Discovery of sulfated metabolites in mycobacteria with a genetic and mass spectrometric approach


*Howard Hughes Medical Institute, †Department of Molecular and Cell Biology, ‡Department of Chemistry, and §School of Public Health, University of California, Berkeley, CA 94720

Edited by Chi-Huey Wong, The Scripps Research Institute, La Jolla, CA, and approved October 23, 2002 (received for review August 16, 2002)

The study of the metabolome presents numerous challenges, first among them being the cataloging of its constituents. A step in this direction will be the development of tools to identify metabolites that share common structural features. The importance of sulfated molecules in cell–cell communication motivated us to develop a rapid two-step method for identifying these metabolites in microorganisms, particularly in pathogenic mycobacteria. Sulfur-containing molecules were initially identified by mass spectral analysis of cell extracts from bacteria labeled metabolically with a stable sulfur isotope ($^{34}$SO₄$^{2-}$). To differentiate sulfated from reduced-sulfur-containing molecules, we employed a mutant lacking the reductive branch of the sulfate assimilation pathway. In these sulfur auxotrophs, heavy sulfate is channeled exclusively into sulfated metabolites. The method was applied to the discovery of several new sulfated molecules in Mycobacterium tuberculosis and Mycobacterium smegmatis. Because a sulfur auxotrophic strain is the only requirement of the approach, many microorganisms can be studied in this manner. Such genetic engineering in combination with stable isotopic labeling can be applied to various metabolic pathways and their products.

The modification of primary and secondary metabolites by the addition or removal of sulfate can have a profound influence on their biological properties (1–5). Typically, sulfated molecules are directed outside the cell, where they serve as modulators of cell–cell interactions. As a notable example pertinent to human health, sulfation of the glycans of endothelial CD34 engenders high-affinity binding with the leukocyte adhesion molecule L-selectin, facilitating an interaction that eventually leads to the recruitment of lymphocytes into peripheral lymph nodes (6). Similarly, sulfation of tyrosyl residues found on the chemokine receptor CCR5 is a modification required for binding of HIV gp120 and therefore efficient viral entry (7).

The roles of sulfated compounds in prokaryotes and other microbes are less clear. In one well-characterized case, however, sulfation acts as a modulator of cell–cell communication, similar to its role in eukaryotes. The nitrogen-fixing bacterium Sinorhizobium meliloti utilizes a sulfated glycolipid as a secondary messenger to induce root nodulation in its plant host alfalfa (4, 8). Mutants lacking the sulfotransferase that installs this sulfate ester are unable to induce root nodulation in alfalfa but gain the ability to colonize the roots of vetch. That sulfation of a single glycolipid plays such a vital role in nitrogen fixation has far-reaching implications for the agricultural community and presents a possible target for chemical or genetic engineering.

Sulfation may also be relevant to the process of bacterial pathogenesis (9). Several mycobacteria, including the human pathogen Mycobacterium tuberculosis and Mycobacterium avium, are known to produce sulfated compounds. One example is Sulfatide-1 (SL-1, Fig. 1A), a sulfated glycolipid from M. tuberculosis that is notable for the correlation of its abundance to strain virulence (10, 11). Other sulfated metabolites include 2-sulfo-6-deoxytalose within the cell wall glycopeptidolipid (GPL) of M. avium, a structure found up-regulated in a drug-resistant strain isolated from a patient with AIDS (12). GPL sulfation has also been detected in the opportunistic pathogen Mycobacterium fortuitum (13). These findings are suggestive of an important biological role for sulfation in mycobacteria.

Given the interesting properties of sulfated molecules, we sought to develop a sensitive and general method for their discovery from microbial metabolomes. Conventional studies using radioactive sulfate as a metabolic label are limited, because it is impossible to distinguish between compounds containing reduced sulfur (r-sulfur, i.e., compounds with thiols or thioethers) and those possessing sulfate esters (14, 15). Furthermore, direct chemical analysis of radiolabeled metabolites requires their laborious chromatographic separation.

