Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites

William R. Wikoffa, Andrew T. Anforab, Jun Liuab, Peter G. Schultzb,1, Scott A. Lesleyb, Eric C. Petersb, and Gary Siuzdakt,a,1

aDepartment of Molecular Biology and Center for Mass Spectrometry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037; and bGenomics Institute of the Novartis Research Foundation, San Diego, CA 92121

Communicated by Steve A. Kay, University of California at San Diego, La Jolla, CA, December 19, 2008 (received for review December 12, 2008)

Although it has long been recognized that the enteric community of bacteria that inhabit the human distal intestinal track broadly impacts human health, the biochemical details that underlie these effects remain largely undefined. Here, we report a broad MS-based metabonomics study that demonstrates a surprisingly large effect of the gut “microbiome” on mammalian blood metabolites. Plasma extracts from germ-free mice were compared with samples from conventional (conv) animals by using various MS-based methods. Hundreds of features were detected in only 1 sample set, with the majority of these being unique to the conv animals, whereas ~10% of all features observed in both sample sets showed significant changes in their relative signal intensity. Amino acid metabolites were particularly affected. For example, the bacterial-mediated production of bioactive indole-containing metabolites derived from tryptophan such as indoxyl sulfate and the antioxidant indole-3-propionic acid (IPA) was impacted. Production of IPA was shown to be completely dependent on the presence of gut microflora and could be established by colonization with the bacterium Clostridium sporogenes. Multiple organic acids containing phenyl groups were also greatly increased in the presence of gut microbes. A broad, drug-like phase II metabolic response of the host to metabolites generated by the microbiome was observed, suggesting that the gut microflora has a direct impact on the drug metabolism capacity of the host. Together, these results suggest a significant interplay between bacterial and mammalian metabolism.

The human body is colonized by hundreds of trillions of microbes, which collectively possess hundreds of times as many genes as coded in the human genome. The combined genetic potential of the endogenous flora is referred to as the “microbiome” (1), and typically results in a mutualistic relationship between microbe and host. Normal host activities, including the processing of nutrients and the regulation of the immune system, are affected by the intestinal microbiome (2, 3), and the microbiome has been implicated in the pathogenesis of diseases such as nonalcoholic steatohepatitis (4), allergy (5), the formation of gallstones (6), and inflammatory bowel disease (7, 8). The composition of the gut microbiome is highly variable (9), and its diversity can be significantly affected by alterations in diet (2, 10) or antibiotic use (11). Alternatively, probiotic therapy is the attempt to alter the extant gut microbial environment through the ingestion of live consumable cultures of beneficial bacteria (12). Differences in commensal microflora are likely to impact human health and disease through any number of ways reflective of the complex nature of the microbiome itself.

Historically, classical microbiology methods including the isolation and culture of individual bacterial species associated with the gut were used to study microbial colonization of higher organisms. More recently, metagenomic techniques have been used to characterize both the composition and the potential physiological effects of entire microbial communities without having to culture individual community members. For example, genes for specific metabolic pathways such as amino acid and glycogen metabolism were found to be overrepresented in the microbiome of the distal gut, reinforcing the notion that human metabolism is an amalgamation of microbial and human attributes (13). Also, obese mice were found to extract energy from their food more efficiently compared with lean counterparts due to alterations in the composition of their gut microflora that resulted in an increased complement of genes for polysaccharide metabolism (10). It has also been observed that bile salt hydrolase encoding genes are enriched in the gut microbiome, and that enteric bacteria carry out a wide range of bile acid modifications (6, 14). These metagenomic studies suggest that the metabolites derived from this diverse microbial community can have a direct role in human health and disease. To date, metabolomics-based investigations of aspects of the impact of the microbiome on mammalian biochemistry have detailed changes in the levels of well-documented metabolites based primarily on NMR-based analysis and subsequent multivariate statistics of untargeted samples, such as urine, gut tissue, or feces extracts (15–17). Recently, this same group reported the multicompartmental effects of the microbiome on mammalian metabolism by using NMR-based analysis of urine and tissue extracts from both conventional (conv) and germ-free (GF) mice (18). Although extremely powerful, these studies provide only limited opportunity for the discovery of differences in unexpected or lower level metabolites.

