

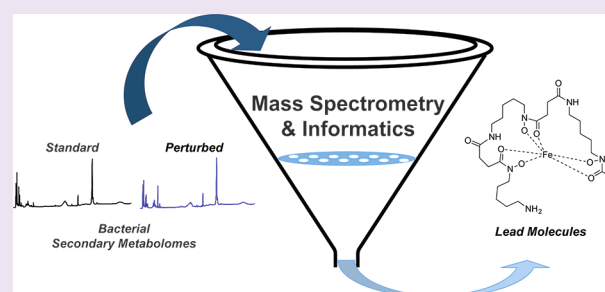
Integrated Metabolomics Approach Facilitates Discovery of an Unpredicted Natural Product Suite from *Streptomyces coelicolor* M145

Ashley M. Sidebottom,[†] Andrew R. Johnson,[†] Jonathan A. Karty,[†] Darci J. Trader,[†] and Erin E. Carlson^{*,†,‡}

[†]Department of Chemistry and [‡]Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, Indiana 47405, United States

S Supporting Information

ABSTRACT: Natural products exhibit a broad range of biological properties and have been a crucial source of therapeutic agents and novel scaffolds. Although bacterial secondary metabolomes are widely explored, they remain incompletely cataloged by current isolation and characterization strategies. To identify metabolites residing in unexplored chemical space, we have developed an integrated discovery approach that combines bacterial growth perturbation, accurate mass spectrometry, comparative mass spectra data analysis, and fragmentation spectra clustering for the identification of low-abundant, novel compounds from complex biological matrices. In this investigation, we analyzed the secreted metabolome of the extensively studied Actinomycete, *Streptomyces coelicolor* M145, and discovered a low-abundant suite of 15 trihydroxamate, amphiphilic siderophores. Compounds in this class have primarily been observed in marine microorganisms making their detection in the soil-dwelling *S. coelicolor* M145 significant. At least 10 of these ferrioxamine-based molecules are not known to be produced by any organism, and none have previously been detected from *S. coelicolor* M145. In addition, we confirmed the production of ferrioxamine D₁, a relatively hydrophilic family member that has not been shown to be biosynthesized by this organism. The identified molecules are part of only a small list of secondary metabolites that have been discovered since sequencing of *S. coelicolor* M145 revealed that it possessed numerous putative secondary metabolite-producing gene clusters with no known metabolites. Thus, the identified siderophores represent the unexplored metabolic potential of both well-studied and new organisms that could be uncovered with our sensitive and robust approach.



Discovery of natural products from the Actinomycete genus has been extremely successful because of the high propensity of these organisms to generate natural products. As a critical source for new drugs and chemical scaffolds, Actinomycetes devote much of their immense genome (>8 Mb) to the production of secondary metabolites, compounds that have led to the development of over half of all FDA-approved microbially derived drugs.¹ In 2003, *Streptomyces coelicolor* M145 became the largest completely sequenced bacterial genome and is an important model organism for this genus.² Prior to sequencing, this organism was reported to produce a number of secondary metabolites including actinorhodin, hopanoids, and prodiginines.² However, sequencing revealed approximately 18 additional orphan gene clusters that are thought to encode for enzymes that likely produce polyketides, chalcones, fatty acids, and siderophores along with several unknown compounds.^{2–4} Efforts to characterize the products of orphan clusters have utilized numerous strategies including genome mining, gene inactivation and targeting,^{5–7} heterologous expression,^{8,9} and/or the one strain many compounds (OSMAC) approach.^{10,11} Despite the

success of these methods, a significant proportion of the metabolites constructed by these clusters remain unidentified.

Compound discovery efforts are impaired by multiple factors; however, the analysis methods utilized to evaluate crude extracts represent a major barrier. Secondary metabolites can range in size (<100–2,000 Da), core-structure (e.g., lipid, glycopeptide), and physicochemical properties, making development of general analytical methods difficult. Accordingly, targeted metabolomics strategies are often designed for particular compound classes that are expected to be produced by a gene(s) of interest.¹² This approach requires *a priori* knowledge and has been important for answering specific biological questions¹³ but does not enable a wide breadth of profiling coverage. In contrast, untargeted strategies utilize robust approaches for the detection of a broad range of compounds but are plagued by two major issues: the vast dynamic range over which secondary metabolites are produced

Received: April 24, 2013

Accepted: June 18, 2013

Published: June 18, 2013

and the overwhelming amount of data generated by these efforts.¹⁴ Liquid chromatography (LC) coupled with ultraviolet (UV) detection has traditionally dominated discovery efforts but can generally detect only highly abundant compounds and yields little structural information for dereplication, or the elimination of known compounds.¹⁵ The recent incorporation of mass spectrometry (MS)-based platforms has increased the dynamic range of compound detection and dramatically decreased the quantity of material required (i.e., detection limit). Despite these advantages, a major challenge of untargeted MS-based analysis is processing and prioritizing the data produced; a LC–MS separation of the secondary metabolites from a single organism can yield hundreds of features. Unlike other -omics fields, comprehensive MS databases are not available for small molecules, meaning that compound deconvolution and identification often requires extensive searching of individual databases that represent only subsets of the metabolome from various organisms. In addition, most databases do not contain fragmentation spectra, which are invaluable for small molecule structure elucidation and unambiguous dereplication. To address these difficulties, we sought to integrate tools for both data acquisition and processing to yield an efficient and untargeted approach capable of quickly detecting, dereplicating, and characterizing novel secondary metabolites.

