Expanding the scope of detectable microbial natural products by complementary analytical methods and cultivation systems

Chantal D. Bader[1†], Patrick A. Haack[1†], Fabian Panter[1], Daniel Krug[1] and Rolf Müller[1*]

[1] Department Microbial Natural Products, Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS)

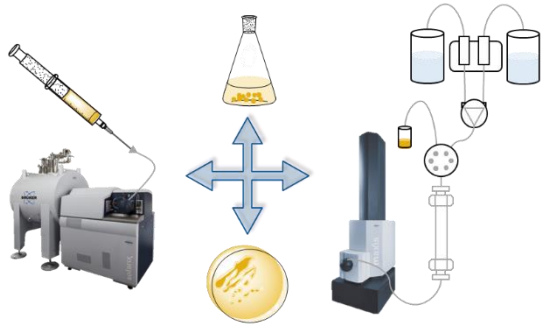
Helmholtz Centre for Infection Research (HZI), German Center for Infection Research (DZIF, Partnersite Hannover-Braunschweig) and Department of Pharmacy, Saarland University Campus E8.1 , 66123 Saarbrücken (Germany)

[†]These authors contributed equally to the manuscript

[*]Correspondence: rolf.mueller@helmholtz-hips.de

KEYWORDS: natural products, myxobacteria, DI-FTICR, UHPLC-Q-TOF, secondary metabolomics

ABSTRACT: Recent advances in genome sequencing unveiled a large discrepancy between the genome-encoded capacity of microorganisms to produce secondary metabolites and the number of natural products (NP) actually detected in their extracts. In order connect biosynthetic gene cluster diversity to NPs, common approaches include genetic manipulation of producers and diversifying the cultivation conditions. However, the range of detectable metabolites is fundamentally limited by the applied analytical method. In this work, we present a two-platform mass spectrometry analysis for the comprehensive secondary metabolomics characterization of nine myxobacterial strains. We evaluated direct infusion (DI) measurements of crude extracts on a Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometer and compared them to measurements conducted on a Time-Of-Flight (TOF) device coupled to liquid chromatography (LC), a setup widely used in NP laboratories. We demonstrate that both methods are likewise successful to detect known metabolites, whereas statistical analysis of unknowns highlights their complementarity: strikingly, 82-99% molecular features were only found with one of the analytical setups. In addition, we evaluated NP profile differences seen from our set of strains grown in liquid culture versus their swarming colonies on agar plates. The detection of 21-96% more molecular features when both liquid and plate cultures were analyzed (compared to only one of them) translates into increased chances to identify new NPs. Determination of strain-specific compounds in combination with GNPS molecular networking revealed strain Mx3 as particularly promising in terms of isolation and structure elucidation of novel secondary metabolites.



INTRODUCTION

There is an imminent threat emerging from antimicrobial resistance development in microbial pathogens.[1] As the need for novel antimicrobial lead structures is becoming more and more urgent, natural product research is experiencing a resurgence.[2,3] Complementary to well-studied microorganisms like streptomycetes and bacilli, myxobacteria represent an ubiquitous but yet underexplored source for natural products.[4,5] Previously identified and isolated myxobacterial secondary metabolites exhibit a wide range of biological activities as well as intriguing structural diversity.[6–13] Many of the compounds isolated early on were identified by activity guided isolation, while genomics and metabolomics driven approaches have recently come into focus.[14–16] This is mainly caused by advances in sequencing technology, as well as *in silico* annotation and predictions tools, that supplied natural product laboratories with an unprecedented amount of ready to use genomic data[17–19] These data revealed that the number of biosynthetic gene clusters (BGCs) encoding secondary metabolites by far exceeds the amount of compounds characterized to date.[20,21] Common approaches to access these "hidden" secondary metabolites include the activation or inactivation of underlying biosynthetic pathways. This is attempted by changing the cultivation conditions, introducing environmental challenges or by genetically manipulating the strains.[22–26] While this relatively slow methodology has proven to be successful for compound discovery, one often overlooked limiting factor for compound detection and identification is the analytical setup applied during microbial crude extract analysis. Common analytical setups include liquid chromatography - mass spectrometry (LC-MS), gas chromatography - MS (GC-MS) and matrix assisted laser desorption ionization - MS (MALDI-MS) based systems. LC-coupled time of flight (TOF)-MS is among the most frequently used high resolution analytical setups in bacterial secondary metabolomics. The main advantages of TOF-MS are its high dynamic range, high sensitivity, high mass accuracy and easily established automation.[27–29] TOF mass spectrometers are also well suited to be coupled to liquid chromatography due to their fast scan rates. This offers

the advantage to link mass signals with LC retention times that serve as an orthogonal identifier, which increases the confidence of annotations and allows association of adducts based on peak congruence.[30–32] Although this setup is capable of detecting a wide range of microbial secondary metabolites, it has some inherent limitations that need to be considered. The choice of eluents as well as columns adds a discriminatory effect that limits the detectable chemical space. An inadequate choice of the chromatographic conditions can lead to the dilution of compounds in the LC gradient, rather than the desired concentration due to the chromatographic retention. To bypass these factors mass spectra can also be generated using direct infusion into the mass spectrometer (DI). In order to use DI-MS as a profiling tool, mass spectrometers need ultrahigh resolution to differentiate between near-isobaric substances.[33] A Fourier transform ion cyclotron (FTICR)-MS offers ultrahigh resolution as well as a low limit of detection and high precision.[34] Nevertheless, the ultrahigh mass accuracy of the system requires a slower scan rate than TOF instruments which makes them less suitable for LC-MS hyphenation. Therefore, direct infusion setups are a plausible choice for FTICR instruments. While DI-FTICR applications remain underrepresented in bacterial secondary metabolomics, there are several applications in fields ranging from biomarker identification to petroleomics already utilizing the strengths of DI-FTICR to rapidly and accurately analyze complex mixtures.[35–39] However, a drawback of DI-MS applications is the higher prevalence of ion suppression effects.[40] Moreover, the injection of complex mixtures like bacterial crude extracts generates a dense ion cloud, which may cause reduction in data quality due to so called space-charge-effects. To minimize these effects in the FTICR cell, a spectral stitching method as developed by Southam *et al.* can be applied. Herein, the quadrupole incorporated in the device is used as an ion filter for a specific m/z value range. The whole m/z range can thus be divided into smaller segments, entering the FT-ICR cell one after another.[41]

Myxobacteria require complex cultivation media to grow, which contain well ionizing molecules including a variety of polar compounds and lipids. This may be one of the reasons why LC-TOF is commonly employed as the analytical method of choice. The LC method separates all components across a polarity gradient, resulting in reduced ion suppression effects.[42] But when cultivated in rich and complex liquid media, the cultivation conditions greatly differ from the ecological niche of myxobacteria which are soil living organisms. A practical approach to cultivate myxobacteria under conditions that resemble their natural habitat is to grow myxobacteria on agar plates. This has a big influence on their unique morphology as they are only able to form their characteristic fruiting bodies or show their characteristic swarming behavior when grown on solid medium.[43] Those changes in phenotype are accompanied by immense changes in the secondary metabolite profiles.[44]

This work aims at extending the detectable space of microbial secondary metabolites, using the characterization of the metabolomes of nine different myxobacterial strains as an example. Furthermore, we study the influence of different cultivation conditions on the metabolites produced and herein evaluate the possible impact of DI-FTICR as a hitherto non-standard analytical setup in bacterial secondary metabolomics research. In a first step, we investigate the production of known myxobacterial metabolites in the two different cultivation systems as well as their detection using DI-FTICR and LC-TOF platforms. Secondly, all metabolites detected in our analyses are analyzed into a non-targeted metabolomics workflow. MS/MS spectra in combination with molecular networking are used to further investigate metabolites only detectable with one of the analytical setups. The generated data is ultimately used to distinguish between primary and secondary metabolome.