The method we developed utilizes Fourier transform–ion cyclotron resonance mass spectrometry (FT-ICRMS) to identify sulfated metabolites by virtue of their metabolic labeling with a stable sulfur isotope ($^{34}$SO$_{4}^{2-}$). Sulfated molecules can be unambiguously distinguished from those possessing r-sulfur by performing the experiment with mutants lacking the reductive branch of the sulfate assimilation pathway (Fig. 2). MS is the preferred technique for this study, because its potential for metabolome-wide parallel analysis obviates the need for purification of individual metabolites and therefore any information, a priori, about the structure or physical

Fig. 1. Structure of SL-1 (A) and mycothiol (B). The SL-1 structure shown is as originally proposed by Goren and colleagues (10, 35).

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SL-1, Sulfatide-1; r-sulfur, compounds containing reduced sulfur; APS, adenosine-5’-phosphosulfate; MS/MS, tandem MS; FT-ICRMS, Fourier transform-ion cyclotron resonance MS.

‡J.D.M and M.D.L. contributed equally to this work.

∥To whom correspondence may be addressed. E-mail: bertozzi@chem.berkeley.edu or leary@oacs.berkeley.edu.
upon growth in $\text{SO}_4^{2-}$, and inorganic sulfate ($\text{SO}_4^{2-}$). The sulfate assimilation pathway begins with the conversion of microbes. ($\text{SO}_4^{2-}$) to enzyme-bound APS by ATP sulfurylase ($\text{CysN,D}$). APS lies at the branchpoint in the pathway in mycobacteria (17). APS kinase ($\text{CysC}$) phosphorylates APS to form APS-phosphoadenosine 5'-phosphosulfate (PAPS), the sulfate donor for sulfotransferases. These enzymes transfer the sulfuryl group onto a variety of substrates in the cell, forming sulfated molecules (red). In the reductive branch of the pathway, APS is reduced by APS reductase ($\text{CysH}$), eventually leading to the production of r-sulfur-containing molecules (blue). These can be distinguished from sulfated molecules (red) by using a mutation (green $\text{CysH}$) that blocks the reductive branch of the pathway. By MS can be used with heavy sulfur isotope labeling and the mutant in A ($\text{CysH}$) to identify and distinguish r-sulfur and sulfated molecules. Sulfur-containing molecules are identified by MS in wild-type cells by virtue of their incorporation of $^{34}$S upon growth in $\text{SO}_4^{2-}$ containing media (Step 1). Sulfated (red) molecules can be distinguished from r-sulfur-containing molecules (blue) by performing the same experiment using a $\text{CysH}$ strain (Step 2).

**Fig. 2.** General method for the identification of sulfated metabolites in microbes. (A) The sulfate assimilation pathway begins with the conversion of inorganic sulfate ($\text{SO}_4^{2-}$) to APS by ATP sulfurylase ($\text{CysN,D}$). APS lies at a branchpoint in the pathway in mycobacteria (17). APS kinase ($\text{CysC}$) phosphorylates APS to form APS-phosphoadenosine 5'-phosphosulfate (PAPS), the sulfate donor for sulfotransferases. These enzymes transfer the sulfuryl group onto a variety of substrates in the cell, forming sulfated molecules (red). In the reductive branch of the pathway, APS is reduced by APS reductase ($\text{CysH}$), eventually leading to the production of r-sulfur-containing molecules (blue). These can be distinguished from sulfated molecules (red) by using a mutation (green $\text{CysH}$) that blocks the reductive branch of the pathway. By MS can be used with heavy sulfur isotope labeling and the mutant in A ($\text{CysH}$) to identify and distinguish r-sulfur and sulfated molecules. Sulfur-containing molecules are identified by MS in wild-type cells by virtue of their incorporation of $^{34}$S upon growth in $\text{SO}_4^{2-}$ containing media (Step 1). Sulfated (red) molecules can be distinguished from r-sulfur-containing molecules (blue) by performing the same experiment using a $\text{CysH}$ strain (Step 2).