For our novel discovery workflow, we united accurate mass data acquisition, comparative MS analysis software (XCMS),^{16,17} tandem mass spectrometry analysis,¹⁸ and a molecular networking platform,¹⁹ with the goal of revealing unknown compounds from *S. coelicolor* M145; the data processing workflow is shown in Figure 1. The organism was

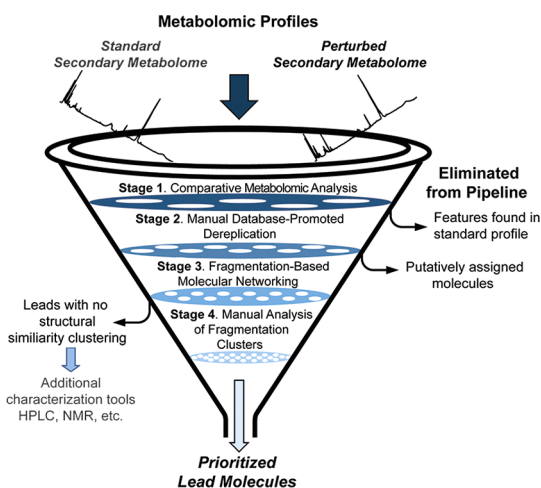


Figure 1. Integration of several tools yields a powerful metabolite discovery platform. Crude secondary metabolite samples are assessed using accurate mass spectrometry. These data undergo comparative analysis, and features found to be in both sample types are eliminated (Stage 1). Ranking of the remaining leads is performed by manual database searching and deprioritization of features with putative identifications (Stage 2). Compounds remaining after Stage 2 are the novel natural product leads. Fragmentation data (MS^2) are obtained for the perturbed metabolite sample and evaluated using molecular networking algorithms to highlight structurally related compounds (Stage 3). Fragmentation clusters containing lead molecules are manually assessed to determine if they contain known natural products (Stage 4). Leads found to cluster with knowns, suggesting structural similarity, are prioritized for further characterization.

grown under normal and stress-inducing conditions (28 and 37 °C incubation) for the generation of standard (28 °C) and perturbed (37 °C) metabolite profiles. The secreted metabolomes were subjected to liquid chromatography–time-of-flight-MS (LC-TOF-MS; Figure 1, input data). This type of analysis typically results in detection of >150 features (i.e., unique compounds and adducts, $M + H^+$, $M + NH_4^+$), creating a major bottleneck in previous discovery efforts because further characterization of this many leads would be impractical.²⁰ To overcome this obstacle, we prioritized and eliminated compounds from our workflow by processing the data with the comparative analysis software, XCMS (Figure 1, Stage 1). This software is designed to detect differences between sample groups and has most commonly been used to monitor primary metabolite levels.²¹ We postulated that compounds produced under typical growth conditions were more likely to be known and treated them as the “standard metabolome” of *S. coelicolor* M145. Accordingly, the two secondary metabolomes were compared, and >100 features were removed from the discovery pipeline due to the high likelihood that they were known compounds. Elimination of these features dramatically decreased the number of compounds that we then subjected to extensive manual database searching (Figure 1, Stage 2). All features surviving Stage 2 were deemed leads, as they have a high likelihood of being novel natural products. Next, tandem mass spectrometry was used to generate fragmentation profiles for the lead molecules. These spectra were compiled into a molecular network, which clusters compounds based on fragmentation similarity (common losses), putatively indicating structural resemblance (Figure 1, Stage 3).¹⁹ Leads not found to cluster are expected to be novel scaffolds but will likely require full purification for structural elucidation. Alternatively, clustered leads may be more readily identified if a common structural motif and/or known natural product is present within the cluster, either of which could simplify the structural elucidation process (Figure 1, Stage 4).

With the combination of these tools, we discovered a suite of 15 amphiphilic siderophores previously unknown to be produced by the model organism, *S. coelicolor* M145. These molecules contain a variable lipid tail and are structurally related to the formerly characterized compound, ferrioxamine B. We also confirmed the production of a hydrophilic family member, ferrioxamine D₁, which differs from B by an acetyl group. The identified siderophores were produced at detectable levels only when the organism was grown at elevated temperature, suggesting that their synthesis is stimulated in response to stress. We estimate that these molecules are present at <170 ng/L after initial broth extraction, illustrating the capacity of our approach to uncover low-abundant compounds. Our ability to reveal an entire family of molecules unknown to *S. coelicolor* M145 despite extensive study is a powerful illustration that the integration of multiple analytical and informatics approaches will accelerate exploration of the natural products repertoire of myriad organisms.

RESULTS AND DISCUSSION

Perturbation of *S. coelicolor* M145 Metabolome.

Bacterial metabolomes remain a valuable source of novel compounds and most organisms are likely to produce many more molecules than have been observed. Intermittent production coupled with an inability to detect low abundance species hinders their identification. To influence compound production, we cultured *S. coelicolor* M145 in various growth

environments (OSMAC).¹⁰ The presence of the blue natural product actinorhodin is considered a gross read-out of secondary metabolite production.²² Cultures grown at 37 °C (standard lab culture is at 28 °C) were not colored with this antibiotic indicating the potential for altered natural product production and were selected as our lead sample. To confirm differing metabolic profiles, data were obtained from the standard growth and elevated temperature growth samples using accurate mass spectrometry (UPLC-ESI-TOF). Manual examination of these data confirmed that features found over the entire intensity range of the acquired data were variable between sample types, indicating the successful perturbation of *S. coelicolor* M145 metabolite production (Supplementary Figure 1).

Comparative Analysis of Secondary Metabolome.