MATERIALS AND METHODS

All materials, including the bacterial strains used for this analysis (as diverse set of representatives of different myxobacterical families), bacterial cultivation media, fermentation protocols, related sample preparation and experimental procedures are described in detail in the supporting information.

Cultivation and Sample Preparation. Myxobacterial strains were cultivated in liquid culture as well as on agar plates using CYH medium. All cultivations were done in biological triplicates. Cells from the agar plate cultures were scraped off and liquid cultures were centrifuged to separate cells and supernatant, respectively. Prior to extraction with methanol, the cell pellets were lyophilized. Blank samples were generated by lyophilizing and extracting the cultivation medium treated in the same way as the bacterial fermentation cultures.

DI-FTICR and LC-TOF Analysis. All MS measurements were performed in duplicates. DI-FTICR measurements were performed using positive as well as negative ionization on the Bruker SolariX XR 7T. Samples were diluted 1:200 with meth-

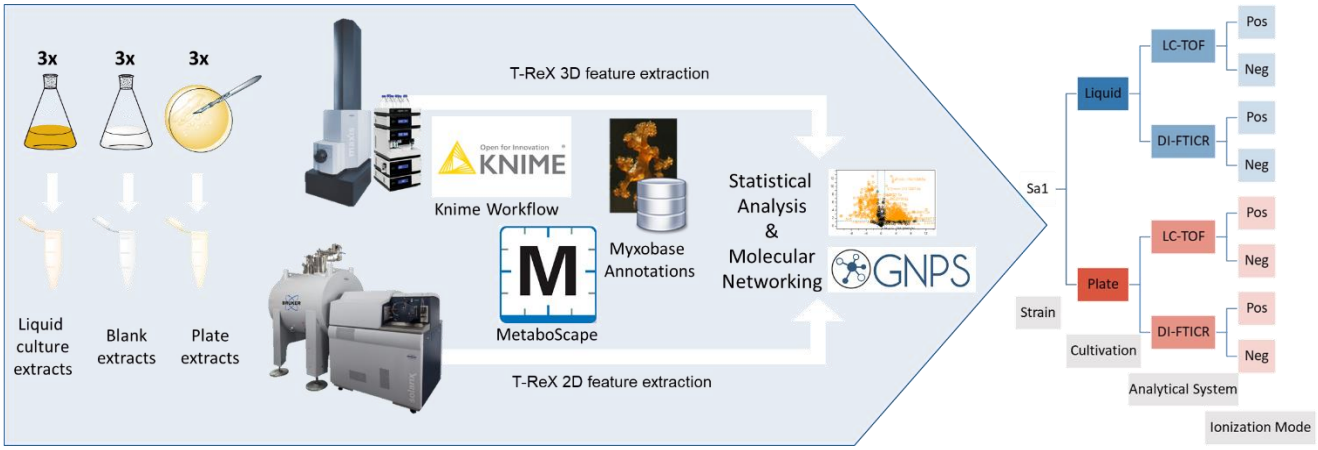


Figure 1. General workflow (left side) and measurements conducted per strain (right side).

anol prior to measurements and the mass spectrometer was externally calibrated to a mass accuracy below 1 ppm. The samples were injected with a preinstalled syringe pump at 2.0 µl/min flowrate. 110 scans were performed within 4 minutes accumulating for 150 ms. The total m/z range from 100-1500 was divided in eight segments that were later combined to a full spectrum. Collision RF Amplitude was optimized for each segment. The data size was set to 4 M with a 2 s transient. LC-TOF measurements were performed on a Dionex Ultimate 3000 SL system coupled to a Bruker maXis 4G UHRTOF. The mobile phase consisted of (A) ddH_2O with 0.1% formic acid and (B) acetonitrile with 0.1% formic acid. For separation a linear gradient from 5-95% B in A on a Waters Acquity BEH C18 column (100 x 2.1 mm, 1.7 µm d_p) was used. The flow rate was set to 0.6 mL/min and the column thermostated at 45 °C. Extracts were diluted 1:20 prior to measurements. The LC flow was split to 75 µL/min before entering the mass spectrometer, which was externally calibrated to a mass accuracy below 1 ppm. Mass spectra were acquired in centroid mode ranging from 150-2500 m/z at a 2 Hz scan rate. Further details of the analytical setups are described in the SI.

Statistical Analysis and Annotations. Statistical interpretation for targeted and non-targeted metabolomics analysis was carried out with MetaboScape 4.0 (Bruker). The minimal intensity threshold for feature detection was set to 1.5 x 10^4 for DI-LC-TOF data and 6 x 10^6 for DI-FTICR data. The maximum charge was set to three (positive and negative) and the minimal group size for creating batch features to five. Known metabolites were annotated using an in-house database containing myxobacterial secondary metabolites (Myxobase).

MS^2 Analysis and Molecular Networking. SPL-guided MS^2 spectra were generated on the LC-TOF system using the same chromatographic conditions as for MS measurements. For enhanced spectra coverage, a specified precursor list was used, containing only features originating from the bacteria.[45] MS^2 spectra generated on the FT-ICR system were recorded manually and parameters were optimized for each precursor ion. A complete list of CID energies can be found in the SI. MS^2 data was uploaded to the GNPS server at the University of California, San Diego.[46] The clustered dataset was visualized using Cytoscape 3.7.2.

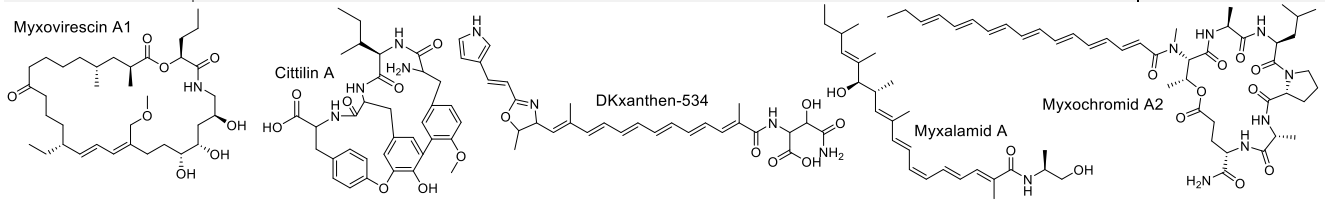
RESULTS AND DISCUSSION

In order to extend the scope of detection for myxobacterial secondary metabolites, we generated a total number of eight different analyses per strain (Figure 1). Each strain was cultivated in liquid cultures and on agar plates. Subsequently, the two extracts per strain were measured on the two analytical systems (LC-TOF and DI-FTICR) in positive and negative ionization mode. This cascade therefore allows metabolomics-based comparison under four different aspects: the ionization polarity, the analytical systems operated in the same ionization mode, the two cultivation systems and the metabolome of the different strains used for our analysis.

