Application Note # ET-29

Screening for novel natural products from myxobacteria using LC-MS and LC-NMR

Introduction

Because known pathogens become increasingly resistant to antibiotics, and at the same time, new threats to human health emerge, there is a constant need for the discovery of novel natural products. Microorganisms have a long-standing tradition as a source of biologically active small molecules – termed “secondary metabolites” – and in addition to well-known producers like the actinomycetes, new sources have been established.

We describe here the application of LC-MS and LC-NMR techniques for the discovery and structural elucidation of novel natural products from myxobacteria. These soil-living bacteria exhibit outstanding biological characteristics such as cooperative swarming on surfaces and formation of multicellular fruiting bodies (Fig. 1). In addition, many myxobacterial strains are also producers of secondary metabolites exhibiting a wide range of biological activities [1] – examples include the antibacterial thuggacins, antifungal compounds like ambruticin and the cytotoxic epothilones.

Authors

Daniel Krug, Kirsten Harmrolfs, Thomas Hoffmann, Niña Cortina and Rolf Müller, Helmholtz-Institute for Pharmaceutical Research, Germany

Read more

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Instrumentation and Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolomics</td>
<td>maXis 4G</td>
</tr>
<tr>
<td>Identification</td>
<td>Avance 700 NMR</td>
</tr>
<tr>
<td>Targeted Analysis</td>
<td>Profile Analysis</td>
</tr>
<tr>
<td>Untargeted Profiling</td>
<td>Target Analysis</td>
</tr>
<tr>
<td>LC-SPE-NMR</td>
<td>Dereplication</td>
</tr>
</tbody>
</table>
Dereplication using maXis 4G UHR-TOF MS measurements

One straightforward and widely used method for discovering bioactive secondary metabolites involves the analysis of crude extracts from small-scale myxobacteria cultivations. These are tested in bioassays using a panel of indicator organisms. Activity-guided fractionation of the extract and subsequent purification steps using liquid chromatography finally yield the active compound, which then typically undergoes full structure elucidation by combining NMR spectroscopy with MS techniques. However, this approach critically depends on efficient early detection of known compounds, termed “dereplication”, prior to bioactivity screening in order to avoid the time-consuming re-isolation of previously identified compounds. High-resolution mass spectrometry offers a powerful solution for obtaining a comprehensive overview of the secondary metabolite content before biological testing. The target screening procedure is based on the combined evaluation of exact mass, isotope pattern and retention time and provides a convenient first-pass screening result, using full scan MS data. Following the pre-treatment of raw datasets by compound-finding, the obtained molecular features are matched against calculated spectra of all known myxobacterial compounds from the in-house database (Fig. 2). The mass accuracy and isotope pattern fidelity offered by the maXis UHR-TOF MS are advantageous for generating target screening reports with increased confidence (Fig. 3).

Statistical data evaluation enables metabolome-mining

Besides activity-guided isolation, genomics-based strategies are becoming increasingly successful in revealing additional natural products. These approaches hold great promise for uncovering novel secondary metabolites from myxobacterial strains, as the number of known compounds identified to date is often significantly lower than that expected from whole-genome sequence information. However, attempts to identify these “hidden” natural products from complex secondary metabolomes frequently encounter considerable analytical challenges due to low initial productivity of the bacterial strain used as a model organism. These so-called “genome mining” studies require sensitive and highly reproducible analysis in combination with high resolution, accuracy and extended sample throughput.

High-resolution mass spectrometry plays a crucial role for bridging the apparent gap between genomic potential and secondary metabolome reality, by enabling a powerful combination of targeted and untargeted analysis. The search for novel myxobacterial secondary metabolites using a comprehensive “metabolome mining” approach is exemplified by the recent discovery of myxoprincomide from the myxobacterial model strain Myxococcus xanthus DK1622 (Fig. 4) [2]. Replicate small-scale fermentations of the wild type strain and mutants created by gene-inactivation experiments were performed. Extracts from all test cultures were separated under UHPLC conditions with coupling to the maXis UHR-TOF mass spectrometer. LC-MS data were subjected to compound-finding and subsequently compared using Principal Component Analysis (PCA), in order to mine the complex datasets for possibly subtle but meaningful systematic differences.
Statistical analysis highlighted a minor signal with 506.271 m/z, representing a doubly-charged ion, as the most significant difference between wild type and mutant extracts. Metabolomic data from a collection of *M. xanthus* wild type strains were then used to select a better producer for this previously unknown metabolite [3]. The yield of the target compound was improved by genetic manipulation of the producing strain prior to upscaling and compound isolation. Purification from a large scale fermentation finally allowed full structural elucidation of myxoprincomide by NMR experiments. At the same time, the assignment of this novel secondary metabolite to a biosynthetic gene cluster with previously unknown function paves the way to investigate the biosynthetic machinery underlying its formation [2].