Quantitative comparison of the produced metabolites was completed with an open-source, platform-independent R-package, XCMS.¹⁶ Incorporation of this comparative analysis strategy was crucial for compound deconvolution. With samples containing between 150 and 200 MS features, complete characterization of all detected species would be prohibitive. We utilized a common growth condition, *S. coelicolor* M145 28 °C in ISP medium 2, as a standard profile for comparison against samples grown under a perturbed condition (37 °C; Supplementary Figure 1). XCMS analysis provided statistical information about the differences between the two sample types including *p*-values, fold change, and LC retention time comparisons. Features deemed to be present in both samples met the following criteria: *p*-value >0.01, low fold change (<20), and <0.2 min deviation in LC retention time. We defined a unique feature as one with a *p*-value of <0.01 and fold change of 20 or higher. A high percentage of detected metabolites were produced in both sample types and removed from the discovery pipeline; 129 features were detected, and of them, 104 were eliminated (~80%) leaving 25 compounds to be further explored (Figure 1, Stage 1).

Manual analysis of the mass spectrum for each of the 25 lead features was completed to confirm the comparative XCMS study. We required that features were distinguishable from noise ($\geq 10^3$ ion intensity on UPLC-ESI-TOF) and displayed realistic isotope distributions. Those features not meeting these specifications were eliminated because of the high possibility that they were false positives from the peak-picking algorithm. In addition, the ionization state and presence of adducts (e.g., $M + Na^+$) was assessed to verify that the correct parent ion species was selected. Representative data obtained for compounds deemed similar and different between the two sample types are in Figure 2, panels a and b, respectively. Next, leads were vetted with preliminary database screening (Figure 1, Stage 2; Supplementary Table 1). If reasonable putative identifications were found for a feature, it was deprioritized, which was the case for 20 of the 25 leads. In total, utilization of a bacterial standard metabolome to assist in lead prioritization followed by manual database assessment enabled removal of ~96% of the features detected, while efficiently highlighting the five most likely candidates for novel compound discovery. We next sought to incorporate a tool that aided structural characterization efforts of our lead compounds within the context of the crude biological matrix.

Identification of Novel Compound Suite. The remaining lead features were further prioritized based on their fragmentation profiles (Figure 1, Stage 3). Comparative studies have shown that members of particular compound classes, such

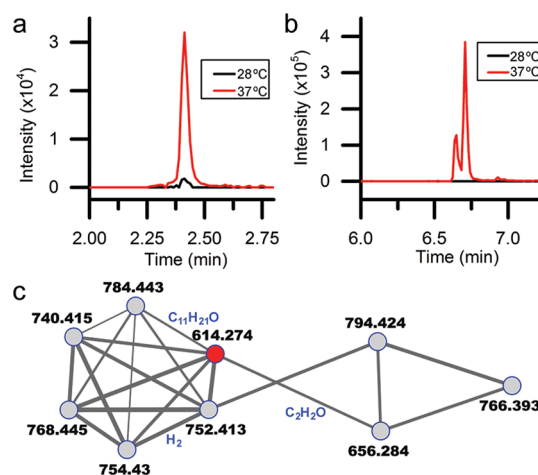


Figure 2. Analysis of *S. coelicolor* M145 grown under normal and perturbing conditions yielded many features likely to represent novel molecules. (a) Sample comparison resulted in detection of molecules present under both growth conditions with altered production levels ($m/z = 614.27$, fold change at 37 °C = 12.2). (b) Identification of features that were unique to the culture grown at 37 °C ($m/z = 754.44$) also occurred. (c) A prioritized lead, $m/z = 754.44$, clustered to several other compounds. Manual database searching revealed that one member of this network was the known natural product, ferrioxamine B, which was produced by the organism at both temperatures ($m/z = 614.27$, part a).

as macrolides or lipids, fragment similarly by collision-induced dissociation (CID) and common losses or fragments can be applied to the classification of unknown structures.¹⁹ The use of this preliminary structure elucidation strategy eliminated the need for purification of each compound and was key in providing additional information about lead molecule substructure. We concentrated our efforts on compounds that were found in a fragmentation cluster, as we anticipated that this information may expedite their structural elucidation. Alternatively, compounds that were not clustered with other features could be prioritized as they have a higher likelihood of representing novel scaffolds, but their full characterization would likely require purification.

Although fragmentation patterns are a powerful small molecule characterization tool,^{23,24} manual interpretation of these data is the most common form of analysis, which inevitably limits throughput. Additionally, comprehensive tandem mass spectral databases for bacterial metabolites do not currently exist. On account of these difficulties, we evaluated the fragmentation (MS^2) spectra with molecular networking,¹⁹ which aligns MS^2 spectra on the basis of common losses for the visualization and classification of secondary metabolites.²³ Data acquired on a nanoLC-ESI-linear ion trap quadrupole-Orbitrap from the *S. coelicolor* M145 37 °C extract was used to generate a molecular network utilizing a recently published procedure.^{19,25} Briefly, fragmentation spectra were compared pairwise accounting for fragment masses and relative intensities.²⁶ Subnetworks or clusters are generated only from data obtained from the analyzed sample and do not incorporate database information. In the resulting spectral map, each node represents a parent ion, connected on the basis of fragmentation similarity (Figure 2, panel c and Supplementary Figure 1). Thus, clustered molecules likely have structural resemblance, greatly simplifying the structure elucidation process. In particular, comparison of the parent

ions in the cluster enables estimation of the differences between their structures, such as an unsaturation or presence of an acetyl group (Figure 2, panel c blue formulas)

