

Targeted Approaches in Natural Product Lead Discovery

Olivier Potterat*

Abstract: Natural products still constitute a prolific source of lead compounds for the development of novel drugs. Due to their unmatched structural diversity they represent an indispensable complement to synthetic compound collections and virtual libraries. However, the tremendous changes which have taken place in the drug discovery process over the last decade have introduced completely new requirements for natural product research. The classical bioactivity-guided approach no longer matches the timelines and the workflow of modern drug discovery. More targeted approaches are needed in order to identify promising candidate molecules at an early stage, and to accelerate the follow-up process. We describe herein our strategy for tracking bioactive components in complex mixtures, and discuss the potential of accurate mass measurements for natural product identification and the impact of mass-controlled preparative HPLC on the isolation process of bioactive compounds.

Keywords: HPLC-based activity profiling · Mass-controlled preparative HPLC · Natural products · Time-of-flight mass spectrometry

Introduction

Natural products have traditionally played a major role in drug discovery and still constitute a prolific source of novel lead compounds for medicinal chemistry. About 40% of current prescription drugs are directly or indirectly derived from compounds of biogenic origin [1][2]. Molecules originating from natural products leads are well represented among the 35 top selling prescription drugs worldwide, with percentages ranging from approximately 20 to 40% over the last five years. Despite increasing competition from combinatorial and classical com-

ound libraries, there has been a steady introduction of natural product-derived drugs in the last years. According to a recent review [3], a total of 15 natural products or natural product-derived drugs has been launched in either the United States, Europe or Japan in the years 2000–2003. Among these are essential medicines such as the anti-malaria drug arteether (Artemotil®), galanthamine (Reminyl®) for the treatment of Alzheimer's disease, and pimecrolimus (Elidel®) for atopic dermatitis. The impact of natural products on the development pipelines of the pharmaceutical industry is unabated. As per December 2003, some 15 compounds were in phase III clinical trials or registration, including representatives of novel compound classes such as the anti-cancer drug ixabepilone, a semisynthetic derivative of epothilone B produced by the myxobacterium *Sorangium cellulosum*. Ziconotide, the synthetic equivalent of ω -conotoxin found in the dart venom of the cone shell *Conus magus*, was approved by the FDA in December 2004 for the treatment of chronic pain [4].

One of the most striking features of natural products is their unmet and still largely untapped structural diversity. In the post-genomic era with its increasing number of druggable targets, natural products remain a rather indispensable complement to synthetic compound collections. Natural products in particular are sterically more complex and differ from synthetic compounds

with respect to the statistical distribution of functionalities [5]. They cover a much larger volume of the chemical space and display a broader dispersion of structural and physicochemical properties than compounds issued from combinatorial synthesis [6]. However, the accelerated transformation of the drug discovery process which has taken place over the last decade has put completely new requirements on natural product research. In particular, the classical and historically successful approach of screening crude or pre-purified extracts followed by several iterative steps of activity-guided fractionation [7] does not match with the short target cycle times in HTS, where testing capacities are only provided for a limited time window. As further pitfalls, the loss of bioactivity in the course of the purification process is not uncommon, and there is little means for early dereplication of known or otherwise uninteresting compounds. There is clearly a compelling need for faster and more effective strategies, delivering early information on the bioactive constituents, and compatible with the modern drug discovery workflow.

