

Assignment of Endogenous Substrates to Enzymes by Global Metabolite Profiling[†]

Alan Saghatelian, Sunia A. Trauger, Elizabeth J. Want, Edward G. Hawkins, Gary Siuzdak, and Benjamin F. Cravatt*

*The Skaggs Institute for Chemical Biology and Departments of Cell Biology and Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037**Received September 11, 2004; Revised Manuscript Received October 1, 2004*

ABSTRACT: Enzymes regulate biological processes through the conversion of specific substrates to products. Therefore, of fundamental interest for every enzyme is the elucidation of its natural substrates. Here, we describe a general strategy for identifying endogenous substrates of enzymes by untargeted liquid chromatography–mass spectrometry (LC–MS) analysis of tissue metabolomes from wild-type and enzyme-inactivated organisms. We use this method to discover several brain lipids regulated by the mammalian enzyme fatty acid amide hydrolase (FAAH) *in vivo*, including known signaling molecules (e.g., the endogenous cannabinoid anandamide) and a novel family of nervous system-enriched natural products, the taurine-conjugated fatty acids. Remarkably, the relative hydrolytic activity that FAAH exhibited for lipid metabolites *in vitro* was not predictive of the identity of specific FAAH substrates *in vivo*. Thus, global metabolite profiling establishes unanticipated connections between the proteome and metabolome that enable assignment of an enzyme's unique biochemical functions *in vivo*.

Enzymes are central components of nearly all signal transduction cascades and metabolic pathways. Genome sequencing projects have provided, for the first time, a view of the total set of enzymes expressed in an organism. Additionally, emerging genetic technologies (e.g., RNAi, targeted gene disruption) coupled with cell- and organism-based phenotypic screens enable a genome-wide analysis of the (patho)physiological activities of enzymes (1–3). In contrast, complementary global approaches for elucidating the endogenous biochemical functions of enzymes are lacking, and as a consequence, the substrates of many enzymes remain unknown. Thus, the assignment of natural substrates to enzymes represents a major problem of the postgenomic era, and systematic methods that directly connect the proteome with the metabolome are needed for its solution.

The substrate selectivities of enzymes are generally determined *in vitro* using purified protein preparations. However, translating these *in vitro* findings into a comprehensive understanding of the scope of substrates utilized by enzymes *in vivo* is problematic for multiple reasons. First, in the living cell and organism, enzymes do not function in isolation, but rather as parts of large and complex biochemical networks (4). Accordingly, *in vitro* assays may fail to account for many potentially competitive metabolic pathways that alter or restrict the substrates utilized by a particular enzyme *in vivo* (5). Second, efforts to assign natural substrates to enzymes *in vitro* are limited by our current knowledge of cellular biochemistry and, therefore, ill-suited

for the discovery of novel metabolites that are regulated by enzymes *in vivo*. Finally, many enzymes are subject to post-translational regulation *in vivo*, including covalent modification (e.g., phosphorylation) and protein–protein interactions, which may alter substrate recognition and catalysis (6).

We postulated that a more direct route for identifying endogenous substrates of enzymes could be achieved by globally profiling the metabolic consequences of enzyme inactivation *in vivo*. In this model, metabolites that accumulated would be considered candidate endogenous substrates for the disrupted enzyme. To enable the simultaneous characterization of both known and unknown natural products regulated by specific enzymes *in vivo*, we developed an “untargeted” (i.e., standard-free) mass spectrometry-based approach for comparative metabolomics, termed discovery metabolite profiling (DMP).¹ We apply DMP to profile the nervous system metabolomes of wild-type mice and mice lacking the enzyme fatty acid amide hydrolase (FAAH), resulting in the identification of several endogenous substrates for this enzyme, including known signaling lipids (e.g., the endogenous cannabinoid anandamide) and a novel structural class of central nervous system (CNS) metabolites, the taurine-conjugated fatty acids.

EXPERIMENTAL PROCEDURES

Tissue Isolation and Extraction. A 2:1:1 CHCl₃/MeOH/1% NaCl solution (8 mL per brain and 4 mL per spinal cord in 8 mL vials) was prepared for tissue extraction to isolate organic soluble metabolites (25, 26). For targeted LC–MS

[†] This work was supported by the NIH (Grants DA015197 and DA017259), the Helen L. Dorris Child and Adolescent Neurological and Psychiatric Disorder Institute, and the Skaggs Institute for Chemical Biology. A.S. is supported by a Merck Fellowship of the Life Sciences Research Foundation.

* To whom correspondence should be addressed. Phone: (858) 784-8633. Fax: (858) 784-8023. E-mail: cravatt@scripps.edu.

¹ Abbreviations: CNS, central nervous system; DMP, discovery metabolite profiling; FAAH, fatty acid amide hydrolase; FTMS, Fourier transform mass spectrometry; LC, liquid chromatography; MS, mass spectrometry; NAE, *N*-acyl ethanolamine; NAT, *N*-acyl taurine; SIM, selected ion monitoring.

measurements, deuterated standards were included in the mixture as described previously (10). FAAH(+/+) and FAAH(-/-) mice (3–6 months of age) were sacrificed at the same time of day and tissues immediately isolated, weighed, placed into the CHCl₃/MeOH/1% NaCl solution, and homogenized using dounce tissue grinders. Each sample was then centrifuged at 2500 rpm for 10 min at 4 °C in a glass vial. After centrifugation, the organic (bottom) and aqueous layers (top) were clearly distinguishable with a layer of insoluble material between them. The organic layer was carefully removed and transferred to another vial. The aqueous layer was extracted with an additional 1 mL of CHCl₃, 0.5 mL of MeOH, and 1 mL of a 1% NaCl/1% formic acid mixture. This solution was mixed vigorously for 30–60 s and then centrifuged at 2500 rpm for 10 min at 4 °C. The organic layers from the first and second extractions were combined and concentrated under a stream of nitrogen. Samples were stored at –80 °C (always for less than 1 day) and dissolved in 120 µL of CHCl₃ prior to analysis by LC–MS.