One of our leads ($m/z = 754.44$) was found to cluster with a feature whose parent mass was similar to that of a molecule found in the METLIN database ($m/z = 614.27$).²⁷ This feature was putatively identified as the siderophore, ferrioxamine B, a known *S. coelicolor* M145 metabolite and the metal-bound form of the drug Desferal, which is used to treat iron toxicity. Comparison of our data with the fragmentation spectra obtained for this metabolite in previous studies confirmed the presence of this trihydroxamate-containing siderophore in our node ($m/z = 614.27$, $[M - 2H + Fe]^+$, $C_{25}H_{46}N_6O_8Fe^+$).²⁸ Examination of the initial XCMS data verified that ferrioxamine B was produced under both growth conditions but that its production was upregulated at higher temperature (Figure 2, panel a, 12.2-fold increase). Upon confirmation of this molecular scaffold, we determined that ferrioxamine D₁ ($m/z = 656.28$, $[M - 2H + Fe]^+$, $C_{27}H_{48}N_6O_9Fe^+$), a related siderophore not previously seen in this organism but identified in *Streptomyces pilosus* (Supplementary Figure 2), was also present in our cluster.²⁹ Ultimately, we discerned that our lead was networked to six other unknown features, suggesting the discovery of an entire suite of structurally related compounds (Figure 2, panel c). Manual spectra analyses confirmed that all unknowns in the cluster possessed fragmentation losses similar to those of ferrioxamine B, revealing that they share a common structural core (Figure 3, and Supplementary Figures 5–7).³⁰ Further manual searching of the MS² data uncovered seven additional molecules with similar fragmentation loss patterns, indicating the compound suite likely contained 17 members, only one of which, ferrioxamine B, was produced at detectable levels when *S. coelicolor* M145 was grown under standard laboratory conditions.

Characterization of Siderophore Suite. Ferrioxamine B, which is both commercially available and has been fully characterized, was utilized as a standard in structure elucidation efforts (Figure 3, panels a and b, compound 1; Supplementary Figure 2).^{31,32} Ferrioxamine D₁, which differs by an acetyl group, was readily identified from its fragmentation pattern (Figure 4, panel a, compound 5 and Supplementary Figure 2). Fragmentation of the remaining 15 molecules indicated that they possess the same core structure as ferrioxamine B with common losses of 17, 118, and 200 but are functionalized with a variable lipid tail. The molecular formulas for these tails indicate that there are three classes of acyl groups present: fully saturated, monounsaturated, and monohydroxylated (C11–C15 chain lengths found for all three tail types; Supplementary Figure 3). Representative fragmentation spectra are shown, one from each acyl chain type (Figure 3, panels c and e, compounds 2–4), and complete fragmentation data for all compounds are provided in Supplementary Figures 5–7. Most spectra contain peaks at $m/z = 454$ and/or $m/z = 572$, fragments that can only form if the acyl chain is located at the indicated position. These spectra are consistent with data obtained from related acylated ferrioxamine derivatives found in other studies.^{33,34}

All detected compounds were observed bound to iron, evidenced by the distinctive isotopic ratio of this element, providing further indication that they were structurally related hydroxamate-containing siderophores (Supplementary Figure 4). None of the 17 molecules were observed in the desferrioxamine forms (without iron) under either growth condition. On the basis of a standard curve generated with

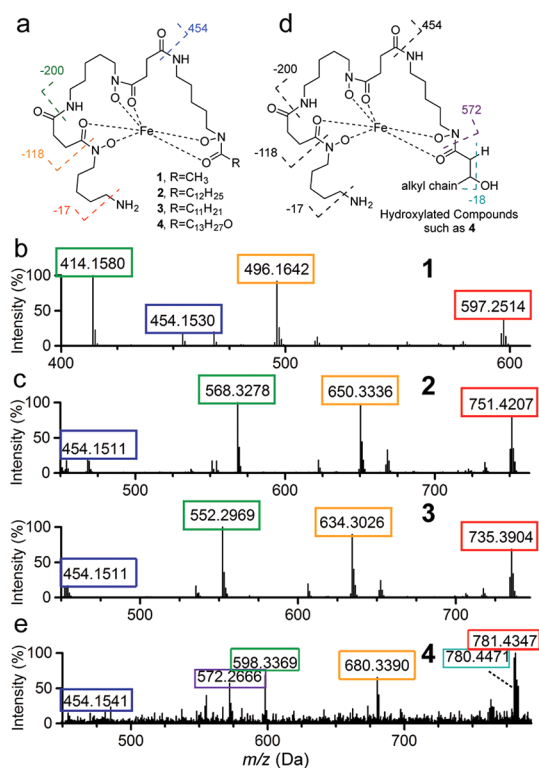


Figure 3. Accurate mass fragmentation (MS²) data obtained from ferrioxamine B and three representative compounds. (a) Structure of ferrioxamine B (1) and three amphiphilic ferrioxamine family members, C13 with saturated acyl tail (2), C12 with monounsaturated acyl tail (3), and C14 with monohydroxylated acyl tail (4). In total, 15 amphiphilic ferrioxamines were identified. (b) Fragmentation spectrum of ferrioxamine B is identical to known data and exhibits characteristic losses of $m/z = 17$, 118, and 200. Additionally, a species of $m/z = 454$ is observed, which is indicative of the core siderophore structure. (c) The characteristic fragmentation pattern was seen in the unknowns, suggesting that they possess the ferrioxamine core. Observation of the fragment at 454 indicates that the lipid tails appended to these molecules must be at the labeled position (part a). (d and e) Compounds predicted to possess a monohydroxylated acyl tail showed a mixed fragmentation pattern containing species from both the ferrioxamine core and additional fragments due to dehydration (-18) and cleavage of a hydroxamate resulting in loss of the lipid tail ($m/z = 572$).

