JOURNAL OF NATURAL PRODUCTS

Article

pubs.acs.org/jnp

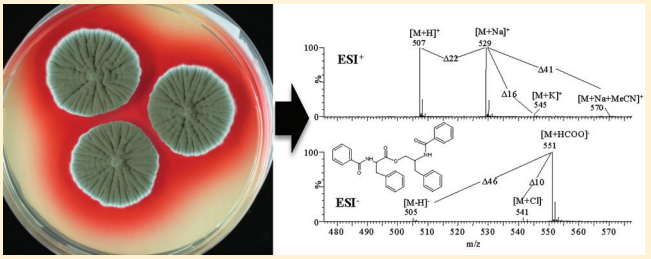
Dereplication of Microbial Natural Products by LC-DAD-TOFMS

Kristian F. Nielsen,* Maria Månsson, Christian Rank, Jens Christian Frisvad, and Thomas O. Larsen

Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800 Kgs. Lyngby, Denmark

Supporting Information

ABSTRACT: Dereplication, the rapid identification of known compounds present in a mixture, is crucial to the fast discovery of novel natural products. Determining the elemental composition of compounds in mixtures and tentatively identifying natural products using MS/MS and UV/vis spectra is becoming easier with advances in analytical equipment and better compound databases. Here we demonstrate the use of LC-UV/vis-MS-based dereplication using data from UV/vis diode array detection and ESI$^+$/ESI$^-$ time-of-flight MS for assignment of 719 microbial natural product and mycotoxin reference standards. ESI$^+$ was the most versatile ionization method, detecting 93% of the compounds, although with 12% ionizing poorly. Using ESI$^+$ alone, 56.1% of the compounds could be unambiguously assigned based on characteristic patterns of multiple adduct ions. Using ESI$^-$, 36.4% of the compounds could have their molecular mass assigned unambiguously using multiple adduct ions, while a further 41% of the compounds were detected only as $[M - H]^-$. The most reliable interpretations of conflicting ESI$^+$ and ESI$^-$ data on a chromatographic peak were from the ionization polarity with the most intense ionization. Poor ionization was most common with small molecules (<200 Da). In ESI$^-$, these were often polar and basic, while in ESI$^+$ they were small aromatic acids or anthraquinones. No single ion-source settings could be applied over a m/z 60−2000 range. However, continuous switching among three settings (e.g., for 0.5 s each) during the chromatographic run allowed MS of both small labile molecules and large peptides, and pseudo MS/MS data on labile molecules since the settings for large molecules often induce fragmentation into small molecules.



Microorganisms and plants are an immense resource of diverse natural products that are candidates for drug development, food and feed additives, and other industrial products.[1−6] Dereplication[7−9] of already known compounds and their potential analogues is a vital part of the discovery process that saves time and resources.[1,10−13] Dereplication processes typically combine chromatographic and spectroscopic methods with database searching. One comprehensive database for natural products from microorganisms is Antibase, which was constructed and is maintained by Laatsch and co-workers.[14]

Our preferred dereplication method for small molecules has gradually shifted from LC-UV/vis diode array detection (DAD) to liquid chromatography combined with mass spectrometry (LC-MS) over the last 10 years, with the LC-UV/vis-MS being our current choice, as the UV/vis spectra still provide important (and inexpensive compared to the MS) information on the presence of conjugated double bonds.[15] This is largely due to the introduction of atmospheric pressure ionization (API) techniques such as ESI and atmospheric pressure chemical ionization (APCI),[9,16−22] which allows a more gentle and versatile compound ionization. Also, dereplication by LC-MS has the major advantage that it can provide accurate mass, which can be used as a query in nearly all NP databases. Dereplication by LC-NMR has advanced substantially in the past few years because of developments in nanoprobes and cryoprobes.[23,24] However, LC-MS is still more sensitive, providing reliable results within the nanogram range.[10,23,25]

MS instruments are continuously being improved, and both positive and negative ionization spectra can now be obtained even during fast ultra-high-performance liquid chromatography (UHPLC). Quadrupole TOF (QTOF) and ion-trap instruments can perform data-dependent MS/MS on all major peaks without predefinition of MS/MS parent masses. This decreases the number of samples that need to be reanalyzed in MS/MS experiments. Moreover, TOF, orbitrap, and Fourier transform ion cyclotron resonance (FT-ICR) instruments provide mass accuracies in the 0.5−1 ppm range, leading to only one or very few possible elemental compositions for a given ion.[9,16−22] However, the most accurate mass instruments currently cannot perform positive/negative switching fast enough for the ultranarrow peaks obtained by UHPLC.