LC–MS of Tissue Metabolomes. LC–MS analysis was performed using an Agilent 1100 LC-MSD SL instrument. For the LC analysis, a HAILIL 300 C18 column (5 µm, 4.6 mm × 100 mm) from Higgins Analytical was used together with a precolumn (C18, 3.5 µm, 2 mm × 20 mm). Mobile phase A consisted of a 95:5 water/methanol mixture, and mobile phase B was made up of 2-propanol, methanol, and water in a 50:45:5 ratio. Solvent modifiers such as 0.1% formic acid, for the positive ionization mode, and 0.1% ammonium hydroxide, for the negative ionization mode, were used to assist ion formation as well as improve the LC resolution. The flow rate for each run started at 0.1 mL/min for 5 min, to alleviate the backpressure associated with injecting CHCl₃, followed by a flow rate of 0.4 mL/min for the duration of the gradient. The gradient started at 0% B and then linearly increased to 100% B over the course of 60 min followed by an isocratic gradient of 100% B for 30 min before equilibration for 10 min at 0% B. The total analysis time, including 5 min at 0.1 mL/min, was 105 min. MS analysis was performed with an electrospray ionization (ESI) source. The capillary voltage was set to 3.0 kV and the fragmentor voltage to 100 V. The drying gas temperature was 350 °C, the drying gas flow rate was 10 L/min, and the nebulizer pressure was 35 psi. Data were collected using a mass range of 200–1200 Da, and each run was performed using 40 µL injections of tissue metabolite extract.

LC–MS Data Analysis. The analysis of the resulting total ion chromatogram was performed manually by generating extracted ion chromatograms (EICs) in 5 Da increments (e.g., 200–205, 205–210, ..., 1195–1200). EICs of FAAH(+/+) and FAAH(-/-) samples were compared in a pairwise manner to identify changes (i.e., new peaks or changes in the magnitude of peaks) between samples for a given mass range and retention time. After peaks that were not shared by all samples had been discarded, the remaining peaks were quantified using the area under the peaks. The measured areas were then normalized to the amount of tissue and averaged ($N = 6$) to afford the mean area for a given peak in the chromatogram. Finally, the peak ratios between FAAH(+/+) and FAAH(-/-) samples provide a quantitative measure of the relative metabolite levels. With signals that fell below the limit of detection in FAAH(+/+) samples, a lower cutoff

ion intensity of 32 500 was used. In those cases where this lower limit was used, the average ion intensity values are reported to be greater than or equal to the calculated ion intensity and resulting FAAH(-/-)/FAAH(+/+) ratios are reported to be greater than or equal to the calculated ratio.

Preparative HPLC Purification of *N*-Acyl Taurines (NATs). The metabolite extracts from five FAAH(-/-) spinal cords were combined for a single LC purification using a Hitachi 7000 series HPLC system. For the purification, a Cliepus C18 column (5 µm, 10 mm × 150 mm) from Higgins Analytical was used. Mobile phase A consisted of water and methanol in a 95:5 ratio, with 0.1% ammonium hydroxide, and mobile phase B consisted of 2-propanol, methanol, and water in a 50:45:5 ratio, with 0.1% ammonium hydroxide. The gradient started at 0% B and then linearly increased to 100% B over the course of 60 min followed by an isocratic gradient of 100% B for 20 min at a flow rate of 2.5 mL/min. Fractions (one per minute) were collected using a Gilson FC 203B fraction collector. Fractions containing the 446, 460, 472, and 474 metabolites were identified by MS analysis. These fractions were then collected, and the solvent was removed using a rotary evaporator. The samples were then dissolved in a minimal amount of solvent B (200–300 µL) for exact mass and MS/MS analysis.

Fourier Transform Mass Spectrometry (FTMS) Experiments. The high-accuracy measurements were performed in the negative ion mode using a Bruker (Billerica, MA) APEX III (7.0 T) FTMS instrument equipped with an Apollo electrospray source. The collected LC fractions were mixed with a collection of small molecule standards and directly infused at 3 µL/min using a Harvard Apparatus (Holliston, MA) syringe pump. Pneumatic assist at a backing pressure of 60 psi was used along with an optimized flow rate of heated counter-current drying gas (300 °C). Ion accumulation was performed using SideKick without pulsed gas trapping. Data acquisition times of approximately 1 min were used, yielding a resolving power of ~130000 at m/z 446 in broadband in the m/z range of 200–2200. Calculated molecular masses for ions generated by a mixture of small molecule standards were used to internally calibrate the data.

Tandem MS Experiments. MS/MS experiments were performed in the negative ion mode using a Micromass (Manchester, U.K.) Q-ToF Micro instrument equipped with a Z-spray electrospray source and a lockmass sprayer. The source temperature was set to 110 °C with a cone gas flow of 150 L/h, a desolvation gas temperature of 365 °C, and a nebulization gas flow rate of 350 L/h. The capillary voltage was set at 3.2 kV and the cone voltage at 30 V. The collision energy was set at 40–45 V. Samples were directly infused at 4 µL/min using a Harvard Apparatus syringe pump. MS/MS data were collected in the centroid mode over a scan range of m/z 50–500 for acquisition times of 2 min.

Representative Synthesis of NATs (described for C18:1 NAT, where 18 and 1 refer to the number of carbons and degrees of unsaturation in the NAT acyl chain, respectively). To a 10 mL round-bottom flask fitted with a magnetic stir bar were added oleic acid (50 mg, 0.17 mmol, 1 equiv) and chloroform (3 mL). To this mixture was added oxalyl chloride (291 mg, 200 µL, 2.3 mmol, 13.5 equiv), and the solution was stirred for 4 h. The reaction mixture was then concentrated using a rotary evaporator, and the residue was then dissolved in chloroform (2 mL) followed by the addition

of taurine (50 mg, 0.4 mmol, 2.4 equiv) as a 25 mg/mL solution in a 50:50:1 dioxane/water/triethylamine mixture. After 1 h, the reaction mixture was concentrated and purified directly by silica gel chromatography (80:10:10 chloroform/methanol/acetic acid) to afford a light brown solid. This solid was then dissolved in a 50:50:2 chloroform/methanol/10 N NaOH solution (10 mL) and passed over a column of Dowex 1x2-400 (3 g). The resin was then washed with methanol (50 mL) and 2% acetic acid in methanol (50 mL). The product was then eluted using 2% HCl in methanol to afford C18:1 NAT (15 mg, 23%): $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 5.34 (m, 2H), 3.66 (t, 2H, $J = 7$ Hz), 3.00 (t, 2H, $J = 7$ Hz), 2.31 (t, 2H, $J = 7.7$ Hz), 2.03 (m, 4H), 1.62 (m, 2H), 1.31 (m, 20 H), 0.90 (t, 3H, $J = 6.95$ Hz). Note that yields of very long chain fatty acids were lower (~ 5 –10%) due to the limited solubility of the intermediates.

Targeted LC–MS Measurements of NATs. A 2:1:1 $\text{CHCl}_3/\text{MeOH}/1\%$ NaCl solution (8 mL per brain and 4 mL per spinal cord in 8 mL vials) was prepared for tissue extraction. Following a previously described procedure for quantifying fatty acids (27), NATs were quantified by using a C17:0 NAT standard (500 pmol), which was synthesized as described above and added to the extraction solution. Mice were sacrificed and tissues immediately isolated, weighed, placed into the $\text{CHCl}_3/\text{MeOH}/1\%$ NaCl solution, and homogenized using dounce tissue grinders. Each sample was then processed as described above and analyzed by targeted LC–MS using selected ion monitoring (SIM). Concentrations of NATs were estimated with respect to the C17:0 NAT standard.