Here, we demonstrate the large effect of the microbiome on mammalian plasma biochemistry. Specifically, a broad, untargeted, mass spectrometry-based profiling of serum from GF and conv mice demonstrates that a significantly large number of chemical species found in systemic circulation arise because of the presence of the microbiome, whereas at least 10% of all detectable endogenous circulating serum metabolites vary in concentration by at least 50% between the 2 mouse lines. Several microbiome-affected molecules identified in the serum of conv mice are either potentially harmful (uremic toxins) or beneficial (antioxidant) to the host. Also, we observed a broad, drug-like phase II metabolic response of the host to species generated by the microbiome, as evidenced by the exclusive presence of numerous sulfated, glycine-conjugated, and glucuronide adducts in the serum of conv mice.

Results and Discussion

Mass Spectrometry Reveals That the Microbiome Has a Broad Effect on Mouse Plasma Biochemistry. Mass spectrometry has become an increasing powerful tool for metabolomics studies due to its wide dynamic range, reproducible quantitative capabilities, and its ability to analyze samples of significant molecular complexity (19). However, the vast chemical diversity of the known metabolome creates numerous challenges for any MS-based “global” profiling effort, including analytical biases introduced by sample preparation, separation, and ionization techniques. Given these challenges, we sought to obtain an initial assessment of the extent of the effect of


The authors declare no conflict of interest.

1To whom correspondence may be addressed. E-mail: schultz@gnf.org or siuzdak@scripps.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812874106/DCSupplemental.

© 2009 by The National Academy of Sciences of the USA
the microbiome on mammalian metabolism by analyzing uniformly processed samples employing various ionization methods, because studies have shown that such multiple ionization mode strategies increase both the number and types of analytes observed (20). We also chose to analyze plasma samples, because this biofluid provides an excellent opportunity to assess the extent of interplay between bacterial metabolic and systemic human pathways.

In Fig. 1 we show the experimental workflow for the metabolomics studies performed. Ten plasma samples each from both individual GF and conv mice were subjected to methanol protein precipitation and subsequent centrifugation, because this technique was previously reported to be the most effective, straightforward, and reproducible extraction method for such samples (21). The resulting supernatants were each analyzed by 4 different methods; reversed-phase liquid chromatography coupled to (i) electrospray ionization (ESI) in the positive mode (+), (ii) ESI in the negative mode (−), (iii) atmospheric pressure chemical ionization (APCI) in the + mode, and (iv) gas chromatography after derivatization with trimethylsilyl chloride. The raw data from the individual analyses were subjected to nonlinear data alignment (22), and the results were analyzed for both global changes by using multivariate statistics to determine group separation, as well as univariate statistics to evaluate the number and percentage of features that are either unique or vary significantly between the 2 sample sets. As seen in Fig. 1B, principle component analysis (plotting principle component 1 vs. principle component 2) shows an excellent separation of the GF and conv sample groups analyzed under + ESI conditions. Similar separations were also seen for the −ESI + APCI, and GC-based analyses (Fig. S1). The majority of individual features detected were present in both the GF and conv samples. However, hundreds of features were detected in only 1 sample set, with the majority of these being unique to the conv samples. Also, as summarized in Table 1, ~10% of all features detected in both sample sets under any of the analysis conditions showed significant changes in their relative signal intensity (defined as a ≥1.5-fold change with a P ≤ 0.0001). It should be noted that a given molecule may be represented by several different features, such as the naturally occurring components of its isotopic cluster or nonspecific adduct ions. Nevertheless, these comparative results show that the microbiome exerts a dramatic effect on mammalian plasma biochemistry.

