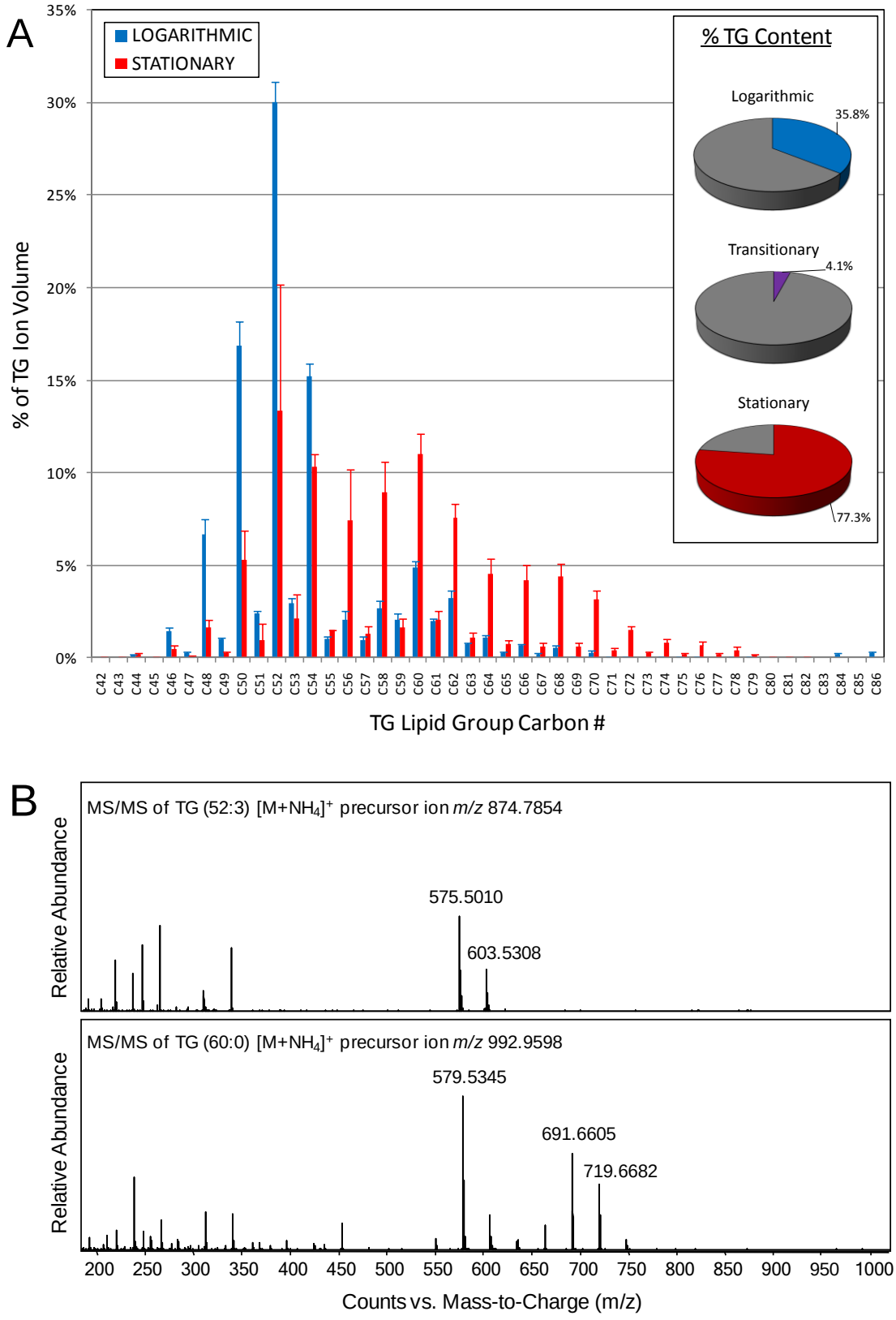

Figure 6

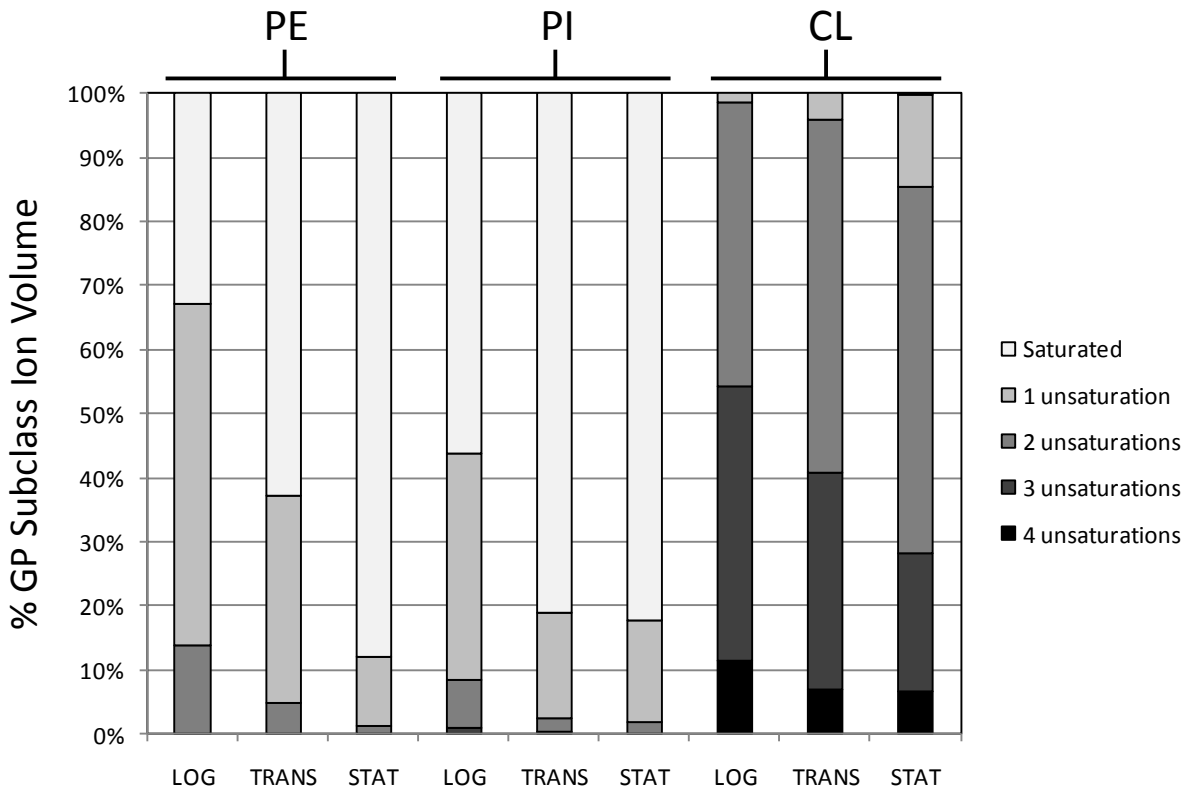

Figure 7