HPLC-based Activity Profiling

Natural product extracts are highly complex mixtures containing hundreds to thousands of constituents, among which only one or just a few are responsible for

*Correspondence: Dr. O. Potterat
Institute of Pharmaceutical Biology
Department of Pharmaceutical Sciences
Pharmazentrum
University of Basel
Klingelbergstrasse 50
CH-4056 Basel
Tel.: +41 61 267 15 64
Fax: +41 61 267 14 74
E-Mail: Olivier.Potterat@unibas.ch
www.unibas.ch

the pharmacological properties. Probably the most challenging task in this context is to combine biological data and chemo-analytical information in a judicious manner in order to identify the active principles at an early stage [8]. The generalization of miniaturized and highly sensitive assays has recently opened new possibilities. HPLC-based activity profiling is one of the most successful approaches, and is widely used in our laboratory for tracking bioactive compounds in crude mixtures. It consists of fractionating bioactive extracts in 96-format by means of analytical or semipreparative HPLC and subsequently assessing the bioactivity of each fraction. In this way, the activity can be assigned to chromatographic peaks and correlated with the spectroscopic information available on-line. The principle is shown in Fig. 1. An extract is separated by analytical gradient HPLC. *Via* a T-split, a portion of the effluent is fractionated in the 96-well format, while the other part serves for the on-line spectroscopic characterization of the eluted peaks. After drying, the fractions are redissolved in a small amount of a suitable solvent, typically DMSO, and assayed for bioactivity. The activity profile is then matched with the HPLC chromatogram and the spectroscopic data recorded on-line. A targeted preparative isolation is subsequently carried out if the active principles are deemed of sufficient interest. As a major asset, the microfractionation step can be performed with the bioactive sample stored in the library, without the need for macroscopic isolation nor time-consuming re-fermentation or recollection activities, and the associated issues of unsatisfactory reproducibility. Preparative purification is then performed using a straightforward peak-guided strategy which does not require biological testing after each chromatographic step. The scale-up of the separation is straightforward, since the chromatographic system can be easily transposed to preparative scale. Compared to other innovative setups which have been recently developed for the on-line integration of biological and chemo-analytical systems [8][9], HPLC-based profiling is particularly versatile and directly applicable to a broad range of mechanism-based and cellular assays. For this reason, it has also become the method of choice in industrial natural product screening programs.

Fig. 2 shows an application of this approach in the course of a screening program for novel glucagon receptor antagonists. An extract of *Streptomyces* sp. strongly inhibited glucagon-induced cAMP elevation. Semipreparative gradient HPLC separation of 30 μ l of extract into 30 one-minute fractions and subsequent testing of each fraction enabled the activity to be attributed to the main peak at 16.1 min. HPLC-UV-MS analysis revealed a UV-spectrum similar to