Targeted metabolomics comparison of the two analytical setups and cultivation systems. To assess the production of known compounds under the two cultivation conditions as well as their detectability with the two analytical setups, the recorded data were annotated using analytical details from our in-house database Myxobase. As most of these metabolites were formerly detected in positive ionization mode, negative ionization data was excluded here. For annotations to be accepted as valid, we used a mass accuracy window of 5 ppm and allowed retention time deviations up to 0.2 min. As there is no second dimension such as retention time for the DI-FTICR measurements, we

Table 1. Known myxobacterial secondary metabolites detectable in Mx1 extracts by analysis with LC-TOF and DI-FTICR. The ion types assigned by automated annotations were manually confirmed. Only hits with mass deviation below 5 ppm and retention time deviation (if given) below 0.2 min were considered as valid. Additional evidence for correct annotations was gained by antiSMASH annotations of the corresponding known biosynthetic gene clusters (BGCs) to confirm the strains ability to produce the detected secondary metabolites.

	LC-TOF+Liquid		*LC-TOF+Plate*		*DI-FTICR+Liquid*		*DI-FTICR+Plate*			
	Δppm	**Ion type**	**Δppm**	**Ion type**	**Δppm**	**Ion type**	**Δppm**	**Ion type**	**ΔRT[min]**	**BGC**
Myxovirescin A	3	$[M+H]^+$	3	$[M+H]^+$	0.7	$[M+K]^+$			0.08	✓
Myxovirescin B	2.7	$[M+H]^+$							0.1	✓
Myxovirescin C	2.6	$[M\text{-}MeOH+H]^+$	0.6	$[M\text{-}MeOH+H]^+$	2.3	$[M+Na]^+$	3.3	$[M+K]^+$	0.11	✓
Myxovirescin G	2.3	$[M+H]^+$			2.6	$[M+K]^+$	3.1	$[M+Na]^+$	0.04	✓
Cittilin A	2.7	$[M+H]^+$							0.02	✓
DKxanthen-534	4	$[M+H]^+$	0.7	$[M+H]^+$			0	$[M+Na]^+$	0.04	✓
Dkxanthen-544	3.5	$[M+H]^+$							0	✓
Dkxanthen-548	3.8	$[M+H]^+$	0.4	$[M+H]^+$	3.1	$[M+K]^+$	0.5	$[M+H]^+$	0.16	✓
Dkxanthen-574	3.3	$[M+H]^+$	0.1	$[M+H]^+$	4.5	$[M+H]^+$	4.8	$[M+H]^+$	0.18	✓
Dkxanthen-560	2.5	$[M+2H]^{2+}$	1.5	$[M+2H]^{2+}$					0.03	✓
Myxalamid A	3.4	$[M\text{-}H_2O+H]^+$	0.6	$[M\text{-}H_2O+H]^+$	0.6	[M+Na]+	0.7	[M+Na]+	0.06	✓
Myxalamid B	3.4	$[M\text{-}H_2O+H]^+$	0.6	$[M+H]^+$	2.9	$[M+K]^+$	2	[M+Na]+	0.08	✓
Myxalamid C	3.3	$[M\text{-}H_2O+H]^+$	0.5	$[M+H]^+$					0.08	✓
Myxochromid A3	1.4	$[M+H]^+$	0.6	$[M+H]^+$	0	[M+Na]+	0.4	[M+Na]+	0.18	✓



only considered hits where more than one member of a secondary metabolite family was detected, or those that were also detectable with the LC-TOF setup, as valid annotations. Most of the strains studied here are new isolates and therefore barely investigated in terms of secondary metabolite richness. This explains why there are only few annotations for most strains, making them on the other hand interesting subjects for screening approaches aimed at novel natural product discovery. Table 1 exemplarily shows secondary metabolites detectable in Mx1 extracts.[47–53] This strain is one of the best characterized myxobacterial strains and produces several natural product families comprising a high number of family members, demonstrated by the high number of annotated features.[53] Detailed annotation results for the targeted investigation for all other strains can be found in the SI.

When comparing the metabolic effects of the cultivation conditions on agar plate and in liquid medium within the LC-TOF datasets, more than 75% of the known compounds produced with liquid cultivation were also found in the extracts from plate cultures. Analyzing DI-FTICR data for differences between the two cultivation systems also shows that 8 out of 9 metabolites can be detected in plate extracts as well as in liquid extracts. Myxovirescin A was only found in the liquid extract and DKxanthene-534 only in the plate extract. Otherwise, all annotated compounds were found in both extracts. Therefore, the general picture emerging from comparison of the two analytical setups is that they can be regarded as similarly capable to detect the known myxobacterial compound classes. When analyzing the individual members of the secondary metabolite families, however, not all derivatives found with the LC-TOF setting were also detectable with DI-FTICR. We furthermore observe major differences in the ion types that are detected. In the LC-TOF analyses $[M+H]^+$ represents the most abundant ion type. $[M+Na]^+$ adducts are often also detected and can be assigned to the other adducts by the chromatographic peak profile. In the DI-FTICR measurements $[M+Na]^+$ and $[M+K]^+$ ions are the most abundant ions, likely because the salts are not separated from the metabolites as it would happen during an LC run. Frequently, $[M+H]^+$ cannot be found in the DI-FTICR analyses at all. Although this does not affect conclusions from the measurements in principle, operators need to be aware of this in order to avoid false annotations.

Generally, most annotations can be found for the combination of liquid cultivation and LC-TOF analysis, showing that the standard cultivation and analytical setup is suitable for a broad range of metabolites. As our in-house database strongly relies on input of natural products isolated after cultivation in liquid medium and detected by LC-TOF, our analyte data collection is biased towards compounds detectable with exactly this setup. In order to make a less biased statement on which setup leads to detection of the largest share of secondary metabolites, including unknowns, we conducted non-targeted metabolomics analysis next.