**Materials and Methods**

**Strains, Growth Analysis, and Metabolite Extraction.** M. smegmatis mcC155 (16) and M. tuberculosis H37Rv (American Type Culture Collection 27294) were the wild-type strains used in this study. The M. smegmatis ΔcysH strain has been reported previously (17). Sulfur auxotrophy of the M. tuberculosis ΔcysH strain was confirmed by growth in sulfur-free medium. Details relating to the construction of M. tuberculosis ΔcysH will be reported elsewhere.

For MS analysis, a 2-ml culture of each strain was grown in modified Hartman deBonte media containing either 2 mM Na$_2$SO$_4$ or Na$_2$$^{34}$SO$_4$ (ICON Isotopes, Summit, NJ) and 1 mM methionine when required for growth, as the sole sulfur sources (18). Once cells had reached stationary phase, they were pelleted by centrifugation and washed with PBS. The resulting cell pellet was extracted with 0.5 ml of 2:1 chloroform:methanol by vigorous shaking for 2 h at room temperature. The organic phase was clarified by centrifugation and removed for MS analysis. Authentic SL-1 was obtained from Colorado State University (National Institutes of Health, National Institute of Allergy and Infectious Diseases Contract N01 AI-75320).

**MS.** Spectra were acquired on a Bruker Apex II FT-ICR MS (Bruker Daltonics, Billerica, MA) equipped with a 7T actively shielded magnet. Cell extracts were sprayed from 2:1 chloroform:methanol at 3 µl/min on a pneumatically assisted electrospray source (Analytica, Branford, CT) in either positive or negative ion mode. Ions were accumulated in an external hexapole between 0.5 and 2 sec to yield strong signal before being transferred to the ion cyclotron resonance cell for mass analysis (19). For tandem MS (MS/MS) analysis, ions of interest were isolated by using a correlated harmonic excitation field isolation sweep (20). In certain cases, cleanup shots were used to eject ions that were not ejected from the initial sweep. Once isolated, ions of interest were collisionally sustained by sustained off-resonance irradiation at 500 Hz above the cyclotron frequency for 250 ms, using Ar as the collision gas (21). Each spectrum is an average of either 16 or 32 transients composed of 256,000 or 512,000 data points acquired by using XMass, Version 5.01 (Bruker, Billerica, MA).

**Synthesis of Trehalose-2-Sulfate.** Triethylammonium 4,6'-dibenzylidene-α,α-trehalose-2-sulfate was prepared according to the procedure of Vasella and coworkers (22). This compound (10 mg, 0.13 mmol) was treated with 2.0 ml of trifluoroacetic acid in 2.0 ml of MeOH. After stirring for 15 min, the solution was concentrated under reduced pressure. The residue was basified with 0.50 ml Et$_3$N and concentrated again.

**Results**

**Identification of R-Sulfur and Sulfated Compounds.** Searching for novel sulfated compounds by MS presented a number of challenges. For example, the chemical nature, abundance, and expected mass of these putative sulfated compounds were not known. Previous studies have shown the utility of isotopic labeling for deconvoluting complex biological samples during MS analysis (23–26). We sought to expand on these techniques for the purposes of discovering metabolites with a specific chemical modification, the sulfate ester. Toward this end, we designed a two-step method that utilizes $^{34}$S as a stable isotopic label, and a specific gene knockout ($\text{CysH}$) in the sulfate assimilation pathway to uniquely identify the sulfated constituents of a metabolome (see the legend to Fig. 2 for details of the sulfate assimilation pathway).

In the first step (Fig. 2B), wild-type cells were grown in minimal medium containing either Na$_2$SO$_4$ or Na$_2$$^{34}$SO$_4$ as the sole sulfur source. In the cells grown with Na$_2$$^{34}$SO$_4$-containing media, the masses of compounds containing either sulfur or sulfate were expected to shift by 1,996 x n Da, where n is the number of sulfur atoms. Extracts from labeled cultures were analyzed using a FT-ICRMS equipped with an electrospray ionization source. This arrangement has many advantages, including gentle ionization to yield intact ions, high resolution, high mass accuracy, and MS/MS capabilities (27).