FAAH Substrate Assays. FAAH assays were performed by following the conversion of substrates to their corresponding fatty acids by LC–MS. FAAH was recombinantly expressed and purified from *Escherichia coli* as described previously (28). Reactions were conducted with 1.25–125 nM FAAH and 12.5–150 μM *N*-acyl ethanolamine (NAE) or NAT substrate in a reaction buffer of 100 mM Tris-HCl, 1 mM EDTA, and 0.1% Triton X-100 (pH 8, adjusted using HCl or NaOH). Reactions were quenched with 0.5 N HCl. LC–MS analysis was performed using an Agilent 1100 LC-MSD SL instrument. For the LC analysis, a HAILSIL 100 C8 column (5 μm , 4.6 mm \times 50 mm) from Higgins Analytical was used. Mobile phase A consisted of water and methanol in a 95:5 ratio, with 0.1% ammonium hydroxide, and mobile phase B consisted of 2-propanol, methanol, and water in a 50:45:5 ratio, with 0.1% ammonium hydroxide. The gradient started at 10% B and then linearly increased to 100% B over the course of 10 min followed by an isocratic gradient of 100% B for 5 min at a flow rate of 0.5 mL/min. Aliquots of the quenched solutions were directly injected into the LC–MS system for analysis. SIM was used to measure both the starting NAE or NAT and the corresponding fatty acid hydrolysis product. Standard curves of fatty acids (C18:1, C22:0, and C24:0) allowed the conversion of ion intensities into molar quantities. Each substrate was tested at four independent concentrations, and at each concentration, four separate time points were measured (4–90 min) such that no greater than 20% formation of product was observed at the final time point. Linear kinetics was observed for each substrate at each concentration that was tested, and from these data, initial velocities were calculated and used to determine k_{cat} , K_m , and k_{cat}/K_m . For very long chain NAE and NATs

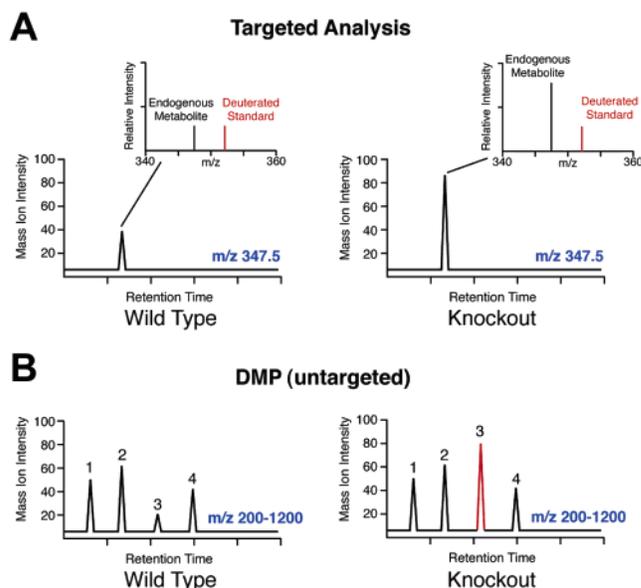


FIGURE 1: Comparison of targeted and untargeted LC–MS methods for comparative metabolite analysis. (A) General scheme for targeted LC–MS analysis, in which metabolites are detected by SIM (shown for a metabolite with a mass of 347.5) and their levels quantified by comparing mass signals to those of isotopically distinct internal standards. (B) General scheme for DMP, an untargeted LC–MS approach, in which metabolites are detected in the broad mass scanning mode (e.g., 200–1200 mass units) and their levels quantified by measuring direct mass ion intensities (i.e., without the inclusion of internal standards). Enzyme-regulated metabolites are identified by comparison of mass ion intensities between wild-type and knockout samples (e.g., metabolite 3).

(≥ 22), K_m and k_{cat} values were not separately determined due to substrate solubility limits (~ 200 μM in reaction buffer [100 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.1% Triton X-100, and 2.5% DMSO]). However, the rates of hydrolysis of these substrates by FAAH increased linearly over a concentration range of 25–150 μM (Figure 2 of the Supporting Information), and from these data, specificity constants (k_{cat}/K_m) were determined.

RESULTS

Development of an Untargeted, Standard-Free Approach for Comparative Metabolite Profiling. Metabolites are typically measured in biological samples by “targeted” gas or liquid chromatography (LC)–MS techniques, in which the levels of specific compounds are determined using isotopic variants as internal standards coupled with MS analysis by selected ion monitoring (SIM) (7, 8) (Figure 1A). However, these targeted methods are restricted to the analysis of a limited number of known metabolites and therefore are not compatible with the discovery of novel substrates of enzymes. Key to the implementation of discovery metabolite profiling (DMP) was the supposition that the absolute mass ion intensities measured during LC–MS analysis could be used in a broad mass-scanning mode to quantify the relative levels of metabolites in tissue samples and thereby obviate the need for internal standards (Figure 1B). To test this premise, we compared the LC–MS metabolite profiles of central nervous system (CNS) tissues from wild-type mice and mice lacking the enzyme FAAH (9) [FAAH(–/–) mice], which degrades several neural signaling lipids *in vivo* (10),