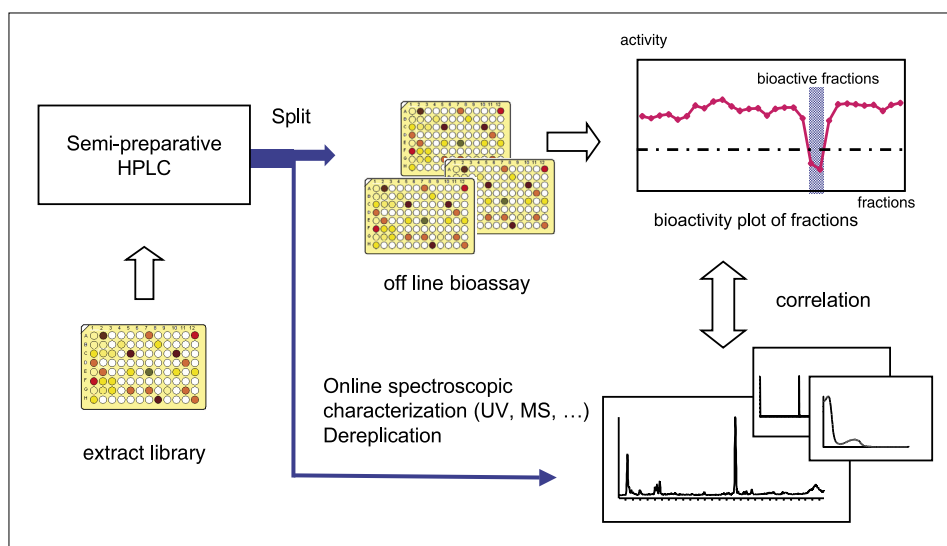


Fig. 1. Schematic representation of HPLC-based extract profiling

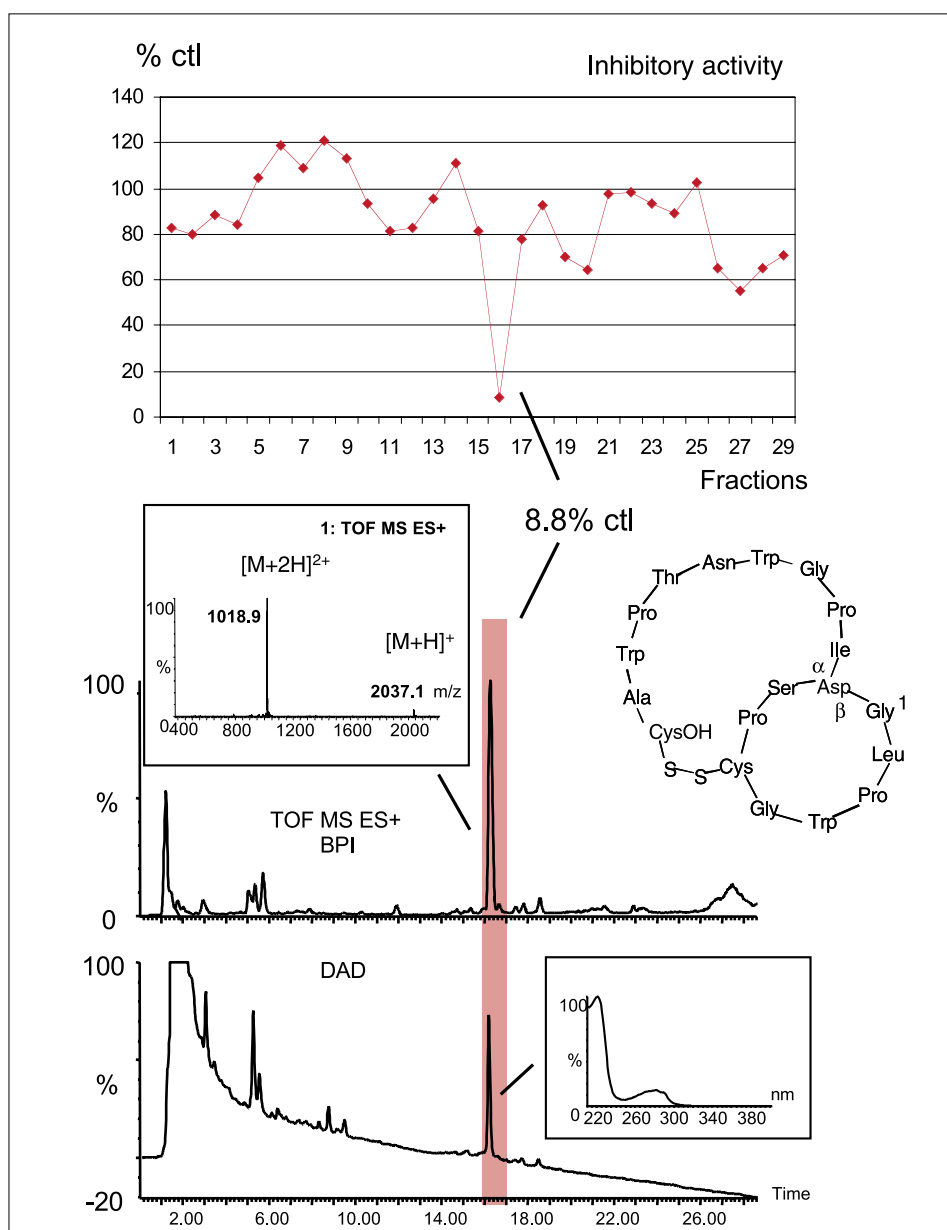


Fig. 2. Activity profiling for glucagon receptor inhibitory activity in a streptomycete extract: Activity plot, HPLC/MS/UV analysis and structure of the bioactive constituent [10]