To differentiate candidate compounds containing sulfate from those containing r-sulfur, a second step (Fig. 2B) was used.
that eliminates the sulfur reduction pathway. Specifically, the same isotope incorporation experiment was performed in a mutant strain, ΔcysH, which lacks the ability to reduce adenosine-5′-phosphosulfate (APS) to sulfite, the first committed step in the biosynthesis of r-sulfur (17). Thus, when a ΔcysH strain was grown in [32S]-methionine (32met) in the biosynthesis of r-sulfur (17), the ΔcysH mutant strain was shown to produce a sulfated molecule (Fig. 3). Similar spectra were obtained from both Na232SO4 and Na234SO4-labeled extracts of M. tuberculosis and M. smegmatis (Table 1).

To demonstrate the method for r-sulfur-containing compounds, we focused our attention on mycothiol, an abundant cysteine derivative found in actinomycetes that is involved in neutralizing potentially damaging cellular electrophiles (Fig. 1B). Using exact mass measurements (m/z 485.1437, within 2 ppm of calculated mass) and MS/MS experiments of one SL-1 ion peak (485.1437), within 2 ppm of calculated mass) and MS/MS experiments of one SL-1 ion peak, we showed that the mycothiol ion was present in both M. tuberculosis and M. smegmatis cultures (Fig. 1B and 3B). We next examined the extract from M. tuberculosis grown in [32S]SO42−. In this sample, the SL-1 ions shifted in mass by the expected 2.00 Da, indicating the presence of a single sulfur atom in the structure of the molecule (Fig. 3C). The masses of SL-1 lipids from extracts taken from M. tuberculosis ΔcysH cultures labeled with [34S]SO42− also shifted by 2.00 Da, as predicted for a sulfated molecule (Fig. 3C). MS/MS experiments of one SL-1 ion peak yielded the HS2PO4− product ion, providing further evidence that the selected ion was sulfated (data not shown).

To demonstrate the method for r-sulfur-containing compounds, we focused our attention on mycothiol, an abundant cysteine derivative found in actinomycetes that is involved in neutralizing potentially damaging cellular electrophiles (Fig. 1B). Using exact mass measurements (m/z 485.1437, within 2 ppm of calculated mass) and MS/MS experiments of one SL-1 ion peak, we showed that the mycothiol ion was present in both M. tuberculosis and M. smegmatis cultures (Fig. 1B and 3B). We next examined the extract from M. tuberculosis grown in [32S]SO42−. In this sample, the SL-1 ions shifted in mass by the expected 2.00 Da, indicating the presence of a single sulfur atom in the structure of the molecule (Fig. 3C). The masses of SL-1 lipids from extracts taken from M. tuberculosis ΔcysH cultures labeled with [34S]SO42− also shifted by 2.00 Da, as predicted for a sulfated molecule (Fig. 3C). MS/MS experiments of one SL-1 ion peak yielded the HS2PO4− product ion, providing further evidence that the selected ion was sulfated (data not shown).

Table 1. Summary of sulfated and r-sulfur-containing molecules found in M. tuberculosis and M. smegmatis