Accurate mass measurements in the dereplication of unknown compounds reduce the number of predicted elemental compositions. This ensures that database searches are conducted with the fewest possible candidates.[26] Recently, Kind et al.[27] demonstrated that compounds within a 0.5−5 ppm mass range often have very different numbers of carbon atoms. Thus, mass accuracy and isotope ratio are almost orthogonal

Received: March 22, 2011
Published: October 25, 2011

parameters that can be used to exclude hypothetical elemental compositions. This has led to mass accuracy and isotope model-fitting being integrated into a single weighted parameter that is part of the software packages of most of the major MS vendors.

For compound classes that are biosynthesized from units with a common elemental composition such as peptides, terpenoids, rhamnolipids, or polysaccharides, accurate mass compared to nominal mass determination provides limited extra information. In these cases, MS/MS, or preferably MS^n, with subsequent fragmentation patterns is more efficient, assuming that reference standards are available for modeling the fragmentation of the compound class.[28,29]

Along with its many clear advantages, LC-MS screening with ESI or APCI ionization faces five major problems: (i) Sensitivity is highly compound-dependent, with some compounds unable to ionize at all in positive and/or negative polarity. This can lead to incorrect assignment of the molecular mass from coeluting impurities, and to ion suppression. (ii) Determining the pattern of adducts such as $[M + H]^+$, $[M + Na]^+$, $[M + NH_4]^+$, $[M - H]^-$, $[M + HCOO]^-$, and others for correct assignment of the molecular ion can be difficult.[24,25,30,31] (iii) Adduct patterns vary substantially from one LC-MS system to another and can even change during a sequence because of sodium extraction from solvent glass bottles.[25,32−35] (iv) Some compounds predominantly form di- and trimeric ions, e.g., $[2M + H]^+$, $[2M + Na]^+$, $[3M + H]^+$, and $[2M - H]^-$, which can further complicate assignment. Finally, (v) some compounds fragment very easily, losing HCOOH, CH_3COOH, one or more H_2O equivalents, and/or CO_2.[36] Any of these problems can result in an erroneous assignment of molecular mass.[25,31]

On the basis of the literature,[25,30,34,35] approximately 20 types of charged adducts and simple fragments are expected from ESI^+ (including $[M + H]^+$ and doubly charged ions), and about 11 different types of adducts and fragments from ESI^- (including $[M - H]^-$ and doubly charged ions), for commonly used buffers and solutions. These studies were performed with a single compound class such as trichothecenes[37] or sugars.[38] To the best of our knowledge, no large survey on adduct formation among different classes of compounds has been reported.

Fragmentation and the adduct formation are highly dependent on the source settings, and especially the potential difference between the skimmers (cones on the Micromass LCT used in this study) has a pronounced effect on ESI spectra.[25] A low potential difference gives low ion transmission of especially high m/z ions, whereas a high potential difference gives better ion transmission but also fragmentation and thus higher abundance of $[M + Na]^+$ and $[M + K]^+$ ions since these are more stable and are not fragmented as easily as $[M + H]^+$. To cope with this problem, some manufactures can ramp the potential between the skimmers during a scan, thus acquiring an "average" mass spectrum over 0.3−2 s (depending on how narrow the chromatographic peaks are). Additionally, some manufactures can acquire fast alternating scans at different settings. We use three shorter separate full scans (scan functions, here 0.5 s) each acquired at 18, 30, and 50 V cone potential differences; this is sometimes referred to as pseudo MS/MS. Comparison of the different spectra from the same chromatographic peak can thus be used to detect labile small m/z ions and high m/z ions ($>m/z$ 700) and in most cases show the $[M + Na]^+$ and $[M + K]^+$ ions as well as reveal losses of H_2O, CH_3COOH, HCOOH, and CO_2 to be used for identification of functional groups.

Dereplication can be an arduous task, especially when a database contains many candidates with similar or identical masses that need to be evaluated. For candidate exclusion, UV/vis[25,39] and MS/MS[33] or pseudo MS/MS data and taxonomic knowledge of the producing organism and its close relatives can be useful.[25] Taxonomy is far more discriminatory for eukaryotes than bacteria, as far less secondary metabolites are found across several fungal genera[15] than across bacteria, due to the lower frequency of horizontal gene transfer in fungi.[40] Candidates can also be differentiated on the basis of the presence of ionizable groups by explorative solid-phase extraction, E-SPE.[41] This approach can be used to separate neutral, basic, acidic, and amphoteric compounds as well as to unmask coeluting compounds. This simplifies the subsequent LC-MS spectra of individual fractions. Other parameters that can be used to eliminate candidates include chromatographic retention. For neutral compounds, the calculated octanol−water partition coefficient $c\log P$ and its pH-corrected log D correlate with retention time.[42−44] This was demonstrated for acylhomoserine lactones, where a convincing $R^2 = 0.993$ was found between retention time (t_R) and log D.[42−44]