ferrioxamine B, we estimate that ferrioxamine D₁ and the other identified metabolites are produced at <170 ng/L, making purification of this compound suite for full characterization of the acyl tails exceptionally difficult, especially given that iron is a paramagnetic element and would complicate assessment by NMR (iron-siderophore affinity $K_d \approx 10^{-23-31}$ M).³⁵ On the basis of their molecular formulas, we can discern that at least 10 of the 15 amphiphilic siderophores were previously undescribed and none of the molecules are found in a metabolite database. The remaining five compounds have the same acyl tail composition as several reported ferrioxamines, C15 promicroferrioxamine from *Promicromonosporaceae* strain SIO-11,³³ C11 promicroferrioxamine from *Micrococcus luteus* KLE1011,³⁴ and C11–C15 microferrioxamines from *Micromonospora* sp.³⁶

Amphiphilic Siderophores from a Terrestrial Organism. Trihydroxamate-containing siderophores are ubiquitous in bacteria and essential for transport of iron. In fact, the ferrioxamines are the most commonly produced siderophore among terrestrial Actinomycetes.^{28,37,38} These compounds

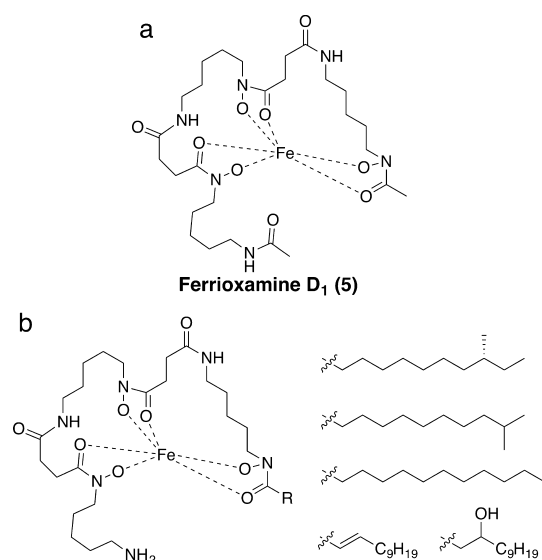


Figure 4. A suite of low-abundant trihydroxamate siderophores was identified from *S. coelicolor* M145. (a) The relatively hydrophilic ferrioxamine analogue, D₁, was in our molecular network. This compound has been observed in other organisms but not noted in *S. coelicolor* M145. (b) A suite of 15 amphiphilic ferrioxamines was discovered. The acyl chain conformation is likely to be branched; however, a linear lipid chain is possible. The location of the unsaturation and hydroxylation is postulated on the basis of acyl chains previously seen in siderophores.

contain three amides that are functionalized with hydroxamic acid groups that together form a hexadentate complex with Fe(III) to facilitate its delivery into cells.³⁹ Previous studies have demonstrated that ferrioxamine B, G₁, and E are produced by *S. coelicolor* M145.^{28,40} Here, we report production of ferrioxamine B and 16 previously unobserved trihydroxamate siderophores. Ferrioxamines G₁ and E, which have been detected when the organism was grown in iron-deficient media, were not seen in our investigations. The siderophores discovered in this study were produced under unique environmental conditions (37 °C) and at low levels (<170 ng/L). At this temperature, we observed a high mycelium density, which would result in low levels of available nutrients, likely prompting the production of these iron scavengers. Siderophore biosynthesis is well-known to respond to alterations in growth temperature, pH, nutrient supply, and the presence of neighboring bacterial species.^{33,34,41,42}

Interestingly, the discovered suite of molecules was largely composed of amphiphilic compounds with acyl chains containing between 11 and 15 carbons. In general, siderophores with amphiphilic character have been found in marine organisms^{43–45} but are rarely seen from terrestrial sources such as *S. coelicolor* M145. It has been postulated that the variable acyl chain lengths observed in suites of siderophores enable association of these compounds with a range of affinities to the cell membrane and facilitate transfer of iron from more hydrophilic molecules, which are soluble in the surrounding aqueous environment, to more hydrophobic siderophores that are closely associated with the cells.⁴⁵ An amphiphilic siderophore suite such as this is unanticipated from *S. coelicolor* M145 because it has been classified a terrestrial organism. However, when this organism was first sequenced, it was noted that it possessed the potential for survival in waterlogged environments, including gene clusters that resemble factors that

aid in buoyancy.² Our discovery of compounds primarily associated with marine organisms is consistent with this hypothesis. It is likely that growth of this organism in liquid culture, combined with application of stress-inducing conditions, stimulated the production of these siderophores. We tracked the production of ferrioxamine B and *m/z* = 754.44 (original lead; C₃₅H₆₅N₆O₈Fe⁺; C12 saturated acyl tail) over the course of the 3-week growth period. Desferrioxamine B biosynthesis started after approximately 2 weeks, while the amphiphilic siderophores were not detected until a week later (data not shown). These data suggest that depletion of the nutrients available to the organism over time may be a factor in stimulating the production of these iron chelators.

The identified compounds are biosynthesized in extremely low quantities, making their full characterization difficult. Strategies to increase their production were attempted and include growth of the organism in low-iron or iron-deficient media;²⁹ however, no measurable difference in siderophore concentration was observed. We also examined nutrient-poor culturing conditions and shorter or longer growth periods but again observed no siderophore increases. Heightened compound production will likely require use of other established strategies such as genetic manipulation,⁴⁶ co-culturing of *S. coelicolor*,⁴² and/or further variations to the growth media such as pH or addition of small molecule perturbants.^{47,48}