We next turned our focus toward the identification of those features that showed significant differences in level between the 2 sample sets. This task is by far the most laborious part of any untargeted MS-based metabolomics profiling study, involving the integrated analysis of accurate mass measurements, tandem MS fragmentation patterns, and literature/database searches to produce candidate structures that still require experimental verification. For the purposes of this article, an observed feature was not designated as unequivocally identified unless an authentic sample of the proposed species was independently analyzed in our laboratories and found to exhibit the same chromatographic and mass spectrometric properties as the feature in question. By comparison, highly probable identification status was assigned when the experimentally determined properties of a given feature matched those reported in the literature for a particular molecule, or were fully consistent with those reported for a close analog. Although this process remains a work in progress, the identifications produced to date as listed in Table 2 (and further detailed in Table S1) already exhibit several striking features. For example, despite the use of a single sample preparation method, we identified various chemically distinct compounds ranging from fatty acids to highly polar charged molecules, along with several groups of functionally related molecules.

**The Microbiome Affects the Diversity of Indole-Containing Compounds in Serum.** Multiple features that differed significantly between conv and GF mice correspond to compounds bearing indole moieties. For example, the plasma concentrations of tryptophan and N-acetyltryptophan in conv mice were 40 and 60% lower than in their respective GF counterparts. A subset of enteric bacteria express tryptophanase, which converts tryptophan to indole, pyru-

---

**Table 1. Summary of comparative metabolomics profiles using different MS-based methodologies**

<table>
<thead>
<tr>
<th></th>
<th>Total peaks</th>
<th>Total significant changed peaks*</th>
<th>Significant change conv &gt; GF</th>
<th>Significant change GF &gt; conv</th>
<th>Unique to conv</th>
<th>Unique to GF</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC ESI +</td>
<td>4,172</td>
<td>466, 11.2%</td>
<td>318, 68%</td>
<td>148, 32%</td>
<td>145</td>
<td>52</td>
</tr>
<tr>
<td>LC ESI −</td>
<td>4,132</td>
<td>271, 6.6%</td>
<td>168, 62%</td>
<td>103, 38%</td>
<td>100</td>
<td>61</td>
</tr>
<tr>
<td>LC APCI +</td>
<td>1,584</td>
<td>152, 9.5%</td>
<td>96, 63%</td>
<td>56, 37%</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>GCMS +</td>
<td>1,803</td>
<td>253, 14.0%</td>
<td>225, 88.9%</td>
<td>28, 11.1%</td>
<td>28</td>
<td>12</td>
</tr>
</tbody>
</table>

*Defined as having a fold-change ≥ 1.5 and P ≤ 0.0001.
species, have been administered as a test probiotic to dialysis non-indole-producing bacteria, such as various
transformation of the bacterial metabolite indole. Because trypto-
suffering from chronic kidney failure (25), arises from hepatic
was identified only in the serum of conv mice. This molecule, a
microbe interaction.
These findings suggest that the increased plasma serotonin levels
resulted from the metabolism of dietary tryptophan by the direct
action of enteric bacteria.
Interestingly, serotonin plasma levels were 2.8-fold higher in conv
animals. However, this increase in serotonin is difficult to attribute
to a direct metabolic transformation mediated by gut bacteria,
because the production of serotonin by characterized enteric bac-
terial species has not been described. Despite the well-documented
role of serotonin as a neurotransmitter, Enterochromaffin cells in
the gut are the largest source of serotonin production in the body,
and this molecule has been implicated in gastrointestinal patholo-
gies such as irritable bowel syndrome and Crohn’s disease (24).
These findings suggest that the increased plasma serotonin levels
observed could indirectly result from an as yet undefined host
microbe interaction.
As seen in Fig. 2, other indole-containing molecules were also
affected by the microflora. For example, indoxyl sulfate (indicam
was identified only in the serum of conv mice. This molecule, a
known nephrotoxin that accumulates in the blood of patients
suffering from chronic kidney failure (25), arises from hepatic
transformation of the bacterial metabolite indole. Because trypto-
phanase activity derives from only a subset of enteric bacteria,
non-indole-producing bacteria, such as various Bifidobacterium
species, have been administered as a test probiotic to dialysis
patients to decrease their plasma levels of indoxyl sulfate (26).
Conversely, a different set of enteric bacteria has been implicated
in the metabolic transformation of indole to indole-3-propionic acid
(IPA) (27). IPA, also identified only in the plasma of conv mice, has
been shown to be a powerful antioxidant (28), and is currently being
investigated as a possible treatment for Alzheimer’s disease (29, 30).