In fungal extracts we rarely have observed ion suppression since organic extracts typically contain low amounts of interfering polar compounds. However for quantitative analysis in fungal extracts, recovery experiments are needed to compensate for both losses during extraction and modulation of the electrospray. For the strongly ionizing fumonisins we have observed that spiking extracts from non-fumonisin-producing *Aspergillus niger* strains with fumonisin B_2 (always making mainly $[M + H]^+$) showed peaks 20−50% higher than from pure solutions due to ionization enhancement.[45,46] Real samples may show a slightly altered adduct pattern with more $[M + Na]^+$ and $[M + K]^+$ adduct ions than pure standards due to the presence of Na^+ and K^+ in the sample, and this may even change over a sequence of samples. Note that using MS/MS one would observe such adduct displacement as ion suppression.

The objective of this study was to obtain a detailed understanding of adduct and fragmentation formation in ESI^+ and ESI^- for a database of 719 known compounds and to highlight compound classes not easily detected by ESI^+ and ESI^-. This has not been done on a large scale for natural products, aside from our 2003 study[25] on UV/vis and ESI^+ spectra of 474 compounds. We also evaluated the impact of mass and isotope pattern accuracy on differentiating elemental compositions of 33 136 valid compounds in Antibase2008. Finally, we explored the accuracy of log D calculated by Advanced Chemistry Development (ACD) software for predicting retention times used to eliminate or verify candidates during dereplication.

RESULTS AND DISCUSSION

Antibase and Mass Accuracy. To characterize the general mass distribution of microbial natural products, the accurate mass was calculated for all compounds in AntiBase 2008.[14] Antibase, which is used extensively in our laboratory, contains data for secondary metabolites from microorganisms and higher fungi, including yeasts, ascomycetes, basidiomycetes, lichens, algae, and cyanobacteria.[14] For important fungal genera such as *Aspergillus*, *Penicillium*, *Alternaria*, *Stachybotrys*, and *Fusarium*, AntiBase contains up to 95−98% of the compounds published in journal papers and patents. Before investigating the AntiBase data, the approximately 34 300 records were stripped for (i) compounds with monoisotopic masses below 60 Da; (ii) compounds not containing hydrogen; and (iii) $C_{24}H_{38}N_{12}O_{12}$ (used in the database for unidentified peptides), resulting in 33 136 valid records.

Table 1. Distribution of Compounds in Antibase 2008 with Different Elemental Compositions That Cannot Be Distinguished at Selected Mass Differences[a], Carbon Isotope Ratio, and A+2 Elements (S, Cl, Br) determination

		number of compounds																	
		mass difference (ppm)																	
		nominal[b]			10	5	5	5	5	5	2	2	2	2	2	2	2	2	1
		isotope constraint																	
mass range (Da)	total compds	none	±5C	±2C	±2C[c]	none	none	±5C	±2C	+1C	±1C[c]	none	±5 C	±2 C	+1C	±1C[c]	none		
60–99	133	45	45	38	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100–199	2969	738	717	461	342	13	0	0	0	0	0	0	0	0	0	0	0	0	0
200–299	6167	1532	1376	811	538	104	48	43	40	32	19	0	0	0	0	0	0	0	0
300–399	6494	2119	1694	899	518	264	84	60	53	45	27	0	0	0	0	0	0	0	0
400–499	5125	2208	1636	872	567	434	195	163	138	114	92	0	0	0	0	0	0	0	0
500–599	3764	1791	1237	600	458	326	132	95	71	64	50	58	46	41	36	32	32	C	
600–699	2065	1234	815	392	288	203	106	85	46	39	23	27	23	20	18	17	17	0	
700–799	1590	985	669	352	273	145	69	55	26	25	23	23	20	16	16	16	16	0	
800–899	1157	710	487	275	221	111	64	45	26	22	17	17	11	10	10	10	10	0	
900–999	639	402	252	118	93	47	26	19	7	3	3	5	3	2	2	2	2	0	
1000–3535	3033	1260	659	369	242	118	1	0	0	0	0	0	0	0	0	0	0	0	
total	33136	13024	9587	5187	3571	1775	730	1775	730	344	254	130	103	89	82	77	77	0	

[a]Mass difference, thus meaning that 3–4-fold better instrument mass accuracy is needed to safely distinguish the two. [b]−0.3 to +0.7. [c]Assuming Cl, Br, and S isotope information is being used.