Non-targeted comparison of the two cultivation and analytical systems. For non-targeted comparison of our measurements, the detected signals were grouped into molecular features to collect different ion types belonging to one metabolite as a first data reduction step. Subsequently, all molecular features from one analysis were compared to the ones measured in the same ionization mode, either to compare the two cultivations or the two analytical setups. In that way, the percentage of

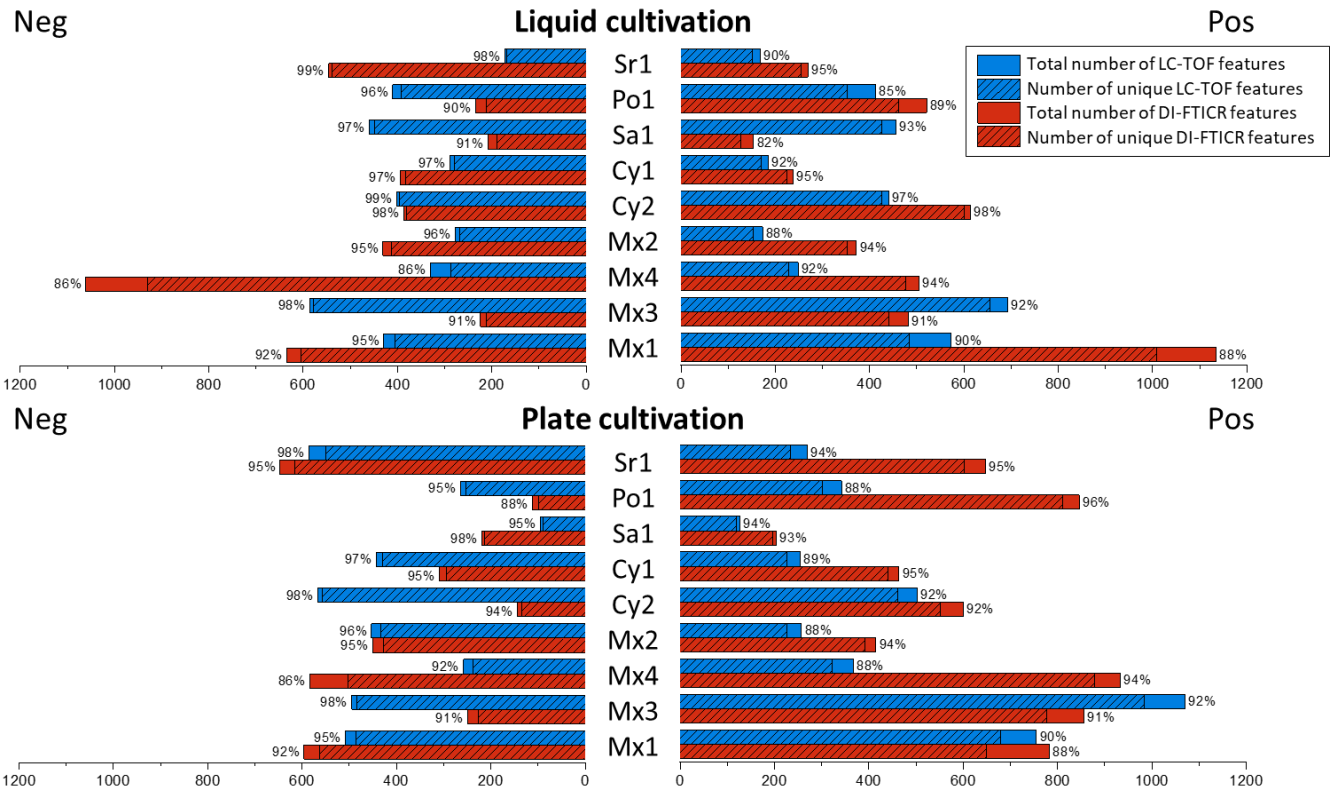


Figure 2. Features detected for the nine different myxobacterial strains comparing the two analytical systems. Blue: DI-FTICR Features, Red: LC-TOF Features. Percentage of unique features to one of the analytical systems shown at the respective end of a bar and marked as dashes. Bars to the left represent molecular features detected in negative mode and bars to the right features in positive mode.

features only present in one of the analyses (later on referred to as unique features) was calculated.

In positive ionization mode we can find between 53 and 93% plate unique molecular features (DI-FTICR). In our liquid cultivation experiment the percentage of unique molecular features is 42-87%. Cultivation on plate therefore delivers more unique features that we can detect in positive mode than cultivation in liquid (see SI). The absolute amount of molecular features detectable in negative mode is generally lower than in positive mode (Figure 2 and SI). In negative ionization mode we find 22-62% plate unique and 21-72% liquid unique molecular features (DI-FTICR). Hence, in negative mode we observe the percentage of liquid unique molecular features a bit higher than the percentage of plate unique molecular features. Comparing our results of the DI-FTICR measurements to the results of the LC-TOF measurements, the percentage of unique molecular features is even higher in the LC-TOF measurements. In positive ionization mode we observe 70-93% plate unique molecular features and 58-92% liquid unique molecular features (LC-TOF). For measurements conducted in negative mode the percentage of unique features is in a similar range, giving 56-96% plate unique molecular features and 48-96% liquid unique molecular features (LC-TOF).

The absolute number of features detectable is highly strain dependent, also accounting for the percentage of features uniquely found in one cultivation system (Figure 2). However, we observe a trend that more positive than negative ionizing substances are detectable in the plate extracts (6 out of 9 strains with DI-FTICR and 7 out of 9 with LC-TOF). Most importantly, our study shows that we lose minimum 21% and maximum 93% of information when just using one cultivation system, as those features are uniquely produced under one cultivation condition. Our results confirm that cultivation on plate significantly changes the metabolite profile and should be seen as important complementary method rather than an occasional amendment to bacterial cultivation in liquid medium.

When now comparing the two analytical setups to each other within the context of one cultivation system, in positive mode we find an average of 88-94% LC-TOF unique molecular features in the plate extracts. With DI-FTICR we find 88-96% features that we cannot detect with the other method. Those results highlight the comparably small overlap of the two systems, and at the same time reveals the huge chemical space not captured, when solely using one of the analytical methods. Analysis of the liquid extracts shows consistent results compared to the plate extracts. Here we detect 85-97% unique LC-TOF molecular features and 82-98% unique DI-FTICR molecular features in positive mode. In negative mode we find 86-99% unique LC-TOF molecular features and 86-99% unique DI-FTICR molecular features for liquid cultivation. Plate cultivation shows similar results, giving 92-98% unique LC-TOF molecular features and 86-98% unique DI-FTICR molecular features. Our analysis therefore suggests, that the analytical system chosen for the detection of metabolites has a considerably bigger influence on the observable metabolome than the choice of cultivation system.