Although the presence of IPA in mammals has long been
ascribed in the literature to bacterial metabolic processes, this
conclusion was based on either the production of IPA in ex vivo
cultures of individual bacterial species (31) or observed decreases in
IPA levels in animals after administration of antibiotics (32). In
our own survey of IPA production by representative members of the
intestinal flora, only Clostridium sporogenes was found to produce
IPA in culture (Table S2). Based on these results, individual GF
mice were intentionally colonized with C. sporogenes strain ATCC
15579, and blood samples were taken at several intervals after
colonization. IPA was undetectable in the samples taken shortly
after introduction of the microbes, and was first observed in the
serum 5 days after colonization, reaching plateau values compara-
ble with conv mice by day 10. These colonization studies demonstrate
that the introduction of enteric bacteria capable of IPA
production in vivo into the gastrointestinal tract is sufficient to
introduce IPA into the bloodstream of the host. Also, other GF
animals were injected i.p. with either IPA (at 10, 20, or 40 mg/kg)
or sterile PBS vehicle, and their serum concentrations of IPA were
measured over time. As seen in Table S3, the high serum levels of
IPA observed 1 h after injection decreased more than 90% within
5 h, showing that IPA is rapidly cleared from the blood, and that its
presence in the serum of conv animals must result from continuous
production from 1 or more bacterial species associated with the
mammalian gut. The ability to measure the metabolic profiles of
animals selectively colonized with individual species or even com-
plex communities of defined bacterial populations promises to be
a powerful new tool for deconvoluting the contributions of the
various species comprising the microbiome and for dissecting
complex microbial-mammalian metabolic interactions.

Sulfation is Observed as a Prevalent Host Response to Bacterial
Metabolites. Serum metabolites from GF and conv mice differed
significantly in their level of sulfation. For example, both phenyl
and p-cresol metabolites were present only in conv animals (Fig. S2). These
molecules likely arose from the sulfation of direct bacterial metabo-
"TABLE 2. Summary of metabolites identified that showed
significant differences in concentration between the two
sample sets

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Fold change</th>
<th>P value</th>
<th>Compound and metabolism class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indole derivatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tryptophan</td>
<td>1.7, GF</td>
<td>8.42 × 10^-12</td>
<td>—</td>
</tr>
<tr>
<td>N-acetyltryptophan</td>
<td>2.4, GF</td>
<td>3.56 × 10^-4</td>
<td>—</td>
</tr>
<tr>
<td>Indoxyl sulfate</td>
<td>conv*</td>
<td>1.34 × 10^-7</td>
<td>Phase I: hydroxyl; phase II: sulfite</td>
</tr>
<tr>
<td>Serotonin</td>
<td>2.8, conv</td>
<td>1.27 × 10^-10</td>
<td>—</td>
</tr>
<tr>
<td>IPA</td>
<td>conv*</td>
<td>7.69 × 10^-7</td>
<td>Bacterial conjugation</td>
</tr>
<tr>
<td>Phenyl derivatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1.05, GF</td>
<td>0.3</td>
<td>—</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>1.44, GF</td>
<td>1.14 × 10^-4</td>
<td>—</td>
</tr>
<tr>
<td>Hippuric acid</td>
<td>17.4, conv</td>
<td>1.98 × 10^-9</td>
<td>Phase II: glycine</td>
</tr>
<tr>
<td>Phenylacetylglycine</td>
<td>3.8, conv</td>
<td>4.70 × 10^-8</td>
<td>Phase II: glycine</td>
</tr>
<tr>
<td>Phenyl sulfate</td>
<td>conv*</td>
<td>9.85 × 10^-7</td>
<td>Phase II: sulfite</td>
</tr>
<tr>
<td>p-Cresol sulfate</td>
<td>conv*</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Phenylpropionylglycine</td>
<td>conv*</td>
<td>3.07 × 10^-7</td>
<td>Phase II: glycine</td>
</tr>
<tr>
<td>Cinnamoylglycine</td>
<td>conv*</td>
<td>2.93 × 10^-7</td>
<td>Phase II: sulfite</td>
</tr>
<tr>
<td>Flavonoids</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equol sulfate</td>
<td>conv*</td>
<td>1.44 × 10^-5</td>
<td>Phase II: sulfite</td>
</tr>
<tr>
<td>Methyl equol sulfate</td>
<td>conv*</td>
<td>2.18 × 10^-6</td>
<td>Phase II: sulfite</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urate</td>
<td>1.99, conv</td>
<td>1.51 × 10^-6</td>
<td>—</td>
</tr>
<tr>
<td>Creatinine</td>
<td>1.08, conv</td>
<td>0.071</td>
<td>—</td>
</tr>
<tr>
<td>Dihydroyxoyquinoline</td>
<td>conv*</td>
<td>7.64 × 10^-6</td>
<td>Phase II: glucuronide</td>
</tr>
<tr>
<td>12-Hydroxy-5Z,8Z,10E,14Z,17Z-eicosapentaenoic acid</td>
<td>4.0, conv</td>
<td>8.20 × 10^-5</td>
<td>Fatty acid</td>
</tr>
<tr>
<td>3-Carboxy-4-methyl-5-pentyl-2-furanpropionic acid</td>
<td>3.4, conv</td>
<td>1.37 × 10^-6</td>
<td>Phase II: glucuronide</td>
</tr>
</tbody>
</table>