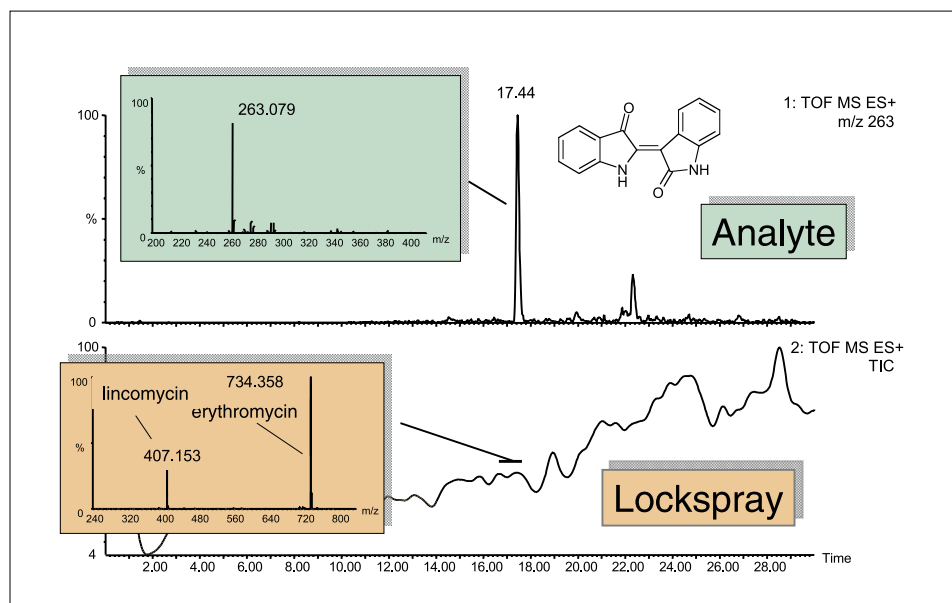


Fig. 3. Lockspray HPLC-ESI TOF MS analysis of indirubin in a crude extract of *Calanthe tricarinata* (Orchidaceae). The accurate mass of indirubin is determined by correcting the measured mass with the expected values for lincomycin (407.2216) or erythromycin (734.4690).