Nevertheless, as already shown for the targeted metabolomics analysis, the ion type detectable with DI-FTICR may differ from the ion type detectable with LC-TOF. This also has a large influence on the output of a non-targeted analysis, as the detection of several ion types is crucial for molecular feature generation. When only one ion-type is detected during the measurement, the feature extracting algorithm cannot distinguish between $[M+H]^+$, $[M+Na]^+$ or $[M+K]^+$. This would result in different calculated molecular masses, leading to mismatching in our analysis. In order to estimate the degree to which this phenomenon may have influenced our non-targeted analysis, we calculated the minimum number of specific features. We repeated the non-targeted comparison in positive ionization mode between LC-TOF and DI-FTICR for the strain Mx1. As $[M+Na]^+$ and $[M+K]^+$ were the most common ion types besides $[M+H]^+$ in our targeted analysis, we calculated the theoretical molecular masses corresponding to the three ion types. For every single feature in our analysis, we therefore obtained three different molecular masses. Subsequently, we calculated the amount of common features under the assumption, that only $[M+Na]^+$ (respectively $[M+K]^+$ or $[M+H]^+$) is detectable. The common features of all three calculations were added up and subtracted from the total number of features, giving the minimum amount of features unique for one of the analytical setups. In this scenario we still found 53-74 % unique features for each platform when comparing LC-TOF and DI-FTICR in positive mode. However, this calculation likely underestimates the amount of unique features, as the amount of features is artificially inflated and the same feature could be matched more than one time. We therefore consider our initial calculations as more suitable for a practical assessment of analytical system complementarity. Moving beyondthis minimum number of unique features calculation, we performed MS/MS fragmentation and subsequent GNPS-clustering, which provides extendes molecular feature comparison.

MS/MS clustering of DI-FTICR unique features. To determine whether molecular features produced by the strain Mx1 uniquely found with DI-FTICR really constitute an extension of the detectable chemical space, 18 DI-FTICR unique molecular features found in our prior analysis were fragmented. Their fragmentation pattern was compared to the fragmentation pattern of all molecular features found for Mx1 in the LC-TOF

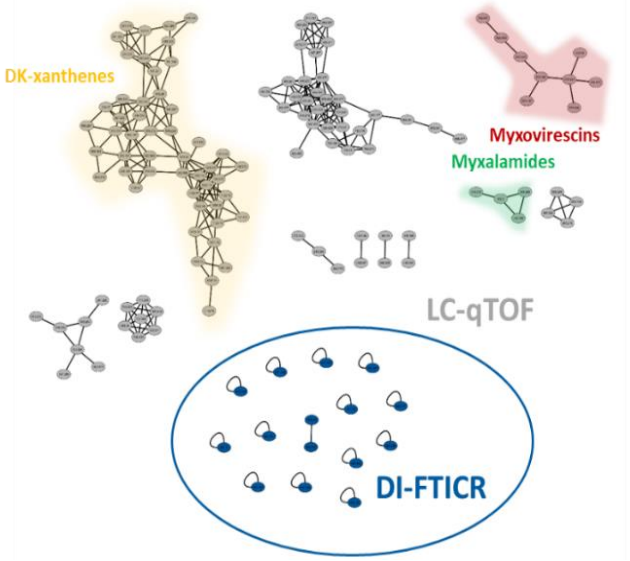


Figure 3. Molecular network of Mx1 features detected with LC-TOF (grey nodes) and eight of the most intense features uniquely found with DI-FTICR (blue nodes). Known compound families produced by this strain are marked in red (Myxovirescins), yellow (DKxanthenes) and green (Myxalamides).

analysis. None of these 18 molecular features found uniquely with DI-FTICR clusters to molecular features found with LC-TOF (see Figure 3). This result suggests that substances we can detect exclusively with DI-FTICR are not simply derivatives of compound families that can also be observed with LC-TOF. Therefore, DI-FTICR analysis of bacterial extracts unlocks previously undetectable metabolites. Final characterization of the chemical structures of all molecular features found with DI-FTICR nevertheless could not be performed with our current setup, as fragmentation here is non-automated and therefore only possible for a limited number of molecular features. Although only a limited set of molecular features uniquely detectable by DI-FTICR were fragmented, we already observe MS/MS based fragmentation pattern clustering of two of the DI-FTICR unique molecular features. Such substances are of special interest, since secondary metabolite families usually have several family members.

Comparison of the influence of the ionization mode on the analysis. In previous studies, Nordström *et al*. have reported that they achieved a 90% increase in detected ion species when performing ESI-MS analysis in both ionization modes instead of just using one mode in a human serum extract analysis.[54] We were therefore interested to assess, whether the influence of the ionization mode is comparably high for myxobacterial metabolite detection. Across all nine strains and after subtraction of blank features, 17,295 molecular features were detected in positive and 13,359 molecular features in negative ionization mode. In positive ionization mode, we are consistently able to detect more molecular features (Figure 4). Only for the combination of LC-TOF and liquid cultivation, the number of detected molecular features are in the same range. This discrepancy is likely due to the variance in ionization efficiencies between positive and negative ionization for different compounds.[55,56] It seems likely that due to their chemical structure, featuring many hydroxyl, carbonyl and amine groups a majority of myxobacterial metabolites has a higher ionization efficiency in positive ionization mode. When comparing positive and negative ionization between the two analytical setups, we observed a difference in the mean m/z values of the measured molecular features. In the DI-FTICR experiments, the mean m/z of all molecular features detected with positive ionization is about 20-30 % higher than the mean m/z of all molecular features detected with negative ionization. For LC-TOF data, we observed the opposite effect, with the mean m/z of all features measured in negative mode being about 10 % higher.

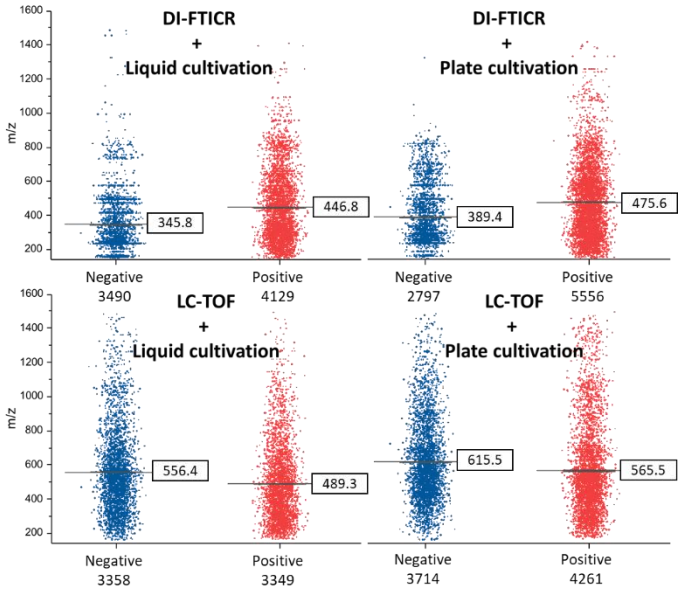


Figure 4. m/z distribution for negative (blue) and positive (red) ionization mode. Molecular features from all strains were grouped according to the analytical setup and cultivation condition. The mean m/z value +/- standard deviation are indicated by the horizontal bars. The total number of features detected for each cultivation condition and analytical setup combination, across all strains for positive and negative ionization measurements is indicated below the scatter plot.