Fold change equals the fold difference in concentration observed between conv and GF samples, with the group indicating which is higher. Names in italics refer to compound identifications that are highly probable, as defined in the text.
*Compound observed only in a single group (conv or GF).
Equol is itself a reported metabolite of the dietary isoflavone dihydrodaidzein. Purported to have the potential to lower the incidence of prostate cancer (37), equol has been shown to arise exclusively as a result of bacterial metabolic processes (38), although only a small percentage of humans harbor the gut bacteria that affect this biotransformation (39). It is important to note that neither equol sulfate nor its methyl ester was listed in any metabolite database, again emphasizing the degree of complexity involving the microbiome that remains to be explored.

Other Phase II Metabolic Products Are also Greatly Elevated in Response to the Microbiome. We next sought to determine whether other bacterial metabolites were modified by mammalian chemical detoxification processes. Indeed, numerous products of another prevalent phase II drug metabolism mechanism, glycine conjugation, were found to be either exclusively present or present at much higher concentrations in conv animals. As shown in Fig. 4, the plasma concentration of hippuric acid was 17-fold higher in conv samples compared with their GF counterparts. Hippuric acid is produced by the conjugation of benzoic acid with glycine, a reaction that occurs not only in the liver (40) but also directly in the intestine (41) and kidney (42). Also, it has been reported that hippuric acid levels increase when animals are fed increased levels of phenylalanine (43). However, the concentration of circulating phenylalanine was not observed to differ significantly between the 2 sample sets, suggesting that unlike the presence of indican or IPA in conv mice, the observed increased levels of hippuric acid may not simply result from the removal of some of the dietary supply of an essential amino acid by the direct action of enteric bacteria. These findings again illustrate the broad interplay of mammalian and enteric bacterial metabolism.

In Fig. 4 we indicate several additional glycine conjugates identified exclusively in the serum of conv animals. For example, cinnamoylglycine was identified as being present only in conv samples despite not being listed in any metabolite database, illustrating the fact that our knowledge of the metabolome even in biofluids as well-studied as plasma is far from complete. Also, phenylpropionylglycine was also detected exclusively in conv sam-
This molecule almost certainly arose from the direct conjugation of glycine to phenylpropionic acid, a known metabolic product of anaerobic bacteria (36). However, in humans, phenylpropionic acid is instead normally converted to benzoic acid by the action of the \( \text{MCAD} \) (44). Thus, phenylpropionylglycine is detected only in the urine of MCAD-deficient patients, and has been used as a biomarker for the diagnosis of this condition (45). These results illustrate the specificity of host–bacteria mutualism, and suggest the likelihood of coevolution between symbionts and their hosts.