With the accurate mass and fragmentation data that we have obtained for this ferrioxamine suite, we determined the molecular composition of each acyl tail and can postulate about the conformation of the atoms within the chain. It is highly likely that the chains are not fully linear but instead contain a methyl branch at the ω terminus, based on previous studies that indicate *Streptomyces* utilize mostly branched starter units for fatty acid biosynthesis (80–90% of total fatty acid content; Figure 4, panel b).^{49–51} Indeed, other ferrioxamines with the same molecular formula that have been identified from marine organisms were all determined to possess branched lipid chains.^{33,34,36} We also anticipate that the mono-unsaturated chains bare the alkene at the position α to the carbonyl because other trihydroxamate-containing siderophores have been discovered with an unsaturation at this position.⁵² Finally, although no closely related compounds are known to contain a hydroxyl group in their acyl chain, several amphiphilic peptide-based siderophores have been shown to be hydroxylated at the β position, suggesting that our compounds could also be functionalized at this site.⁵³

In this study, we demonstrate that with our integrated discovery platform, previously unpredicted, low-abundant compounds can be identified from complex bacterial matrices, as shown with the highly characterized organism *S. coelicolor* M145. Current untargeted discovery techniques are largely dependent on unique UV signals, the presence of bioactivity, and/or extensive database searching to enable facile molecule detection and dereplication. Instead, we utilized mass spectrometry in combination with several data analysis tools to facilitate novel compound identification. In our studies, we performed deconvolution of the ~150 detected features by comparative analysis (XCMS) between a *S. coelicolor* M145 standard metabolome and the perturbed sample of interest, as an alternative to exhaustive searches to validate leads. Features present only in the stress-induced secondary metabolome were prioritized in our discovery workflow (~20% of detected features). Leads were further ranked by utilization of a mass spectral fragmentation data (MS²) networking protocol that

enabled rapid identification of structurally related metabolites. This strategy facilitated the discovery of 16 siderophores not previously known to be produced by *S. coelicolor* M145. Intriguingly, the identified compounds comprise a suite of amphiphilic siderophores, molecules that are often generated by marine organisms, making their detection in the soil-dwelling *S. coelicolor* unexpected. We anticipate that our devised strategy for data processing and prioritization will expedite the discovery of novel compound scaffolds and low-abundant metabolites from both unique and well-studied systems because it does not require prior knowledge of an organism's biosynthetic potential, it facilitates rapid lead prioritization largely independent of database screening, and it can be used to quickly compare well-characterized organisms to lesser studied species.

METHODS

Culturing and Extraction of Metabolites from *S. coelicolor* M145. *Streptomyces coelicolor* M145 (ATCC BAA-471) was incubated in a series of liquid ISP2 (4 g dextrose, 4 g yeast extract, 10 g malt extract, per liter at pH 7.2) cultures at 28 and 37 °C. An initial 5-mL culture was inoculated with 10 μ L of spore stock and incubated for 1 week, shaking at 105 rpm. Next, 2.5 mL was transferred into 100 mL of ISP2 media for 2 weeks, from which 50 mL was transferred into a 1-L ISP2 culture and grown for an additional 3 weeks, each shaken at 105 rpm. Cells were removed by centrifugation (1,693g; 2 \times 15 min), and the remaining supernatant was extracted for the initial comparative analysis studies. Metabolites from the spent media were extracted as described in the Supporting Information for subsequent MS analysis.

Mass Spectrometry Analysis of Secondary Metabolomes. Accurate mass data was acquired in triplicate using UPLC-ESI-TOF MS instrumentation (Agilent, 6220). Data was acquired in centroid mode with source/fragmentor voltage of 120 V, positive mode ion detection between 150–1700 m/z , gas temperature of 250 °C, and capillary voltage of 4,000 V. Comparative analysis of the acquired data was performed using the XCMS platform and is described in detail in the Supporting Information. Fragmentation studies of the saturated and monounsaturated compounds were completed with UPLC-ESI-CID-TOF-MS² (Waters, Acquity LC-Synapt G2-S). Ions isolated by the quadrupole were fragmented in the pre-IMS cell trap over a collision energy ramp of 30–55 V or 45–70 V in 5.3×10^{-3} mbar of N₂. Leucine enkephalin was used as a Lock Mass. Data was processed in Masslynx 4.2 and Origin 9.0. Fragmentation studies of the monohydroxylated compounds were completed by nanoLC-ESI-linear ion trap quadrupole-Orbitrap (Thermo Scientific, LTQ Orbitrap XL). Ions were fragmented with a normalized collision energy 35.0, FT resolution 7500, and 30.0 ms activation time. Accurate mass spectra on the parent ions were acquired by the Orbitrap, and nominal fragment spectra were acquired in the linear ion trap. Data acquired on a LC-ESI-linear ion trap quadrupole-Orbitrap from the *S. coelicolor* M145 37 °C extract was used to generate a molecular network utilizing a recently published procedure.^{19,25} For more details see Supporting Information.

ASSOCIATED CONTENT

Supporting Information

Additional methods descriptions, MS analysis information, MS and MS² data, and data tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: carlsone@indiana.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank L. Brown and J. Trinidad for helpful discussions. Access to the molecular networking algorithms was generously provided by P. Dorrestein. This work was supported by NIH R00GM82983, an NSF CAREER Award, a Pew Biomedical Scholar Award (E.E.C.) and an Indiana University Quantitative and Chemical Biology Training Fellowship (A.R.J.).