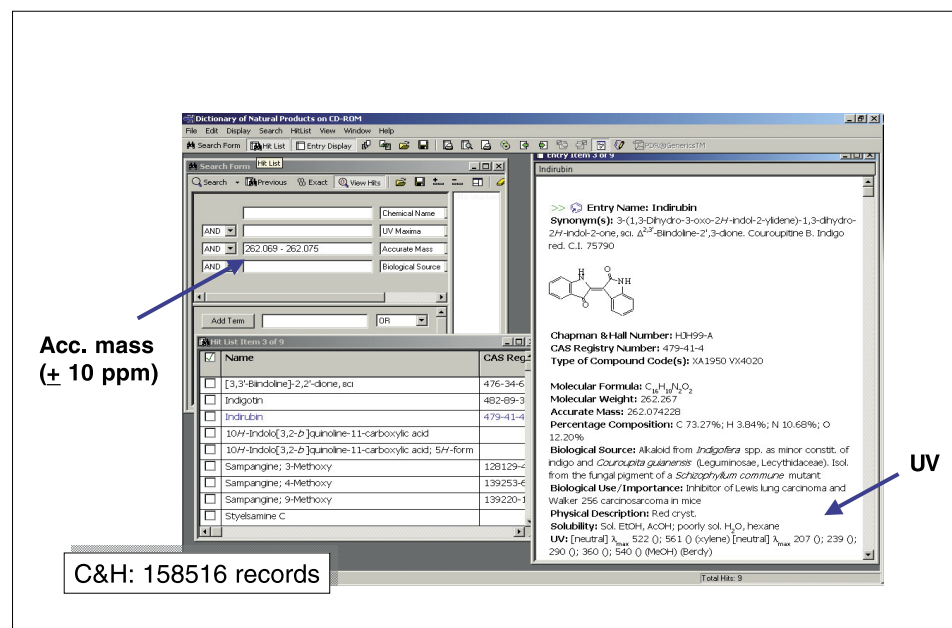


Fig. 4. Identification of indirubin from the accurate MS and UV data with the Dictionary of Natural Products (Chapman & Hall)

that of tryptophan and a MW of 2036 amu. Since no record corresponding to these data could be found in the literature, purification of the compound was undertaken on a preparative scale. The compound, a new bicyclic peptide, exhibited potent and selective antagonist activity towards the human glucagon receptor in a functional assay (IC₅₀ 0.44 μM) [10]. In this project, the micro-scale fractionation also reliably filtered out other extracts in which the inhibitory activity could be assigned to known compounds of little interest, such as cephalochromin or resistomycin.

On-line Accurate Mass Measurements and Database Search

In recent years, the impact of HPLC-coupled spectroscopy on natural product analysis has been tremendous. The concerted use of HPLC-DAD, -MS and -NMR has opened entirely new possibilities for the characterization of secondary metabolites in biological extracts. These techniques provide a wealth of structural information on-line from minute amounts of sample [11]. Nevertheless, rapid identification

and high-throughput dereplication of compounds remains challenging when no preliminary information is available. Not all spectroscopic data are equally well suited to database searching. Moreover, full integration of spectroscopic data of different types and formats remains unsatisfactory.

The molecular formula, which can be deduced from an accurate mass determination, is one of the most useful pieces of information for searching large databases such as CAS Registry. It can now be routinely obtained with time of flight (TOF) mass spectrometers which are increasingly used for HPLC-MS analysis of natural products [12][13]. TOF mass spectrometry combines high accuracy with high sensitivity due to high-frequency sampling of all ions simultaneously in the full mass range. By using an external reference, which can be introduced into the source in alternation with the analyte through a second spray (lockspray), accuracy down to ca. 3 ppm can be routinely achieved.

While nominal mass and fragmentation information provided by quadrupole instruments often proves to be difficult to use for database searching and is insufficient for compound identification, the mass accuracy obtained by HPLC-ESI-TOF-MS enables one to calculate a limited number of elemental composition hypotheses. By combining this information with additional data (UV spectral data, taxonomy of the producer), identification of a large number of known substances is possible. Moreover, accurate mass and, as a corollary, molecular formula are well suited to automated database searching, a condition for high-throughput dereplication.

The identification of indirubin in a crude extract of the orchid *Calanthe tricarinata* (Orchidaceae) is shown as an example of this approach. By using a lockspray device, two separate chromatograms were recorded for the analyte and the reference compounds, respectively (Fig. 3). The mass observed for the pseudomolecular [M+H]⁺ ion of the peak at 17.44 min (263.079 amu) was corrected with the data simultaneously recorded for the external reference lincomycin. A search in the Dictionary of Natural Products (DNP) database (Chapman & Hall) with the accurate mass of 262.072 amu and a mass tolerance of 10 ppm gave a match list of only nine compounds. Inclusion of the UV spectral data then unambiguously identified indirubin (Fig. 4).