<table>
<thead>
<tr>
<th>m/z 32S*</th>
<th>m/z 34S</th>
<th>Shift in ΔcysH</th>
<th>HSO4− ion observed in MS/MS</th>
<th>Form of sulfur</th>
<th>Metabolite annotations†‡‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>m/z 34S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td>421.07</td>
<td>423.06</td>
<td>Y</td>
<td>Y</td>
<td>Sulfate</td>
</tr>
<tr>
<td>485.14</td>
<td>487.13</td>
<td>N</td>
<td>N</td>
<td>r-sulfur§</td>
<td>Mycothiol</td>
</tr>
<tr>
<td>881.57</td>
<td>883.57</td>
<td>Y</td>
<td>Y</td>
<td>Sulfate</td>
<td>None</td>
</tr>
<tr>
<td>1,277.97</td>
<td>1,279.97</td>
<td>Y</td>
<td>2 ppm</td>
<td>Sulfate</td>
<td>SL-1 precursor†</td>
</tr>
<tr>
<td>2,543.22*</td>
<td>2,545.21</td>
<td>Y</td>
<td>Y</td>
<td>Sulfate</td>
<td>SL-1</td>
</tr>
<tr>
<td>M. smegmatis</td>
<td>421.07</td>
<td>423.06</td>
<td>Y</td>
<td>Y</td>
<td>Sulfate</td>
</tr>
<tr>
<td>485.14</td>
<td>487.13</td>
<td>N</td>
<td>N</td>
<td>r-sulfur§</td>
<td>Mycothiol</td>
</tr>
<tr>
<td>521.12</td>
<td>523.12</td>
<td>N</td>
<td>N</td>
<td>r-sulfur§</td>
<td>None</td>
</tr>
<tr>
<td>583.13</td>
<td>585.12</td>
<td>N</td>
<td>N/A</td>
<td>r-sulfur</td>
<td>None</td>
</tr>
<tr>
<td>597.14</td>
<td>599.14</td>
<td>N</td>
<td>N/A</td>
<td>r-sulfur</td>
<td>None</td>
</tr>
<tr>
<td>632.20</td>
<td>634.20</td>
<td>N</td>
<td>N/A</td>
<td>r-sulfur</td>
<td>None</td>
</tr>
<tr>
<td>639.19</td>
<td>641.18</td>
<td>Y</td>
<td>Y</td>
<td>Sulfate</td>
<td>None</td>
</tr>
<tr>
<td>753.23</td>
<td>755.22</td>
<td>N</td>
<td>N/A</td>
<td>r-sulfur</td>
<td>None</td>
</tr>
<tr>
<td>1,009.25†</td>
<td>1,013.24</td>
<td>N</td>
<td>N/A</td>
<td>r-sulfur</td>
<td>Mycothiol dimer§§</td>
</tr>
</tbody>
</table>

*Unless otherwise noted, masses given are monoisotopic and from negative ion mode acquisitions. All ions were singly charged. †Name or structure assigned to the metabolite. ‡Assignment made in this study. §Ion containing a reduced form of sulfur, either a thioether or a thiol. †Due to low abundance, we were unable to perform MS/MS on this ion. ¶Calculation show this ion likely corresponds to a specific diacylated trehalose-2-sulfate (see text). **Monoisotopic mass of a major lipiform. It should be noted that the calculated mass of the originally proposed SL-1 structure does not match with our measured mass of the SL-1 standard. The standard used for our analysis is of the same batch used to elucidate the structure of SL-1. ††Not applicable; MS/MS not performed on r-sulfur-containing compounds, with the exception of mycothiol. §§This ion observed in positive mode. †‡‡Two mycothiol molecules linked by a disulfide bond, potassiated adduct.
ing ambiguity in the assignment of sulfate. Since no sulfated molecules have ever been characterized from *M. smegmatis*, we chose to examine one of them, \( m/z \) 421.07, in more detail. Exact mass measurements of this ion gave \( m/z \) 421.0661, which is within 1 ppm of the expected mass of a sulfated disaccharide. To further test this possibility, we performed MS/MS on \( m/z \) 421.07 (Fig. 4A). Several product ions characteristic of a sulfated disaccharide were observed. The product ions at \( m/z \) 259.0 and \( m/z \) 241.0 correspond to cleavage on either side of the glycosidic bond with charge retention on the sulfate. The ion at \( m/z \) 97.0 corresponds to an elemental composition of HSO\(_4\)^-, whereas the ion observed at \( m/z \) 139.0 results from a cross-ring cleavage, yielding an ion with an elemental composition of \([C_2H_3O_5S]^-\). We propose that the sulfated disaccharide in *M. smegmatis* corresponds to trehalose-2-sulfate, the core structure of SL-1. As further proof of structure, we compared MS/MS spectra from \( m/z \) 421.07 and a chemically synthesized trehalose-2-sulfate standard (Fig. 5B). An identical set of product ions was found in both samples, strongly supporting the identity of \( m/z \) 421.07 as trehalose-2-sulfate. Assignment of the major product ions of \( m/z \) 421.07, based on the structure of trehalose-2-sulfate, is shown in the inset of Fig. 5B.