Several products of yet another mammalian chemical detoxification mechanism, glucuronidation, were also elevated in the presence of bacterial metabolism. For example, a glucuronide of 3-carboxy-4-methyl-5-pentyl-2-furanpropionic acid, a known uremic retention solute shown to inhibit the binding of salicylate and 5,5-diphenylhydantoin to albumin (46), was identified exclusively in the serum of conv animals. The greatly increased concentrations of this furan dicarboxylic acid, as well as other classic uremic retention solutes such as indoxyl sulfate, \( p \)-cresol sulfate, urate, and hippuric acid in the blood of these animals, point to the large burden that the metabolic processes of enteric bacteria impose on the chemical detoxification mechanism of their hosts.

**Conclusion**

In summary, we have demonstrated a significant effect of the microbiome on mammalian blood metabolites by applying an untargeted MS-based metabolomics approach to a germ-free mouse model system. Numerous circulating molecules were determined to arise exclusively in the presence of gut microflora, whereas \( \sim 10\% \) of commonly observed features differed significantly in concentration between GF and conv animals. Several pathways including the metabolic processing of indole-containing molecules were seen to particularly interact with the microbiome. Numerous molecules apparently resulting from phase II drug-like chemical processing of microbial metabolites were also significantly elevated. Such broad metabolomics profiling studies combined with the selective colonization of germ-free animals with defined bacterial population promise to be a powerful tool for the deconvolution of the complex interplay between mammalian and bacterial metabolic processes.

**Materials and Methods**

**Serum Sample Collection and Preparation.** Blood from 10 conv and 10 GF male Swiss Webster mice age 8–10 weeks was ordered from Taconic Farms. Both sets of animals were maintained on autoclaved NIH-31 chow. Blood was collected by
retro-ocular bleed into heparinized tubes, spun down at 3,000 g for 10 min, and the entire serum content of each animal was frozen immediately at −80 °C and shipped on dry ice. Metabolites were extracted from plasma with methanol. We added 4 volumes of cold methanol to 50 μL of plasma, vortexed, and incubated at −20 °C for 1 h. Samples were centrifuged 10 min at 14,000 × g, the supernatant was collected, and the centrifugation was repeated. The supernatant was dried in a SpeedVac and resuspended in 50 μL of 95:5 water:acetonitrile, and clarified for 5 min at 14,000 × g.

Metabolomics Profiling with ESI. The HPLC system consisted of a degasser, capillary pump, binary pump, and autosampler (1100 series, Agilent); 3 μL of extracted plasma maintained at 4 °C in a thermostated autosampler was applied to a capillary reversed-phase column (Zorbax C18 SB-300, Agilent) with dimensions 300-μm ID × 150-mm length. The flow rate was 4 μL/min with solvent A composed of water + 0.1% formic acid and solvent B composed of acetonitrile + 0.1% formic acid in positive ion mode. For negative ion mode data, 5 mM ammonium acetate was substituted for the formic acid in solvents A and B. The gradient consisted of 5% B for 5 min, followed by a gradient to 95% B over 45 min, held at 95% B for 5 min, and a reequilibration at 5% B for 10 min.

To reduce systematic error associated with instrumental drift, samples were run in an order that alternated between wild type and knockout. Data were collected in + and −ESI mode in separate runs on a TOF (Agilent 6210) operated in full scan mode from 100 to 1,000 m/z. The capillary voltage was 3,500 (±2,000 V) with a scan rate of 0.5 scan per second; the nebulizer flow rate was 12 L/hr.

Data Analysis and Statistics. Data in instrument specific format (.wiff) were converted to common data format (.cdf) files. The program XCMS (22) was used for nonlinear alignment of the data in the time domain and automatic integrat-


23. Botsford JL, Demos RD (1972) Escherichia coli tryptophanase in the enteric environ-

1. Wikoff W, et al. PNAS March 10, 2009 vol. 106 no. 10 3703