Characterization of the unique secondary metabolome for each myxobacterial strain. In order to cope with the large number of features in our analyses and to prioritize features for isolation and structure elucidation, we devised a workflow to estimate which molecules stem from primary and which from secondary metabolism. This distinction is of great importance when searching for antimicrobial natural products, as secondary metabolites are often produced as defending mechanism for the producing organism against competitors. A large number of metabolites detectable in a metabolomics investigation, however stem from primary metabolism. The presented workflow enables us to remove all features from the primary metabolism and focus on the unique metabolome per strain which makes the discovery of novel secondary metabolites much easier. As fully automated MS/MS fragmentation is only possible on the LC-TOF system and our previous analysis showed that more myxobacterial metabolites ionize in positive mode, we only performed this analysis for the data generated with positive LC-TOF data. All molecular features detected in one strains' analysis were compared to all the features of the other strains. Each features found in at least one other strain is removed and the remaining features are submitted to MS/MS analysis and GNPS based molecular networking (Figure 5).[46] By subtracting all molecular features that appear in several analyses, we believe to remove most of the compounds belonging to the primary metabolism, as these would appear in most of the analyzed strains. We likely also remove frequently occurring secondary metabolites from the analysis. Secondary metabolites appearing in several strains are, in principle, of less interest to us than a unique molecular family, as the likelihood of discovering novel chemistry is increased in these cases and rediscovery of already described secondary metabolites additionally is more frequent in those cases. Furthermore, if the unique metabolome of one strain appears to be comparably small, our workflow also allows for the calculation of a shared metabolome between two or more strains to cope with this aspect.

After removal of common molecular features from each strains analysis the average of 6662 features that we detect per strain (combined plate and liquid) could be reduced by 94% to an average of 409 features per strain. Remarkably, the amount of unique features per strain is very diverse. The strain with the lowest amount of unique features is Mx2 only showing 16 unique molecular features, whereas Mx3 also belonging to the family of *Myxococcaceae* shows 1242 unique molecular features. After selectively targeting only these molecular features for an MS/MS experiment and GNPS clustering, we obtain a molecular network of solely unique molecular features across all strains. When now analyzing the molecular network of those features, 2/3 of the MS-clusters comprise features solely produced by one strain. 1/3 of the MS-clusters contain mixed features from different strains, meaning that the different strains produce different variants of related natural product families.

This finding is also reflected on the genetic level, as many strains contain BGCs of the same BGC family with less than 100% similarity.[12] In total, we find 81 MS-clusters where all features are just produced by one single strain. Interestingly, 35 MS-clusters thereof solely contain molecular features that are plate-unique molecular features. Thirteen MS-clusters only contain molecular features that are just produced in liquid culture. For more than half of the secondary metabolite families in our analysis the change of the cultivation conditions therefore is an all-or-nothing criterion. Under one condition a set of derivatives is produced, whereas under the other condition the BGC remains inactive.

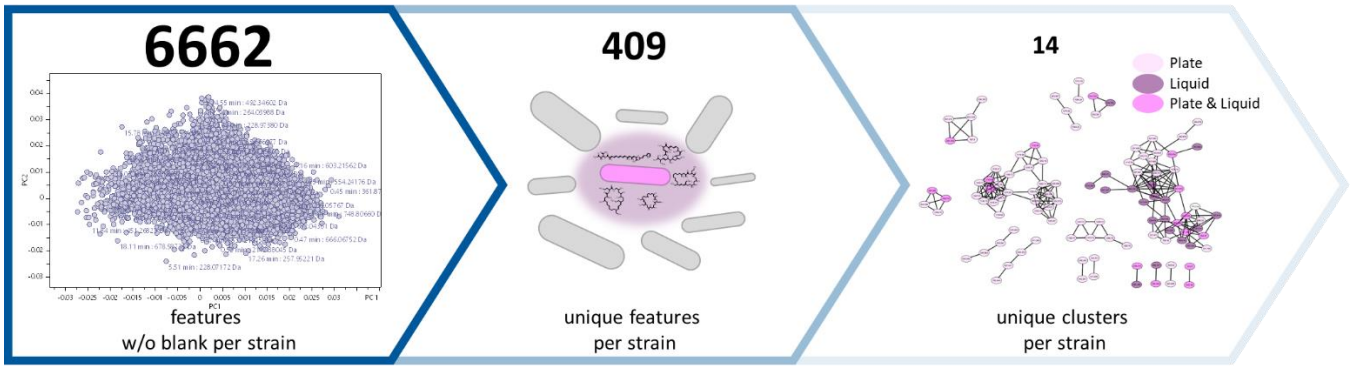


Figure 5. Prioritization workflow to cope with the large number of molecules deriving from primary metabolism. Molecular features depicted are the average number of molecular features detected per strain. Molecular network shown on right side derives from strain Sr1. Nodes depicted in light purple derive from plate cultivation, dark purple nodes from liquid cultivation and pink nodes can be found in both analyses.

CONCLUSION

In this work, we showed that both LC-TOF as well as DI-FTICR can be used to detect myxobacterial secondary metabolites in a targeted metabolomics workflow. Certainderivatives of known metabolites were only found when cultivating in liquid medium and subsequently analyzing the corresponding extracts with LC-TOF. Nevertheless, this analysis is severely biased, as secondary metabolites known so far and used for the annotation are mainly found in LC-MS experiments. In the DI-FTICR setup $[M+Na]^+$ and $[M+K]^+$ adducts are often the only observed ion type, which should be kept in mind for further analysis and when searching for and annotating specific metabolites. The non-targeted analysis of our data showed, that the two myxobacterial cultivation methods (agar plates and liquid culture) are complementary to each other and 21-93% of the myxobacterial metabolome cannot be detected when just applying one cultivation method. The amount of information lost strongly depends on the investigated strain as well as on the employed analytical setup. The choice of the analytical system used for myxobacterial metabolomics has an even bigger impact on the output of the study than the cultivation system. 82-99% of molecular features are unique for one of the two analytical systems that were compared in this study. MS/MS spectra of a subset of unique DI-FTICR features furthermore confirmed that many features uniquely detectable in DI-FTICR measurements belong to a different chemical space than the features detected in LC-TOF analyses. In general, positive ionization mode allows the detection of more myxobacterial molecular features than negative ionization mode. We could also observe a difference in the mean m/z value detectable depending on the used analytical setup, giving a lower mean m/z value detectable with DI-FTICR than with LC-TOF. The immense amount of data created from this metabolomics analysis requires further prioritization strategies to optimize secondary metabolite identification. One suitable approach presented in this work is to exclude molecular features found in more than one analysis. This results in over 90% data reduction. Remaining molecular features can be further grouped with GNPS molecular networking. This whole workflow results – for the set of example strains presented here - in a manageable amount of 14 unique MS-cluster families per strain that can be readily followed up. Comparing the nine myxobacterial strains investigated, Mx3 showed the highest potential for metabolomics-guided isolation and structure elucidation because of its high amount of unique MS-clusters besides the few known natural products it produces.