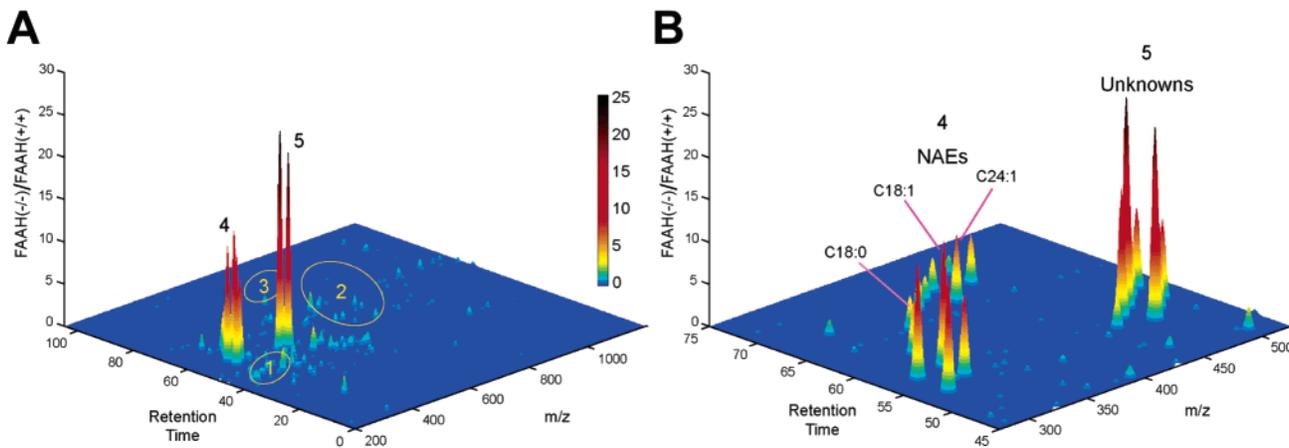


FIGURE 2: Discovery metabolite profiling (DMP) of FAAH(-/-) and FAAH(+/+) brains, where mass ion intensity ratios [FAAH(-/-)/FAAH(+/+)] of metabolites are presented on three-dimensional surface plots. (A) Global view of the relative levels of metabolites in FAAH(-/-) and FAAH(+/+) brains, plotted over a mass range of 200–1200 Da and liquid chromatography retention times of 0–105 min (plot shown for the negative ionization mode). FAAH(-/-) brains possessed highly elevated levels of NAEs (lipid group 4) and an unknown class of metabolites (group 5), while other lipids [e.g., free fatty acids (group 1), phospholipids (group 2), and ceramides (group 3)] were unaltered in these samples. (B) Close-up view of the LC–MS region containing elevated levels of metabolites in FAAH(-/-) brains. Highlighted are representative known (e.g., C18:1) and novel (e.g., C24:1) NAE substrates of FAAH, as well as an unknown family of metabolites also upregulated in FAAH(-/-) brains. Data represent the ratios of the averages \pm standard errors (SE) of six independent experiments per group.

including the endogenous cannabinoid *N*-arachidonoyl ethanolamine [anandamide (11)]. FAAH(-/-) mice have been shown by targeted LC–MS analysis to possess highly elevated brain levels of anandamide and a select number of other *N*-acyl ethanolamine (NAE) substrates (10, 12) and thus provided an excellent model for testing the capacity of DMP to identify both known changes in the levels of specific FAAH substrates and potentially novel metabolites regulated by this enzyme *in vivo*.

General conditions for DMP were established for the optimal extraction, LC separation, and mass ionization of tissue lipid metabolites using a diverse set of purified standards, including NAEs, ceramides, phospholipids, fatty acids, and glycerol esters (Table 1 of the Supporting Information). These conditions were then applied to CNS tissues from FAAH(+/+) and FAAH(-/-) mice using LC–electrospray ionization (ESI) MS in both the positive and negative ion modes scanning across a mass range of 200–1200 Da. Comparative analysis of the resulting total ion chromatograms was performed manually by generating extracted ion chromatograms in 5 Da increments (e.g., 200–205, 205–210, ..., 1195–1200). Six independent experiments were run for each tissue of each genotype, and the intensities of peaks that were reproducibly detected in all six experiments were averaged (Tables 2 and 3 of the Supporting Information). The intensity ratios of averaged peaks showing equivalent masses and retention times in the FAAH(+/+) and FAAH(-/-) samples provided a quantitative measure of changes in the metabolome resulting from the inactivation of FAAH (Figure 2). Known endogenous NAE substrates of FAAH (e.g., C18:1 and C18:0) were readily identified in these comparative metabolite profiles, and their levels were estimated to be elevated 12–18-fold in the brains and spinal cords of FAAH(-/-) mice (Figure 2, group 4; also see Table 1 and Tables 2 and 3 of the Supporting Information). Notably, these DMP measurements were all within 1.6-fold of results obtained by targeted LC–MS methods (Table 1 and Table 4 of the Supporting Information), demonstrating that the relative levels of endogenous metabolites can be

Table 1: Relative Levels of Representative Lipids Measured by DMP in Central Nervous System Tissue from FAAH(+/+) and FAAH(-/-) Mice^a

lipid class acyl chain	FAAH(-/-)/FAAH(+/+) brain	FAAH(-/-)/FAAH(+/+) spinal cord
NAEs		
C16:1	≥ 4	≥ 5
C18:1	12.8 (20.2)	17.7 (17.1)
C18:0	13.8 (10.3)	12.0 (8.6)
C20:4	≥ 4 (10.9)	≥ 5 (10.4)
C22:0	≥ 3.8	6.3
C24:1	4	8.1
MAGs		
C16:0	0.82	0.97
C18:1	0.72	0.91
C18:0	0.68	0.98
ceramides		
C16:0	0.78	1.1
C18:1	0.73	1.12
NATs		
C22:0 (446)	11 (16)	13 (21)
C23:0 (460)	16 (22)	13 (19)
C24:1 (472)	24 (43)	26 (42)
C24:0 (474)	26 (41)	23 (37)
phospholipids		
C34:1 PA	0.97	0.89
C36:2 PA	0.89	0.95
C34:1 PE	0.86	0.96
C36:3 PE	0.83	0.89
C36:2 PE	0.85	0.89
C36:2 PC	0.92	0.98
C36:2 PS	1.01	1.03
FFAs		
C16:0	1.23	0.82
C18:1	1.32	1.04
C18:0	0.93	0.95

^a Data in parentheses were determined by targeted LC–MS methods. NAEs, *N*-acyl ethanolamines; MAGs, monoacylglycerols; NATs, *N*-acyl taurines; FFAs, free fatty acids. Data represent the mass ion intensity ratios of the averages of six independent experiments per group. For a more complete summary of lipid expression profiles, including raw mass ion intensity data, see Tables 2 and 3 of the Supporting Information.

accurately quantified in tissue extracts by untargeted methods (i.e., without the need for internal standards and SIM). Very low-abundance NAEs, such as anandamide (C20:4), were

also detected by DMP, but only in FAAH(−/−) tissues, establishing a sensitivity limit for this method that was within 5–10-fold of the sensitivity of targeted LC–MS analysis. In addition to known NAEs, several heretofore unrecognized members of this lipid family were also found by DMP to be significantly elevated (~4–10-fold) in CNS tissues from FAAH(−/−) mice, including C16:1, C22:0, and C24:1 (Table 1 and Tables 2 and 3 of the Supporting Information). These data indicate that the NAE class of bioactive lipids is much larger than previously appreciated and support the notion that FAAH controls the levels of most, if not all, of these compounds *in vivo*.