**Structure elucidation of novel metabolites facilitated by LC-SPE-NMR**

Genomics-based mining of myxobacterial metabolomes for novel natural products exhibits a high probability for the discovery of entirely new structural scaffolds. Similar chances are not expected for the activity-guided isolation strategy: targeted screening can effectively prevent the re-isolation of specific known compounds, but still there remains a significant risk that a putatively novel metabolite could be revealed – following time-consuming large-scale purification – as “only” a distinct derivative of a known compound class. Tandem-MS data can help in some cases to shed light on suspected structural relationship. The most important key for the rapid identification of known compound classes, however, is the ability to generate a partial (or even complete) structure based on NMR information as early as possible in the course of the discovery workflow. LC-SPE-NMR is an important tool to achieve this objective. After initial separation of a myxobacterial extract obtained from small-scale fermentation into crude fractions by size-exclusion chromatography, analytical HPLC with MS monitoring is used to track fractions of interest. Candidate peaks are then trapped on SPE cartridges and transferred via a Cryofit flowprobe to the NMR system using deuterated solvent. The acquisition of 1D and 2D NMR spectra sets from small scales (~ 20 µg) was achieved by applying multiple solvent suppression pulse programs. This process facilitates the rapid identification of known and novel structural scaffolds (Fig. 5). Thus, a well-founded decision whether to include a candidate compound into the labor-intensive upscaling, fermentation and purification pipeline can be rapidly made.

---

**Dereplication: Detection of known compounds**

Figure 2: Mass precision and isotope pattern accuracy are crucial for the identification of natural products from myxobacteria during first-pass screening.

Chromatogram: Separation of a crude extract from a *Sorangium* strain. Right inset: Comparison of measured and calculated spectra for Ambruticin, a myxobacterial compound that exhibits potent antifungal activity.
Conclusion
Microbial natural products, as exemplified here by secondary metabolites from myxobacteria, represent a rich and still largely untapped resource for novel chemical scaffolds. Their discovery enables characterization of the underlying biosynthetic pathways and at the same time chances are good to find compounds with potent biological activity. UHPLC-coupled UHR-TOF mass spectrometry and LC-NMR represent most valuable analytical tools for the discovery of novel natural products, supporting the screening workflow with efficient dereplication, comprehensive mining of metabolomes and rapid structure elucidation of novel metabolites.

Methods

- TOF-MS: Bruker Daltonics maXis 4G equipped with electrospray ion source
- LC/MS (analytical): Dionex RSLC 3000 system, using a Waters Acquity C-18 100 x 2 mm (1.7µm) column and precolumn
- LC/MS conditions: Linear gradient, runtime 19 min + 5 min post-run time, flow rate 0.6 mL/min, column temperature 45°C, injection volume 2 µL, solvent A: 0.1% formic acid, solvent B: acetonitrile with 0.1% formic acid, flow split 1:8 into ESI source (75 µL/min)
- LC/MS/NMR: VWR 2400 series solvent delivery system, using a Phenomenex Luna C-18 150 x 4.6 mm (2.5 µm) column, SPE unit and Bruker Biospin Avance 700 NMR platform with cryoprobe

TOF-MS instrument settings

- Positive ion polarity, capillary voltage 4500 V, endplate offset -500 V, nebulizer 1.5 bar, dry gas flow 6 L/min, dry heater 200°C, flight tube voltage 9000 V, detector voltage 2200 V, sample rate 2 Hz

Parameters for targeted analysis

A screening library containing molecular formulae and experimentally-determined retention times of ions typically observed from myxobacterial natural products was generated by an export function from Myxobase, a specialized and constantly updated in-house database. The TargetAnalysis software was used with a retention time window of 0.6 min, mass window 3 Da, mSigma value < 60.

Statistical treatment for multivariate analysis

The ProfileAnalysis software was used to perform principal component analysis (PCA). Following molecular feature extraction, advanced bucketing was carried out in the retention time range 2-18 min spanning the mass range 200-2000 m/z.

NMR measurements

Fractions containing target compounds were collected on SPE cartridges (Hysphere resin GP), dried under nitrogen flow and eluted into the NMR flowprobe system using deuterated solvent.

Target screening report

The above table provides a summary of targeted screening results. The compounds were identified based on their mass-to-charge ratio (m/z) and retention time (RT). The table includes columns for compound name, molecular formula, m/z range, and retention time. The results are analyzed using statistical tools, such as principal component analysis (PCA), to identify and quantify the target compounds.

Figure 3: Target screening report generated by TargetAnalysis for a Myccococcus xanthus extract, highlighting the presence of several known myxobacterial secondary metabolites.
Rapid structure elucidation combining MS and NMR

Identification of Myxoprincomides by „metabolome-mining“

Figure 4: Statistical treatment of extracts derived from wildtype and mutant strains of *Myxococcus xanthus*. Principal component analysis (PCA, plot shows PC1 vs. PC) facilitates the discovery of a novel compound class, the Myxoprincomides, by using a combined genome- and metabolome-mining strategy.

Rapid structure elucidation combining MS and NMR

Figure 5: Rapid structure elucidation of metabolites from myxobacterial crude extracts enabled by LC-SPE-NMR, following semi-preparative liquid chromatography coupled to MS- and UV/vis detection. The NMR spectrum (right path) corresponds to a new derivative of Myxochelin.
Figure 6: Natural products screening and structure elucidation platform at HIPS (Helmholtz-Institute for Pharmaceutical Research Saarland), consisting of the maXis4G UHPLC-UHR-TOF system, semipreparative LC-SPE unit, and 700 MHz cryoprobe NMR system.

References


For research use only. Not for use in diagnostic procedures.