In conclusion, this study highlights that the choice of cultivation and analytical system massively influences the results of any microbial secondary metabolome analysis. In targeted workflows one well-chosen setup may be considered sufficient to detect all metabolites of interest. In a non-targeted workflow however, we were able to show using at least two analytical setups as well as different cultivation conditions leads to a drastic increase of the metabolome coverage. The additional molecular features uncovered by complementary growth conditions and analytical setups directly translate into increased chances for the discovery of new natural products. Thus, we argue that the simplification of screening workflows should not come at the cost of decreased metabolomics characterization, as this not only paints an incomplete picture of the bacterial secondary metabolome complexity but also results in missed compelling target molecules.

Data availability: The recorded MS data, molecular networking files, and other files used for data evaluation are available for download at ChemRxiv.org.

References

(1) U.S. Centers for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States, 2019*.

(2) Bernardini, S.; Tiezzi, A.; Laghezza Masci, V.; Ovidi, E. Natural products for human health: an historical overview of the drug discovery approaches. *Nat. Prod. Res.* **2018**, *32*, 1926–1950.

(3) Baker, D. D.; Chu, M.; Oza, U.; Rajgarhia, V. The value of natural products to future pharmaceutical discovery. *Nat. Prod. Rep.* **2007**, *24*, 1225–1244.

(4) Bader, C. D.; Panter, F.; Müller, R. In depth natural product discovery - Myxobacterial strains that provided multiple secondary metabolites. *Biotechnol Adv.* **2020**, *39*, 107480.

(5) Wenzel, S. C.; Müller, R. Myxobacteria - unique microbial secondary metabolite factories. In *Comprehensive Natural Products Chemistry II, Vol 2: Structural Diversity II - Secondary Metabolite Sources, Evolution and Selected Molecular Structures*; Moore, B.S., Ed.; Elsevier: Oxford, 2010; pp 189–222.

(6) Schäberle, T. F.; Lohr, F.; Schmitz, A.; König, G. M. Antibiotics from myxobacteria. *Nat. Prod. Rep.* **2014**, *epub ahead of print*.

(7) Herrmann, J.; Fayad, A. A.; Müller, R. Natural products from myxobacteria: novel metabolites and bioactivities. *Nat. Prod. Rep.* **2017**, *34*, 135–160.

(8) Dehhaghi, M.; Mohammadipanah, F.; Guillemin, G. J. Myxobacterial natural products: An under-valued source of products for drug discovery for neurological disorders. *Neurotoxicology* **2018**, *66*, 195–203.

(9) Irschik, H.; Jansen, R.; Höfle, G.; Gerth, K.; Reichenbach, H. The corallopyronins, new inhibitors of bacterial RNA synthesis from Myxobacteria. *J. Antibiot.* **1985**, *38*, 145–152.

(10) Irschik, H.; Gerth, K.; Höfle, G.; Kohl, W.; Reichenbach, H. The myxopyronins, new inhibitors of bacterial RNA synthesis from *Myxococcus fulvus* (Myxobacterales). *J. Antibiot.* **1983**, *36*, 1651–1658.

(11) Yamaguchi, H.; Nakayama, Y.; Takeda, K.; Tawara, K.; Maeda, K.; Takeuchi, T.; Umezawa, H. A New Antibiotic, Althiomycin. *J. Antibiot.* **1957**, *10*, 195–200.

(12) Irschik, H.; Reichenbach, H.; Höfle, G.; Jansen, R. The thuggacins, novel antibacterial macrolides from Sorangium cellulosum acting against selected Gram-positive bacteria. *J Antibiot.* **2007**, *60*, 733–738.

(13) Steinmetz, H.; Mohr, K. I.; Zander, W.; Jansen, R.; Gerth, K.; Müller, R. Indiacens A and B: prenyl indoles from the *myxobacterium Sandaracinus amylolyticus*. *J. Nat. Prod.* **2012**, *75*, 1803–1805.

(14) Corre, C.; Challis, G. L. Exploiting Genomics for New Natural Product Discovery in Prokaryotes. In *Comprehensive Natural Products Chemistry II, Vol 2: Structural Diversity II - Secondary Metabolite Sources, Evolution and Selected Molecular Structures*; Moore, B.S., Ed.; Elsevier: Oxford, 2010; pp 429–453.

(15) Adamek, M.; Spohn, M.; Stegmann, E.; Ziemert, N. Mining Bacterial Genomes for Secondary Metabolite Gene Clusters. In *Antibiotics: Methods and protocols / edited by Peter Sass*; Sass, P., Ed.; Springer protocols 1520; Humana Press: New York, 2017; pp 23–47.

(16) Bachmann, B. O.; van Lanen, S. G.; Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 175–184.

(17) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, W81-W87.

(18) Fraser, C. M.; Eisen, J. A.; Salzberg, S. L. Microbial genome sequencing. *Nature* **2000**, *406*, 799–803.

(19) Fraser, C. M.; Eisen, J. A.; Nelson, K. E.; Paulsen, I. T.; Salzberg, S. L. The Value of Complete Microbial Genome Sequencing (You Get What You Pay For). *J. Bacteriol.* **2002**, *184*, 6403–6405.

(20) Wenzel, S. C.; Müller, R. The biosynthetic potential of myxobacteria and their impact on drug discovery. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 220–230.

(21) Jensen, P. R.; Chavarria, K. L.; Fenical, W.; Moore, B. S.; Ziemert, N. Challenges and triumphs to genomics-based natural product discovery (Review). *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 203–209.

(22) Adnani, N.; Chevrette, M.; Adibhatla, S. N.; Zhang, F.; Yu, Q.; Braun, D. R.; Nelson, J.; Simpkins, S. W.; McDonald, B. R.; Myers, C. L.*; et al.* Co-culture of Marine Invertebrate-Associated Bacteria and Interdisciplinary Technologies Enable Biosynthesis and Discovery of a New Antibiotic, Keyicin. *ACS Chem. Biol.* **2017**, *12*, 3093–3102.

(23) Piddock, L. J. V. Teixobactin, the first of a new class of antibiotics discovered by iChip technology? *J Antimicrob Chemother* **2015**, *70*, 2679–2680.

(24) Li, P. F.; Li, S. G.; Li, Z. F.; Zhao, L.; Wang, T.; Pan, H. W.; Liu, H.; Wu, Z. H.; Li, Y. Z. Co-cultivation of *Sorangium cellulosum* strains affects cellular growth and biosynthesis of secondary metabolite epothilones. *FEMS Microbiol. Ecol.* **2013**, *85*, 358–368.

(25) Brakhage, A. A.; Schuemann, J.; Bergmann, S.; Scherlach, K.; Schroeckh, V.; Hertweck, C. Activation of fungal silent gene clusters: a new avenue to drug discovery. *Prog. Drug Res.* **2008**, *66*, 3–12.

(26) Panter, F.; Krug, D.; Baumann, S.; Müller, R. Self-resistance guided genome mining uncovers new topoisomerase inhibitors from myxobacteria. *Chem. Sci.* **2018**, *9*, 4898–4908.

(27) Saito-Shida, S.; Hamasaka, T.; Nemoto, S.; Akiyama, H. Multiresidue determination of pesticides in tea by liquid chromatography-high-resolution mass spectrometry: Comparison between Orbitrap and time-of-flight mass analyzers. *Food chemistry* **2018**, *256*, 140–148.