Identification and Structural Characterization of a Novel Class of FAAH-Regulated Metabolites by DMP. In contrast to the NAEs, most of the other lipid species that were examined remained unchanged in FAAH(+/+) and FAAH(−/−) tissues, including free fatty acids, phospholipids, and ceramides (Figure 2, groups 1–3, respectively; also see Table 1 and Tables 2 and 3 of the Supporting Information). Interestingly, however, DMP detected a class of unknown metabolites that was dramatically elevated in the brains and spinal cords of FAAH(−/−) mice (Figure 2, group 5). These compounds, which were selectively observed in the negative ionization mode, were larger in mass than the NAEs (m/z 446–502 vs 296–438) and also appeared to be more hydrophilic based on earlier elution times on LC. To determine the structures of these novel FAAH-regulated metabolites, the most abundant representatives (m/z 446, 472, and 474) were purified by preparative HPLC and their exact masses determined using ESI-FTMS to be 446.3310, 472.3466, and 474.3623, affording predicted molecular formulas of $C_{24}H_{48}NO_4S$, $C_{26}H_{50}NO_4S$, and $C_{26}H_{52}NO_4S$, respectively (Figure 3A). The convergence of these metabolites to a consensus molecular formula (i.e., a conserved heteroatomic component, NO_4S , with a variable hydrocarbon portion) provided strong evidence that they belonged to the same structural class. Consistent with this premise, tandem MS analysis generated highly related fragmentation patterns for the 446, 472, and 474 metabolites, which included the serial loss of 14 mass units indicative of a lipid alkyl chain and a series of shared lower-molecular mass ions (m/z 80, 107, and 124) suggesting the presence of taurine (Figure 3B and Figure 1 of the Supporting Information), a bioactive amino acid highly enriched in the brain (13, 14). On the basis of these criteria and our knowledge of the general reactions catalyzed by FAAH (e.g., amide/ester hydrolysis), we postulated that this novel class of FAAH-regulated CNS lipids represented taurine-conjugated analogues of very long chain fatty acids [*N*-acyl taurines (NATs)], with the 446, 472, and 474 peaks corresponding to C22:0, C24:1, and C24:0 NATs, respectively. The levels of these very long chain NATs were elevated 4–40-fold in the brains and spinal cords of FAAH(−/−) mice (Table 1 and Table 3 of the Supporting Information). Shorter chain NATs (e.g., C18:1) were also identified in CNS tissue, but the levels of these lipids were only slightly elevated in the absence of FAAH (Table 5 of the Supporting Information).

Although taurine conjugates of sterols (bile salts) are well-established liver metabolites (15), to our knowledge, NATs have not previously been identified as natural products in the CNS or peripheral tissues of mammals. To confirm the structure of this unusual class of lipids, we prepared synthetic

samples of C18:1, C22:0, and C24:0 NATs for direct comparison with the natural products. The synthetic and natural substances exhibited identical retention times by LC (Figure 3C), similar mass ionization characteristics (e.g., selective ionization in the negative mode), and equivalent fragmentation patterns by MS/MS analysis (Figure 3B). Collectively, these analytical data provided strong evidence that the unknown FAAH-regulated CNS metabolites identified by DMP were NATs. We next quantified the levels of these lipids in tissues from FAAH(+/+) and FAAH(−/−) mice using targeted LC–MS methods. For these studies, a non-natural NAT (C17:0) was included as an internal standard and several endogenous NATs (chain lengths from 16 to 26) were measured concurrently by MS using SIM. The levels of very long chain NATs (\geq C22) were determined to be, in general, 10–40-fold elevated in CNS tissues from FAAH(−/−) mice, and these targeted data were all within 2-fold of the estimates obtained by DMP (Table 1 and Table 5 of the Supporting Information). Targeted MS studies also provided a measure of the absolute concentrations of NATs, revealing that these lipids were enriched in the CNS compared to peripheral tissues and exhibited an acyl chain distribution favoring very long chains (Tables 3 and 5 of the Supporting Information). This latter property contrasted sharply with the acyl chain distribution of NAEs, which was strongly weighted toward shorter chain species (e.g., C16:0, C18:0, and C18:1) (Figure 2 of the Supporting Information).

The Relative Activity that FAAH Displays for Lipid Substrates in Vitro Is Not Predictive of the Metabolites Regulated by This Enzyme in Vivo. Given that several of the lipids identified by DMP in the CNS of FAAH(−/−) mice represented novel natural products, we next tested whether these compounds were substrates for FAAH using a general LC–MS assay (Figure 4). FAAH was found to hydrolyze representative long and very long chain members of both the NAE and NAT classes of lipids, as well as members of the monoacylglycerol family of fatty acid esters (Figure 4 and Table 2). The relative k_{cat}/K_m values (specificity constants) for these substrates varied by more than 5 orders of magnitude (Table 2), with saturated, very long chain substrates being hydrolyzed by FAAH with markedly lower efficiencies than shorter chain unsaturated lipids. Interestingly, this *in vitro* substrate selectivity profile for FAAH did not correlate with the relative expression levels of its lipid substrates in CNS tissue from FAAH(+/+) and FAAH(−/−) mice. For example, the C24:0 NAT was a several hundred-fold poorer substrate for FAAH than the C18:1 NAT, but its level was more dramatically elevated in FAAH(−/−) tissues (Table 2). Similarly, FAAH hydrolyzed the C18:1 NAE with a specificity constant ~50000 times greater than that for the C22:0 NAE, yet both of these substrates were still highly upregulated (6–18-fold) in FAAH(−/−) CNS tissues (Table 2). Conversely, the fatty acid ester 2-OG, which was one of the most active FAAH substrates *in vitro* (Table 2), showed no change in expression in FAAH(−/−) mice (Table 1). These data indicate that the relative activity displayed by FAAH for substrates *in vitro* is not necessarily indicative of the specific contribution that this enzyme makes to the regulation of lipid metabolites *in vivo*, as even the levels of very poor substrates (e.g., C22:0 NAE and C24:0 NAT) were highly elevated in FAAH(−/−) mice, while other