Preparative HPLC-MS

HPLC-based activity profiling in conjunction with hyphenated spectroscopic methods enables promising compounds to be tracked in crude extracts, and often their chemical structures can be determined on-

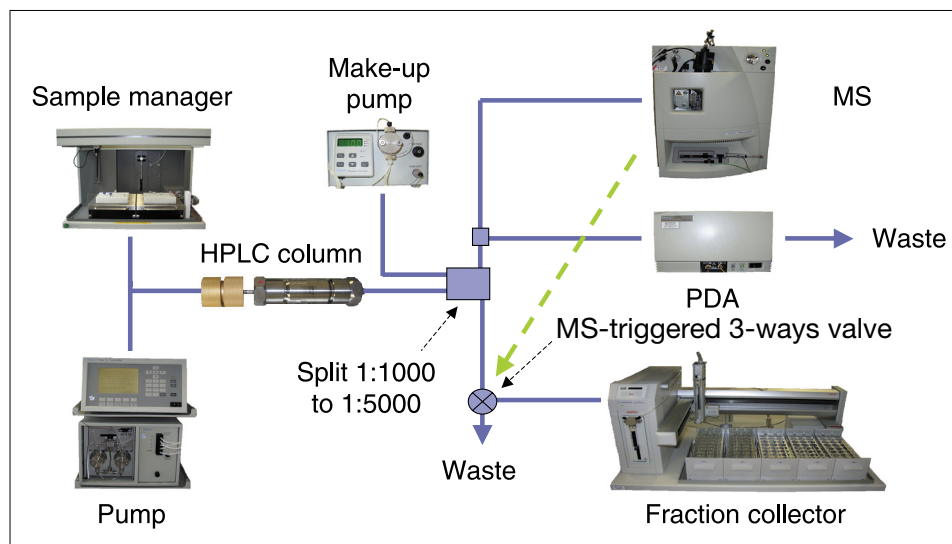


Fig. 5. Typical preparative HPLC-MS setup

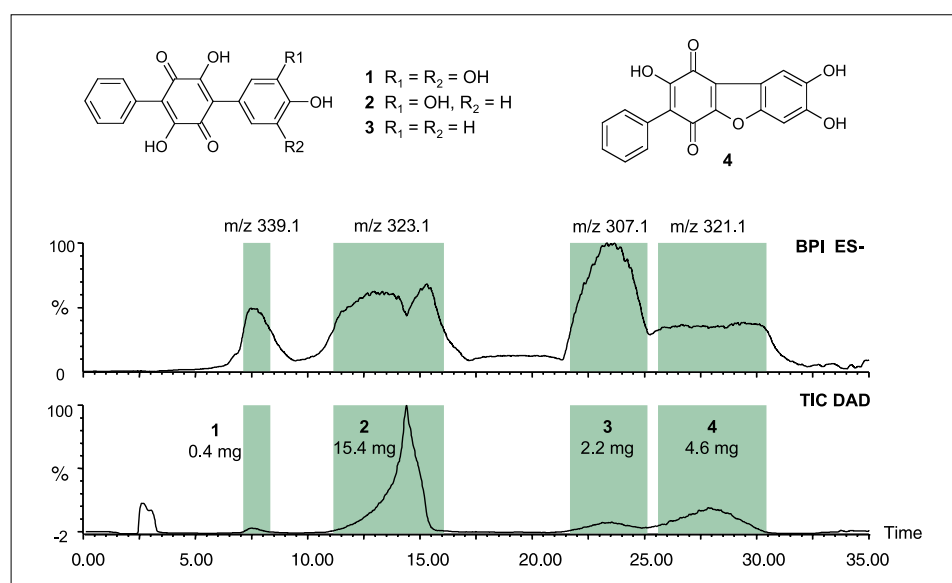


Fig 6. Preparative HPLC-MS separation of four terphenylquinones from a crude fraction of the fungus *Stilbella* sp. [14]. C18 column (30 × 100 mm i.d.), 0.02% aq. AcOH-MeCN (28:72), flow rate: 28 ml/min, applied amount: 32 mg in 0.5 ml MeOH. Fractions collected automatically are highlighted.