(28) Fels, H.; Dame, T.; Sachs, H.; Musshoff, F. Liquid chromatography-quadrupole-time-of-flight mass spectrometry screening procedure for urine samples in forensic casework compared to gas chromatography-mass spectrometry. *Drug testing and analysis* **2017**, *9*, 824–830.

(29) Boesl, U. Time-of-flight mass spectrometry: Introduction to the basics. *Mass spectrometry reviews* **2017**, *36*, 86–109.

(30) Kim, J.; Choi, J. N.; Kim, P.; Sok, D. E.; Nam, S. W.; Lee, C. H. LC-MS/MS Profiling-Based Secondary Metabolite Screening of *Myxococcus xanthus*. *J. Microbiol. Biotechnol.* **2009**, *19*, 51–54.

(31) Hoffmann, T.; Krug, D.; Hüttel, S.; Müller, R. Improving natural products identification through targeted LC-MS/MS in an untargeted secondary metabolomics workflow. *Anal. Chem.* **2014**, *86*, 10780–10788.

(32) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Xueting, L.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; Felicio, R. de; Fenner, A.*; et al.* Molecular networking as a dereplication strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699.

(33) Sleighter, R. L.; Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **2007**, *42*, 559–574.
(34) Brown, S. C.; Kruppa, G.; Dasseux, J.-L. Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom. Rev.* **2005**, *24*, 223–231.
(35) Sommella, E.; Conte, G. M.; Salviati, E.; Pepe, G.; Bertamino, A.; Ostacolo, C.; Sansone, F.; Prete, F. D.; Aquino, R. P.; Campiglia, P. Fast Profiling of Natural Pigments in Different Spirulina (Arthrospira platensis) Dietary Supplements by DI-FT-ICR and Evaluation of their Antioxidant Potential by Pre-Column DPPH-UHPLC Assay. *Molecules (Basel, Switzerland)* **2018**, *23*.
(36) Wu, Z.; Rodgers, R. P.; Marshall, A. G. Characterization of vegetable oils: detailed compositional fingerprints derived from electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *J. Agric. Food Chem.* **2004**, *52*, 5322–5328.
(37) Park, K. H.; Kim, M. S.; Baek, S. J.; Bae, I. H.; Seo, S.-W.; Kim, J.; Shin, Y. K.; Lee, Y.-M.; Kim, H. S. Simultaneous molecular formula determinations of natural compounds in a plant extract using 15 T Fourier transform ion cyclotron resonance mass spectrometry. *Plant methods* **2013**, *9*, 15.
(38) Witt, M.; Barsch, A.; Wolff, J.; Krug, D.; Hoffmann, T.; Müller, R. *Ultrafast Statistical Profiling of Bacterial Metabolite Extracts*, 2013.
(39) Fernandez-Lima, F. A.; Becker, C.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G.; Russell, D. H. Petroleum crude oil characterization by IMS-MS and FTICR MS. *Anal. Chem.* **2009**, *81*, 9941–9947.
(40) Annesley, T. M. Ion Suppression in Mass Spectrometry. *Clin. Chem.* **2003**, *49*, 1041–1044.
(41) Southam, A. D.; Payne, T. G.; Cooper, H. J.; Arvanitis, T. N.; Viant, M. R. Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal. Chem.* **2007**, *79*, 4595–4602.
(42) Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M. Ion suppression; a critical review on causes, evaluation, prevention and applications. *Talanta* **2013**, *115*, 104–122.
(43) Muñoz-Dorado, J.; Marcos-Torres, F. J.; García-Bravo, E.; Moraleda-Muñoz, A.; Pérez, J. Myxobacteria: Moving, Killing, Feeding, and Surviving Together. *Front. Microbiol.* **2016**, *7*.
(44) Hoffmann, M.; Auerbach, D.; Panter, F.; Hoffmann, T.; Dorrestein, P. C.; Müller, R. Homospermidine Lipids: A compound class specifically formed during fruiting body formation of *Myxococcus xanthus* DK1622. *ACS Chem. Biol.* **2018**, *13*, 273–280.
(45) Panter, F.; Krug, D.; Müller, R. Novel Methoxymethacrylate Natural Products Uncovered by Statistics-Based Mining of the *Myxococcus fulvus* Secondary Metabolome. *ACS Chem. Biol.* **2019**, *14*, 88–98.
(46) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.*; et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837.
(47) Gerth, K.; Irschik, H.; Reichenbach, H.; Trowitzsch, W. The myxovirescins, a family of antibiotics from Myxococcus virescens (Myxobacterales). *J. Antibiot.* **1982**, *35*, 1454–1459.
(48) Hug, J. J.; Dastbaz, J.; Adam, S.; Revermann, O.; Koehnke, J.; Krug, D.; Mueller, R. Biosynthesis of cittilins, unusual ribosomally synthesized and post-translationally modified peptides from *Myxococcus xanthus*. *ACS Chem. Biol.* **2020**, *submitted*.
(49) Meiser, P.; Bode, H. B.; Müller, R. The unique DKxanthene secondary metabolite family from the myxobacterium *Myxococcus xanthus* is required for developmental sporulation. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19128–19133.
(50) Gerth, K.; Jansen, R.; Reifenstahl, G.; Höfle, G.; Irschik, H.; Kunze, B.; Reichenbach, H.; Thierbach, G. The myxalamids, new antibiotics from *Myxococcus xanthus* (Myxobacterales) I. production, physico-chemical and biological properties, and mechanism of action. *J Antibiot* **1983**, *36*, 1150–1156.
(51) Trowitzsch Kienast, W.; Gerth, K.; Reichenbach, H.; Höfle, G. Myxochromid A: Ein hochungesättigtes Lipopeptidlacton aus *Myxococcus virescens*. *Liebigs Ann. Chem.* **1993**, 1233–1237.
(52) Krug, D.; Garcia, R.; Müller, R. Myxobakterielle Naturstofffabriken. *Biospektrum* **2020**, *26*, 32–36.
(53) Krug, D.; Zurek, G.; Revermann, O.; Vos, M.; Velicer, G. J.; Müller, R. Discovering the Hidden Secondary Metabolome of *Myxococcus xanthus*: a Study of Intraspecific Diversity. *Appl. Environ. Microbiol.* **2008**, *74*, 3058–3068.
(54) Nordström, A.; Want, E.; Northen, T.; Lehtiö, J.; Siuzdak, G. Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics. *Anal. Chem.* **2008**, *80*, 421–429.
(55) Oss, M.; Kruve, A.; Herodes, K.; Leito, I. Electrospray ionization efficiency scale of organic compounds. *Anal. Chem.* **2010**, *82*, 2865–2872.
(56) Huffman, B. A.; Poltash, M. L.; Hughey, C. A. Effect of polar protic and polar aprotic solvents on negative-ion electrospray ionization and chromatographic separation of small acidic molecules. *Anal. Chem.* **2012**, *84*, 9942–9950.