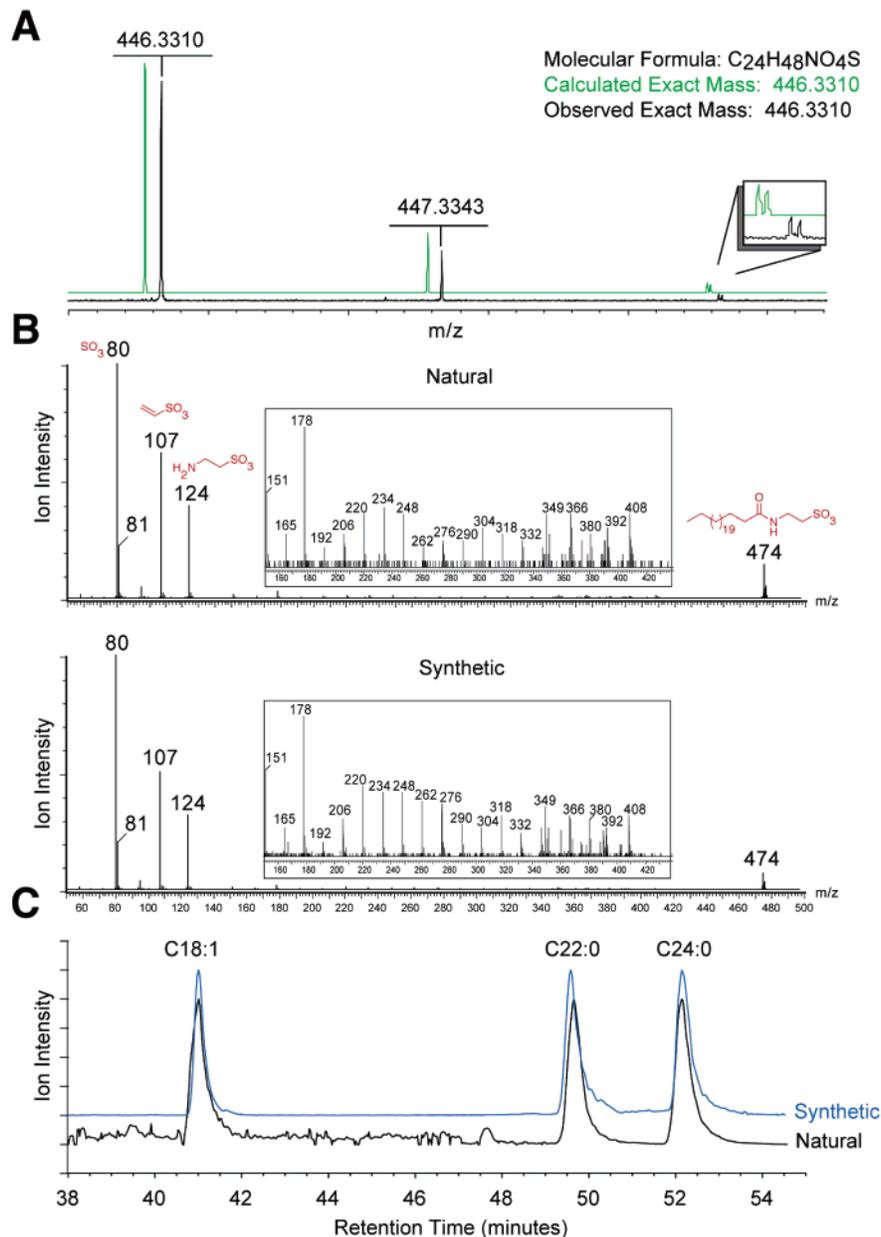


FIGURE 3: Chemical characterization of an unknown class of FAAH-regulated brain metabolites as *N*-acyl taurines (NATs). (A) Analysis of the *m/z* 446 metabolite by ESI-FTMS (Bruker APEX III instrument) provided an exact mass of 446.3310, corresponding to a molecular formula of $C_{24}H_{48}NO_4S$. The calculated isotope pattern for $C_{24}H_{48}NO_4S$ (green) overlaid well with the experimental spectrum, including the splitting pattern of the $M + 2$ isotope (inset) caused by differences in the mass between two ^{13}C isotopes and one ^{34}S isotope. Similar spectra were obtained for the *m/z* 472 and 474 metabolites, providing exact masses of $C_{26}H_{50}NO_4S$ (472.3466) and $C_{26}H_{52}NO_4S$ (474.3623), respectively (Table 6 of the Supporting Information). (B) MS/MS analysis of the natural *m/z* 474 metabolite (top trace) leading to its structural assignment as C24:0 NAT. MS/MS data were obtained on a Micromass Q-TOF instrument. Highlighted are prominent fragments corresponding to taurine (124), vinylsulfonic acid (107), and sulfur trioxide (80), as well as a pattern of progressive loss of 14 mass units from *m/z* 150–430 indicative of a fatty acyl chain (inset). This fragmentation spectrum matched closely the MS/MS data of a synthetic C24:0 NAT standard (bottom trace). (C) Comigration by LC-MS of natural and synthetic samples of C18:1, C22:0, and C24:0 NATs.

lipids that were excellent substrates *in vitro* (e.g., 2-OG) were unperturbed by the inactivation of FAAH.

DISCUSSION

The levels and distribution of metabolites in cells and tissues are regulated by an intricate biochemical network of enzymes acting in both coordination and competition. Accordingly, the assignment of endogenous substrates to enzymes requires analytical methods that can evaluate the function of these proteins directly in their native environment. Here, we have introduced a strategy termed discovery

metabolite profiling (DMP) for globally evaluating the metabolic consequences of enzyme inactivation *in vivo* and have applied this method to identify several physiological substrates for the mammalian enzyme fatty acid amide hydrolase (FAAH), including known signaling lipids (e.g., anandamide) and a novel family of nervous system-enriched natural products, the *N*-acyl taurines (NATs). DMP also determined that other lipid classes were unaffected by the inactivation of FAAH, indicating that this enzyme performs unique and specific roles in CNS metabolism. Importantly, however, these endogenous activities could not be directly

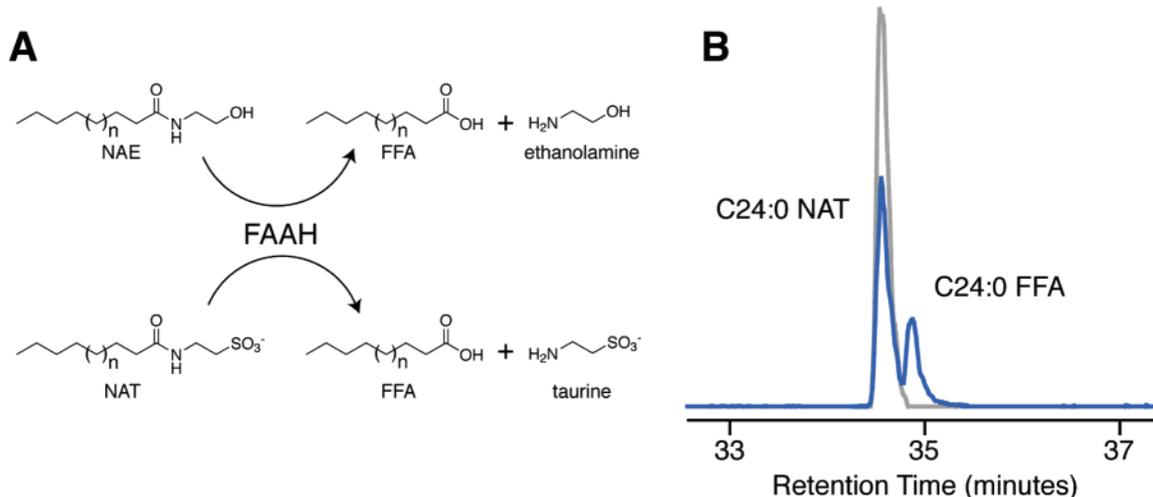


FIGURE 4: Determination that NATs are substrates for FAAH. (A) Scheme showing the FAAH-catalyzed hydrolysis of NAEs and NATs. (B) Representative LC-MS trace showing FAAH-catalyzed hydrolysis of C24:0 NAT to C24:0 free fatty acid (FFA): (blue) reaction that included NAT and FAAH and (gray) reaction that included NAT alone.