**Discovery of New Sulfated Molecules in M. tuberculosis.** Based on heavy sulfur incorporation in both wild-type and ΔcysH \(^{34}\text{SO}_4^-\)-labeling experiments, three additional sulfated compounds were identified in *M. tuberculosis*, corresponding to \( m/z \) 421.07, 881.57, and 1,277.97 (Table 1). As shown in Fig. 4B, the \( m/z \) 881.57 ion differs by only 0.03 and 0.07 Da from its closest isobar in the \(^{32}\text{S}\) and \(^{34}\text{S}\)-forms, respectively. This example underscores the need for high-resolution data when searching a complex metabolome for mass shifts within unknown ions. The \( m/z \) 421.07 ion, also observed in *M. smegmatis*, corresponds to trehalose-2-sulfate (discussed above), the disaccharide core of SL-1. The possibility that \( m/z \) 1,277.97 and 881.57 were also related to SL-1, as either metabolic precursors or ionization fragments, was also explored. If trehalose-2-sulfate is the first intermediate in the biosynthesis of SL-1, as observation of this free sulfated disaccharide in *M. tuberculosis* extracts would suggest, it is likely that \( m/z \) 1,277.97 is a downstream intermediate in the same pathway. The [M−H]^- ion at \( m/z \) 1,277.97 agrees with the deprotonated mass of trehalose-2-sulfate acylated with a palmitoyl group (C16 fatty acid) and a single hydroxyphthioceranoyl group (\( m/z \) 1,277.93, Fig. 1A). The observation of these proposed intermediates offers a possible biosynthetic route to SL-1, in which sulfation of the trehalose core precedes lipid modification.

By contrast, \( m/z \) 881.57 appears to be structurally unrelated to SL-1. To further characterize this compound, we subjected the \(^{34}\text{S}\) form to MS/MS analysis. This revealed the H\(^{34}\text{SO}_4^-\)-product ion (\( m/z \) 99.0) and a product ion corresponding to the
loss of $^{34}\text{SO}_3^-$ ($m/z$ 801.6, Table 1, and data not shown). MS/MS therefore provided direct chemical evidence that $m/z$ 881.57 was sulfated. Exact mass measurements and additional preliminary MS/MS experiments showed that the ion is likely to have both lipid and peptide components. Further structural details of $m/z$ 881.57 are forthcoming and will be reported elsewhere.

**Discussion**

We developed a simple two-step procedure for discovering sulfated metabolites from complex cellular extracts. The application of this method to *M. tuberculosis* and *M. smegmatis* resulted in a number of significant findings, including the first examples of sulfated molecules in *M. smegmatis*, observation of a biosynthetic precursor to SL-1, and the identification of a novel sulfated molecule in *M. tuberculosis*. The ability of the method to differentiate between sulfated and r-sulfur-containing compounds was demonstrated by using SL-1 and mycothiol as appropriate internal controls.

Labeling with $^{34}\text{SO}_3^-$ can also be used for MS-based discovery of sulfated proteins, polysaccharides, and small molecules in mammalian systems. In higher organisms, the experiment is simplified by the absence of sulfate reduction pathways; heavy sulfate is only incorporated into sulfated molecules. We anticipate that this metabolic labeling/MS approach will be amenable to studies of protein sulfation and the sulfated glycosaminoglycans that populate the extracellular matrix.

In addition to its use in the discovery of sulfated molecules, this method provides a framework and strategy for discovering other types of secondary metabolites with relatively little information about their chemical structure. If a known metabolic pathway is required for the biosynthesis of a compound, mutations that block undesired branchpoints in this pathway can be made in order to channel the isotope to the target molecule class. This technology, coupled with advances in MS, will aid in studies of metabolic flux, the identification of new biosynthetic pathways, and the annotation of cellular metabolomes.

J.D.M. was supported by a Ford Foundation Predoctoral Fellowship. S.J.W. was supported by a Howard Hughes Medical Institute Fellowship of the Life Sciences Research Foundation. Work in the authors’ laboratories is supported by grants to C.R.B. (AI51622) and J.A.L. (GM47356) from the National Institutes of Health.