REFERENCES

- (1) Berdy, J. (2005) Bioactive Microbial Metabolites. *J. Antibiot.* 58, 1–26.
- (2) Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M.-A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147.
- (3) Udvary, D. W., Zeigler, L., Asolkar, R. N., Singan, V., Lapidus, A., Fenical, W., Jensen, P. R., and Moore, B. S. (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10376–10381.
- (4) Challis, G. L., and Hopwood, D. A. (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10362–10366.
- (5) Song, L., Barona-Gomez, F., Corre, C., Xiang, L., Udvary, D. W., Austin, M. B., Noel, J. P., Moore, B. S., and Challis, G. L. (2006) Type III polyketide synthase beta-ketoacyl-ACP starter unit and ethyl-malonyl-CoA extender unit selectivity discovered by *Streptomyces coelicolor* genome mining. *J. Am. Chem. Soc.* 128, 1475–14755.
- (6) Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M., and Takano, E. (2010) Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology* 156, 2343–2353.
- (7) Lin, X., Hopson, R., and Cane, D. E. (2006) Genome mining in *Streptomyces coelicolor*. Molecular cloning and characterization of a new sesquiterpene synthase. *J. Am. Chem. Soc.* 128, 6022–6023.
- (8) Magnolo, S. K., Leenutaphog, D. L., DeMonena, J. A., Curtis, J. E., Bailey, J. E., Galazzo, J. L., and Hughes, D. E. (1991) Actinorhodin production by *Streptomyces coelicolor* and growth of *Streptomyces lividans* are improved by the expression of a bacterial hemoglobin. *Nat. Biotechnol.* 9, 473–476.
- (9) Izumikawa, M., Shipley, P. R., Hopke, J. N., O'Hare, T., Xiang, L., Noel, J. P., and Moore, B. S. (2003) Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from *Streptomyces coelicolor* A3(2). *J. Ind. Microbiol. Biotechnol.* 30, 510–515.
- (10) Bode Bjorn, H., Bethe, B., Hofs, R., and Zeeck, A. (2002) Big effects from small changes: Possible ways to explore nature's chemical diversity. *ChemBioChem* 3, 619–627.
- (11) Tanaka, Y., Hosaka, T., and Ochi, K. (2010) Rare earth elements activate the secondary metabolite–biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J. Antibiot.* 63, 477–481.
- (12) Ito, Y., and Butler, A. (2005) Structure of synechobactins, new siderophores of the marine cyanobacterium *Synechococcus* sp. PCC 7002. *Limnol. Oceanogr.* 50, 1918–1923.
- (13) Wen, Y., Clark, P. M., Mason, D. E., Keenan, M. C., Hill, C., Goddard, W. A., III, Peters, E. C., Driggers, E. M., and Hsieh-Wilson, L. C. (2012) Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. *Science* 337, 975–980.