Table 2: Kinetic Analysis of FAAH with Endogenous Lipid Substrates^a

substrate	k_{cat} (s ⁻¹)	K_m (μ M)	k_{cat}/K_m (M ⁻¹ s ⁻¹)	$(k_{cat}/K_m)/(k_{cat}/K_m)_{C18:1NAE}$	FAAH(-/-)/FAAH(+/+) brain (spinal cord)
C18:1 NAE	8 ± 0.3	35 ± 4	2.3 × 10 ⁵	1	12.8 (17.7)
C18:1 MAG	3 ± 0.2	28 ± 6	1.1 × 10 ⁵	0.48	0.72 (0.91)
C18:1 NAT	1.8 ± 0.1	25 ± 5	7.2 × 10 ⁴	0.31	2.0 (2.9)
C22:0 NAE	≥ 0.0006	≥ 150	4.3	2 × 10 ⁻⁵	≥ 3.8 (6.3)
C24:0 NAT	≥ 0.017	≥ 150	110	5 × 10 ⁻⁴	26 (23)

^a NAEs, *N*-acyl ethanolamines; NATs, *N*-acyl taurines; MAGs, monoacylglycerols. Data represent averages ± standard errors (SE). For details about the FAAH substrate assay, see Experimental Procedures.

gleaned from an *in vitro* analysis of FAAH's substrate selectivity, which was unpredictable of the lipids regulated by this enzyme *in vivo*. We suggest that this discordance can be largely explained by the presence or absence of competing metabolic pathways *in vivo*. For example, the esterified lipid, 2-OG, which is an excellent FAAH substrate *in vitro*, but unaltered in FAAH(-/-) mice (Table 2), is also hydrolyzed by additional enzymes, such as monoacylglycerol lipase (16), suggesting that other metabolic pathways may control the levels of this lipid *in vivo*. Consistent with this notion, brain homogenates from FAAH(+/+) and FAAH(-/-) mice have been found to hydrolyze 2-OG at equivalent rates (17). On the other hand, poor substrates for FAAH *in vitro*, such as the very long chain NATs and NAEs, may still be regulated by this enzyme *in vivo* if alternative routes for catabolism are absent. Further complicating predictions of *in vivo* activity for enzymes is the potential for tissue-specific functions. For example, hormone-sensitive lipase, an enzyme originally considered to represent the primary cholesterol esterase in most cells and tissues, has recently been found to serve this function in testis and adipose tissue, but not in macrophages, where other (as yet unidentified) cholesterol esterases predominate (18). Collectively, these findings stress that, to understand the endogenous functions played by an enzyme, one must not only determine its catalytic properties *in vitro* but also distinguish which subset of these activities is uniquely attributable to this enzyme versus those that can be performed by other metabolic pathways *in vivo*.

In addition to analyzing genetic models of enzyme inactivation, DMP should also be suitable for evaluating the acute or chronic effects of enzyme inhibitors. In the case of

FAAH, such studies may have important biomedical implications, as this enzyme is an emerging therapeutic target for the treatment of pain (19) and neuropsychiatric disorders (20). FAAH inhibitors would be predicted to reduce pain and anxiety by elevating bioactive NAEs (e.g., anandamide) in the CNS; however, if NATs are similarly upregulated [as they were in FAAH(-/-) mice], one must consider whether these novel fatty acid amides are inert metabolites or signaling molecules in their own right. Given the key role that taurine plays as a chemical transmitter in the CNS (13, 14), combined with the selective enrichment of NATs in these tissues (Table 5 of the Supporting Information), we speculate that this novel family of lipids may serve a signaling function in the brain, possibly as a neuronal source for the local (FAAH-catalyzed) production of taurine, which is otherwise biosynthesized, transported, and released primarily by astrocytes (21, 22).

In summary, DMP addresses a problem of fundamental importance for all enzymes, the assignment of endogenous substrates, and can, in principle, be applied to any metabolic enzyme, even those that function in highly complex tissues such as the mammalian brain. The value of DMP over more conventional, targeted analytical methods is underscored by our finding of a structurally novel class of CNS lipids, the *N*-acyl taurines (NATs), as endogenous substrates of FAAH. The structure determination of NATs was expedited by our knowledge of the enzymatic mechanism of FAAH, highlighting an additional virtue of DMP, which provides a functional link between the proteome and the metabolome that adds significant predictive value to the often laborious process of characterizing unknown natural products. In our pursuit of FAAH substrates, we elected to analyze the organic

extracts of mouse tissues to enrich for lipophilic metabolites. However, other portions of the metabolome (e.g., aqueous soluble metabolites) can be similarly fractionated (23, 24) and thus should be amenable to analysis by DMP. Additionally, it is important to note that, while the natural products identified by DMP in this study constituted direct substrates for FAAH, secondary metabolites of enzymatic pathways may also be uncovered with this method. By forging specific connections between the proteome and the metabolome, DMP should elucidate endogenous functions for many enzymes and facilitate their integration into larger metabolic and signaling networks in the cell. The correlation of these biochemical outputs with changes in cell and organismal biology that accompany enzyme inactivation should, in turn, lead to a deeper understanding of the metabolic basis for a number of physiological and pathological states.

ACKNOWLEDGMENT

We thank T. Bartfai, R. A. Lerner, M. McKinney, M. P. Patricelli, P. Schimmel, and I. A. Wilson for helpful discussions, K. Masuda for preparation of recombinant FAAH, S. E. Palmer for assistance with figure preparation, A. E. Clement for assistance with tissue isolation, and M. Greig and B. Bolanos (Pfizer, La Jolla, CA) for their assistance and generous use of their Fourier transform mass spectrometer.