line without requiring any time- and cost-consuming isolation work. However, preparative isolation remains indispensable for detailed biological characterization of potential lead compounds. The emergence of mass spectrometry-controlled preparative HPLC has recently led to a paradigm change in preparative chromatography. The high sensitivity and selectivity of mass detection enables a highly efficient purification of compounds directly from complex matrices, even when they are present only as minor components. Since this technology is well suited to automation and unattended operation, large autopurification platforms combining both mass and UV triggering modes have been set up, in particular in the context of industrial drug discovery. A typical setup is shown in Fig. 5. The eluent is split after the HPLC column

in a ratio of 1:1000 to 1:5000 between the collection and detection lines. A make-up pump is used to create a sufficient flow in the detection line and dilute the eluent in order to avoid detector saturation. The delay between detection and collection (typically 10–25 s) is then determined as a function of the flow rate. The DAD detector and the mass spectrometer can be coupled in parallel, as shown in the Fig., or in series. When the target mass signal exceeds a given threshold, a three-way valve, triggered by the mass spectrometer, switches the eluent flow from the waste line to the fraction collector. For additional specificity, mass and UV collection criteria can be combined using the logical operators ‘and’, ‘and not’, and ‘or’.

The purification of four terphenylquinones with src inhibiting activity from the

fungus *Stilbella* sp. is shown in Fig. 6 as an example [14]. The four compounds could be obtained in a spectroscopically pure form in a single step. Up to 32 mg of the fraction could be separated per chromatographic run. Mass triggered detection allowed optimal fraction cutting even when some compounds gave asymmetric and incompletely resolved peaks.

Conclusion

The approaches described here are part of numerous innovative technological developments which have taken place during the last decade [8]. The unique diversity of natural products and their successful track record are undisputed. However, continuing improvements will be needed if natural products research should keep pace with the ongoing changes in drug discovery, in particular with the unbroken trend towards miniaturization and shorter project cycle times. Further improvement in combining chemoanalytical and biological data, and full automation of the dereplication process are probably the greatest challenges for the next few years. In view of recent achievements and ongoing developments it is reasonable to think that, despite increased competition from synthetic and virtual libraries, natural products will remain an invaluable source of new pharmaceuticals in the future.

Received: December 12, 2005

- [1] D.J. Newmann, G.M. Cragg, K.M., Snader, *J. Nat. Prod.* **2003**, 66, 1022.
- [2] A. Harvey, *Drug Discov. Today* **2000**, 5, 294.
- [3] M.S. Butler, *J. Nat. Prod.* **2004**, 67, 2141.
- [4] G. Stix, *Scientif. Amer.* **2005**, 292, 88.
- [5] T. Henkel, R.M. Brunne, H. Müller, F. Reichel, *Angew. Chem. Int. Ed.* **1999**, 38, 643.
- [6] M. Feher, J.M. Schmidt, *J. Chem. Inf. Comput. Sci.* **2003**, 43, 218.
- [7] M. Hamburger, K. Hostettmann, *Phytochemistry* **1991**, 30, 3864.
- [8] O. Potterat, M. Hamburger, *Curr. Org. Chem.*, in press.
- [9] D.A. van Elswijk, H. Irth, *Phytochem. Rev.* **2003**, 1, 427.
- [10] O. Potterat, K. Wagner, G. Gemmecker, J. Mack, C. Puder, R. Vettermann, R. Streicher, *J. Nat. Prod.* **2004**, 67, 1528.
- [11] J.-L. Wolfender, K. Ndjoko, K. Hostettmann, *J. Chromatogr. A* **2003**, 1000, 437.
- [12] J.-L. Wolfender, P. Waridel, K. Ndjoko, K.R. Hobby, H.J. Major, K. Hostettmann, *Analisis* **2000**, 28, 895.
- [13] O. Potterat, K. Wagner, H. Haag, *J. Chromatogr. A* **2000**, 872, 85.
- [14] C. Puder, K. Wagner, R. Vettermann, R. Hauptmann, O. Potterat, *J. Nat. Prod.* **2005**, 68, 323.