- (14) Bowen, B. P., and Northen, T. R. (2010) Dealing with the unknown: Metabolomics and metabolite atlases. *J. Am. Soc. Mass Spectrom.* 21, 1471–1476.
- (15) Lang, G., Mayhudin, N. A., Mitova, M. I., Sun, L., van der Sar, S., Blunt, J. W., Cole, A. L. J., Ellis, G., Laatsch, H., and Munro, M. H. G. (2008) Evolving trends in the dereplication of natural product extracts: New methodology for rapid, small-scale investigation of natural product extracts. *J. Nat. Prod.* 71, 1591–1599.
- (16) Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. *Anal. Chem.* 78, 779–787.
- (17) Patti, G. J., Tautenhahn, R., and Siuzdak, G. (2012) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. *Nat. Protoc.* 7, 508–516.
- (18) Patti, G. J., Yanes, O., and Siuzdak, G. (2012) Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell. Bio.* 13, 263–269.
- (19) Watrous, J., Roach, P., Alexandrov, T., Heath, B. S., Yang, J. Y., Kersten, R. D., van der Voort, M., Pogliano, K., Gross, H., Raaijmakers, J. M., Moore, B. S., Laskin, J., Bandeira, N., and Dorrestein, P. C. (2012) Mass spectral molecular networking of living microbial colonies. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10150–10151.
- (20) Neumann, S., and Bocker, S. (2010) Computational mass spectrometry for metabolomics: identification of metabolites and small molecules. *Anal. Bioanal. Chem.* 398, 2779–2788.
- (21) Wikoff, W. R., Gangotri, J. A., Barshop, B. A., and Siuzdak, G. (2007) Metabolomics identifies perturbations in human disorders of propionate metabolism. *Clin. Chem.* 53, 2169–2176.
- (22) Hindra, Pak, P., and Elliot, M. A. (2010) Regulation of a novel gene cluster involved in secondary metabolite production in *Streptomyces coelicolor*. *J. Bacteriol.* 192, 4973–4982.
- (23) Kersten, R. D., Yang, Y.-L., Xu, Y., Cimermancic, P., Nam, S.-J., Fenical, W., Fischbach, M. A., Moore, B. S., and Dorrestein, P. C. (2011) A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat. Chem. Biol.* 7, 794–802.
- (24) Kim, J., Choi, J. N., Kim, P., Sok, D. E., Nam, S. W., and Lee, C. H. (2009) LC-MS/MS profiling-based secondary metabolite screening of *Myxococcus xanthus*. *J. Microbiol. Biotechnol.* 19, 51–54.
- (25) Bandeira, N., Tsur, D., Frank, A., and Pevzner, P. A. (2007) Protein identification by spectral networks analysis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6140–6145.
- (26) Stein, S. E., and Scott, D. R. (1994) Optimization and testing of mass spectral library search algorithms for compound identification. *J. Am. Soc. Mass Spectrom.* 5, 859–866.
- (27) http://metlin.scripps.edu/metabo_search_alt2.php.
- (28) Barona-Gomez, F., Wong, U., Giannakopoulos, A. E., Derrick, P. J., and Challis, G. L. (2004) Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J. Am. Chem. Soc.* 126, 16282–16283.
- (29) Müller, G., and Raymond, K. N. (1984) Specificity and mechanism of ferrioxamine-mediated iron transport in *Streptomyces pilosus*. *J. Bacteriol.* 160, 304–312.
- (30) Gledhill, M. (2001) Electrospray ionisation-mass spectrometry of hydroxamate siderophores. *R. Soc. Chem.* 126, 1359–1362.
- (31) Mawji, E., Gledhill, M., Worsfold, P. J., and Achterberg, E. P. (2008) Collision-induced dissociation of three groups of hydroxamate siderophores: ferrioxamines, ferrichromes and coprogens/fusigens. *Rapid Commun. Mass Spectrom.* 22, 2195–2202.
- (32) Simionato, A. V. C., de Souza, G. D., Rodrigues-Filho, E., Glick, J., Vouros, P., and Carrilho, E. (2006) Tandem mass spectrometry of coprogen and deferoxamine hydroxamic siderophores. *Rapid Commun. Mass Spectrom.* 20, 193–199.
- (33) Yang, Y.-L., Xu, Y., Kersten, R. D., Liu, W.-T., Meehan, M. J., Moore, B. S., Bandeira, N., and Dorrestein, P. C. (2011) Connecting chemotypes and phenotypes of cultured marine microbial assemblages by imaging mass spectrometry. *Angew. Chem., Int. Ed.* 50, 5839–5842.
- (34) D'Onofrio, A., Crawford, J. M., Stewart, E. J., Witt, K., Gavrish, K., Epstein, S., Clardy, J., and Lewis, K. (2010) Siderophores from neighboring organisms promote the growth of uncultured bacteria. *Chem. Biol.* 17, 254–264.
- (35) Helman, R., and Lawrence, G. D. (1989) The increase in ferrioxamine B reduction potential with increasing acidity of the medium. *Bioelectrochem. Bioenerg.* 22, 187–196.
- (36) Schulze, C. J., Bray, W. M., Woerhmann, M. H., Stuart, J., Lokey, R. S., and Lington, R. G. (2013) "Function-first" lead discovery: Mode of action profiling of natural product libraries using image-based screening. *Chem. Biol.* 20, 285–295.
- (37) Lee, H. S., Shin, H. J., Jang, K. H., Kim, T. S., Oh, K. B., and Shin, J. (2005) Cyclic peptides of the nocardamine class from a marine-derived bacterium of the genus *Streptomyces*. *J. Nat. Prod.* 68, 623–625.
- (38) Yamanaka, K., Oikawa, H., Ogawa, H. O., Hosono, K., Shinmachi, F., Takano, H., Sakuda, S., Beppu, T., and Ueda, K. (2005) Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*. *Microbiology* 151, 2899–2905.
- (39) Bickel, H., Bosshardt, R., Gäumann, E., Reusser, P., Vischer, E., Voser, W., Wettstein, A., and Zähler, H. (1960) Metabolic products of Actinomycetaceae. XXVI. Isolation and properties of ferrioxamines A to F, representing new sideramine compounds. *Helv. Chim. Acta* 43, 2118–2128.
- (40) Imbert, M., Béchet, M., and Blondeau. (1995) Comparison of the main siderophores produced by some species of *Streptomyces*. *Curr. Microbiol.* 31, 129–133.
- (41) Sandy, M., and Butler, A. (2009) Microbial iron acquisition: Marine and terrestrial siderophores. *Chem. Rev.* 109, 4580–4595.
- (42) Traxler, M. F., Seyedsayamdost, M. R., Clardy, J., and Kolter, R. (2012) Interspecies modulation of bacterial development through iron competition and siderophore piracy. *Mol. Microbiol.* 86, 628–644.
- (43) Butler, A., and Theisen, R. M. (2010) Iron(III)-siderophore coordination chemistry: Reactivity of marine siderophores. *Coord. Chem. Rev.* 254, 288–296.
- (44) Martinez, J. S., Carter-Franklin, J. N., Mann, E. L., Martin, J. D., Haygood, M. G., and Butler, A. (2003) Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3754–3759.
- (45) Martinez, J. S., and Butler, A. (2007) Marine amphiphilic siderophores: marinobactin structure, uptake, and microbial partitioning. *J. Inorg. Biochem.* 101, 1692–1698.
- (46) Murakami, T., Burian, J., Yanai, K., Bibb, M. J., and Thompson, C. J. (2011) A system for the targeted amplification of bacterial gene clusters multiplies antibiotic yield in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16020–16025.
- (47) Craney, A., Ozimok, C., Pimentel-Elardo, S. M., Capretta, A., and Nodwell, J. R. (2012) Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chem. Biol.* 19, 1020–1027.
- (48) Ahmed, S., Craney, A., Pimentel-Elardo, S. M., and Nodwell, J. R. (2013) A synthetic, species-specific activator of secondary metabolism and sporulation in *Streptomyces coelicolor*. *ChemBioChem* 14, 83–91.
- (49) Kaneda, T. (1991) Iso- and anteiso-fatty acids in bacteria: Biosynthesis, function and taxonomic significance. *Microbiol. Rev.* 55, 288–302.
- (50) Wallace, K. K., Zhao, B., McArthur, H. A. I., and Reynolds, K. A. (1995) In vivo analysis of strain-chain and branched-chain fatty acid biosynthesis in three actinomycetes. *FEMS Microbiol. Lett.* 131, 227–234.
- (51) Moore, B. S., and Hertweck, C. (2002) Biosynthesis and attachment of novel bacterial polyketide synthase start units. *Nat. Prod. Rep.* 19, 70–99.
- (52) Luo, M., Fadeev, E. A., and Groves, J. T. (2005) Membrane dynamics of the amphiphilic siderophore, acinetoferrin. *J. Am. Chem. Soc.* 127, 1726–1736.
- (53) Martin, J. D., Ito, Y., Homann, V. V., Haygood, M. G., and Butler, A. (2006) Structure and membrane affinity of new amphiphilic

siderophores produced by *Ochrobactrum* sp. SP18. *J. Biol. Inorg. Chem.* 11, 633–641.