SUPPORTING INFORMATION AVAILABLE

Tables of ion intensities, a table of exact mass data for C22:0, C24:1, and C24:0, and three figures depicting MS/MS spectra of compounds from FAAH(-/-) mice, the acyl chain distribution of NAEs and NATs, and substrate concentration versus velocity plots. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Carpenter, A. E., and Sabatini, D. M. (2004) Systematic genome-wide screens of gene function, *Nat. Rev. Genet.* 5, 11–22.
- Boutros, M., Kiger, A. A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S. A., Consortium, H. F., Paro, R., and Perrimon, N. (2004) Genome-wide RNAi analysis of growth and viability in *Drosophila* cells, *Science* 303, 832–835.
- Mitchell, K. J., Pinson, K. I., Kelly, O. G., Brennan, J., Zupicich, J., Scherz, P., Leighton, P. A., Goodrich, L. V., Lu, X., Avery, B. J., Tate, P., Dill, K., Pangilinan, E., Wakenight, P., Tessier-Lavigne, M., and Skarnes, W. C. (2001) Functional analysis of secreted and transmembrane proteins critical to mouse development, *Nat. Genet.* 28, 241–249.
- Remy, I., and Michnick, S. W. (2001) Visualization of biochemical networks in living cells, *Proc. Natl. Acad. Sci. U.S.A.* 98, 7678–7683.
- Cascante, M., Boros, L. G., Comin-Anduix, B., de Atauri, P., Centelles, J. J., and Lee, P. W. (2002) Metabolic control analysis in drug discovery and disease, *Nat. Biotechnol.* 20, 243–249.
- Kobe, B., and Kemp, B. E. (1999) Active site-directed protein regulation, *Nature* 402, 373–376.
- Jones, P. M., Quinn, R., Fennessey, P. V., Tjoa, S., Goodman, S. I., Fiore, S., Burlina, A. B., Rinaldo, P., Boriack, R. L., and Bennett, M. J. (2000) Improved stable isotope dilution-gas chromatography–mass spectrometry method for serum or plasma free 3-hydroxy-fatty acids and its utility for the study of disorders of mitochondrial fatty acid β -oxidation, *Clin. Chem.* 46, 149–155.
- Giuffrida, A., Rodriguez de Fonseca, F., and Piomelli, D. (2000) Quantification of bioactive acylethanolamides in rat plasma by electrospray mass spectrometry, *Anal. Biochem.* 280, 87–93.
- Cravatt, B. F., Giang, D. K., Mayfield, S. P., Boger, D. L., Lerner, R. A., and Gilula, N. B. (1996) Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides, *Nature* 384, 83–87.
- Cravatt, B. F., Demarest, K., Patricelli, M. P., Bracey, M. H., Giang, D. K., Martin, B. R., and Lichtman, A. H. (2001) Supersensitivity to anandamide and enhanced endogenous cannabinoid signaling in mice lacking fatty acid amide hydrolase, *Proc. Natl. Acad. Sci. U.S.A.* 98, 9371–9376.
- Devane, W. A., Hanus, L., Breuer, A., Pertwee, R. G., Stevenson, L. A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., and Mechoulam, R. (1992) Isolation and structure of a brain constituent that binds to the cannabinoid receptor, *Science* 258, 1946–1949.
- Clement, A. B., Hawkins, E. G., Lichtman, A. H., and Cravatt, B. F. (2003) Increased seizure susceptibility and proconvulsant activity of anandamide in mice lacking fatty acid amide hydrolase, *J. Neurosci.* 23, 3916–3923.
- Renteria, R. C., Johnson, J., and Copenhagen, D. R. (2004) Need rods? Get glycine receptors and taurine, *Neuron* 41, 839–841.
- Sergeeva, O. A., Chepkova, A. N., Doreulee, N., Eriksson, K. S., Poelchen, W., Monnighoff, I., Heller-Stilb, B., Warskulat, U., Haussinger, D., and Haas, H. L. (2003) Taurine-induced long-lasting enhancement of synaptic transmission in mice: role of transporters, *J. Physiol.* 550, 911–919.
- Hay, D. W., and Carey, M. C. (1990) Chemical species of lipids in bile, *Hepatology* 12, 65–145.
- Dinh, T. P., Carpenter, D., Leslie, F. M., Freund, T. F., Katona, I., Sensi, S. L., Kathuria, S., and Piomelli, D. (2002) Brain monoglyceride lipase participating in endocannabinoid inactivation, *Proc. Natl. Acad. Sci. U.S.A.* 99, 10819–10824.
- Lichtman, A. H., Hawkins, E. G., Griffin, G., and Cravatt, B. F. (2002) Pharmacological activity of fatty acid amides is regulated, but not mediated, by fatty acid amide hydrolase in vivo, *J. Pharmacol. Exp. Ther.* 302, 73–79.
- Osuga, J., Ishibashi, S., Oka, T., Yagyu, H., Tozawa, R., Fujimoto, A., Shionoiri, F., Yahagi, N., Kraemer, F. B., Tsutsumi, O., and Yamada, N. (2000) Targeted disruption of hormone-sensitive lipase results in male sterility and adipocyte hypertrophy, but not in obesity, *Proc. Natl. Acad. Sci. U.S.A.* 97, 787–792.
- Cravatt, B. F., and Lichtman, A. H. (2003) Fatty acid amide hydrolase: an emerging therapeutic target in the endocannabinoid system, *Curr. Opin. Chem. Biol.* 7, 469–475.
- Gaetani, S., Cuomo, V., and Piomelli, D. (2003) Anandamide hydrolysis: a new target for anti-anxiety drugs? *Trends Mol. Med.* 9, 474–478.
- Tappaz, M., Almaghni, K., and Do, K. (1994) Cysteine sulfinate decarboxylase in brain: identification, characterization and immunocytochemical location in astrocytes, *Adv. Exp. Med. Biol.* 359, 257–268.
- Pow, D. V., Sullivan, R., Reye, P., and Hermanussen, S. (2002) Localization of taurine transporters, taurine, and ^3H taurine accumulation in the rat retina, pituitary, and brain, *Glia* 37, 153–168.
- Soo, E. C., Aubry, A. J., Logan, S. M., Guerry, P., Kelly, J. F., Young, N. M., and Thibault, P. (2004) Selective detection and identification of sugar nucleotides by CE-electrospray-MS and its application to bacterial metabolomics, *Anal. Chem.* 76, 619–626.
- Tolstikov, V. V., and Fiehn, O. (2002) Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry, *Anal. Biochem.* 301, 298–307.
- Bligh, E. G., and Dyer, W. J. (1959) A rapid method of total lipid extraction and purification, *Can. J. Med. Sci.* 37, 911–917.
- Folch, J., Lees, M., and Sloane Stanley, G. H. (1957) A simple method for the isolation and purification of total lipids from animal tissues, *J. Biol. Chem.* 226, 497–509.
- Patterson, B. W., Zhao, G., Elias, N., Hachey, D. L., and Klein, S. (1999) Validation of a new procedure to determine plasma fatty acid concentration and isotopic enrichment, *J. Lipid Res.* 40, 2118–2124.
- Patricelli, M. P., Lashuel, H. A., Giang, D. K., Kelly, J. W., and Cravatt, B. F. (1998) Comparative characterization of a wild type and transmembrane domain-deleted fatty acid amide hydrolase: identification of the transmembrane domain as a site for oligomerization, *Biochemistry* 37, 15177–15187.