In Table 1, elemental compositions differing by less than 10, 5, 2, and 1 ppm are divided into mass intervals of 100 Da. For example, the mass range 500−599 Da contains 132 compounds with a mass difference of 5 ppm and 58 with a mass difference of 2 ppm. The number of candidates can be reduced by 25−50% when isotope ratios such as the number of carbon atoms (A + 1 corresponding to $^{12}C/^{13}C$ ratio) and halogens or sulfur (A + 2 corresponding to $^{79}Br/^{81}Br$, $^{35}Cl/^{37}Cl$, or $^{32}S/^{34}S$) are used. The AntiBase 2008 distribution confirms the results of Kind et al.[27] on hypothetically generated compositions within the mass range 0−500 Da. One sulfur atom can be difficult to detect because of the low natural abundance of the heavier isotope; nonetheless it is found in 6.9% of the records and is therefore relevant for dereplication purposes. Another way to recognize sulfur-containing compounds is low mass defect (−0.0279, as sulfur has the monoisotopic mass 31.9721 Da). The newest TOF products from Agilent, Waters, Bruker, and Sciex are capable of providing a strong combination of a high mass accuracy (1.5 ppm) and an isotope ratio accuracy of 1−2%.

The observed mass distribution within AntiBase 2008 clearly stresses the importance of high mass accuracy in a high-throughput dereplication setup: if differences of less than 1 ppm mass can be detected, AntiBase 2008 has no overlapping candidates (Table 1), even without information on isotopes. However, in practice, to differentiate between two candidate elemental compositions the mass accuracy of the MS instrument must be 3- to 4-fold better than the mass difference (depending on the statistical way one defines mass accuracy and the "accepted" risk of confusing the two elemental compositions).[30] This level of accuracy can currently be achieved on FT-ICR, Orbitrap, and the newest TOF instruments, if internal mass correction is used.

Figure 1 presents the number of compounds in AntiBase 2008 with the same elemental composition, as a function of the monoisotopic mass. Interestingly, the compounds with the highest redundancy in composition were all terpenes. This makes sense considering the numerous variations in fungi in, for example, nonoxygenated mono- ($C_{10}H_{16}$) and sesquiterpenoids ($C_{15}H_{24}$), with the incorporation of repetitive isoprene units into these compounds. For most, the proper isotope ratio, t_R's, MS/MS data, and/or prerequisite knowledge of microbial origin is imperative for successful dereplication. This is very time-consuming, especially for compositions such as $C_{15}H_{22}O_3$, with 113 candidates in Antibase2008. Thus, *in-silico* tools are

Figure 1. Distribution of AntiBase 2008 compounds (n = 33 136) with identical elemental compositions by monoisotopic mass.

needed for automated fragmentation analysis, especially for compounds with the same elemental composition.[30]

Statistical Analysis on the Adduct Formation and Ionization Efficiency for the 719 Reference Standards. Finding compounds in databases requires unambiguous establishment of the mass or the molecular formula. However, defining the molecular ion from ESI-generated data may be difficult. In addition, knowledge of the adduct pattern is paramount. With a goal of developing a general strategy for interpretation of ESI spectra, we compiled data from all 719 compounds (95% fungal metabolites) in our in-house collection and stored these in an ACD Chemfolder database. Each entry contained chemical structure, chromatographically validated (deconvoluted) monoisotopic ions from ESI^+ and ESI^-, UV/vis spectrum (pH 3−3.3), UV absorption maxima, and retention index (t_R relative to alkylphenones).[47] Analyses were performed by a 15 → 100% MeCN gradient in 20 min on a Luna C_{18} column with 20 mM formic acid in the solvents. The ions detected with relative abundance higher than 5% of the most intense ion from three continuously changing full scan functions with different in-source fragmentation settings (different potential differences between the skimmers/cones) are compiled in Table 1 in the Supporting Information (Excel format). Doubly charged ions are not included in the table because they were rarely observed and their formation is dependent on the specific ESI ion[48] (on a recently acquired Bruker maXis G3 QTOF, we observed doubly charged ions in ESI for most compounds). Paracelsins constituted the only compound class in which double-charged ions were regularly observed, and all were in the 1900−2000 Da range, which is the upper range for secondary metabolites (Figure 1). Likewise, we did not include dimeric $[2M + X]^+$ and trimeric ions $[3M + X]^+$;[30,35] the intensities of these types of ions were not reliably reproducible because their production is concentration-dependent (Figure 2).[48]

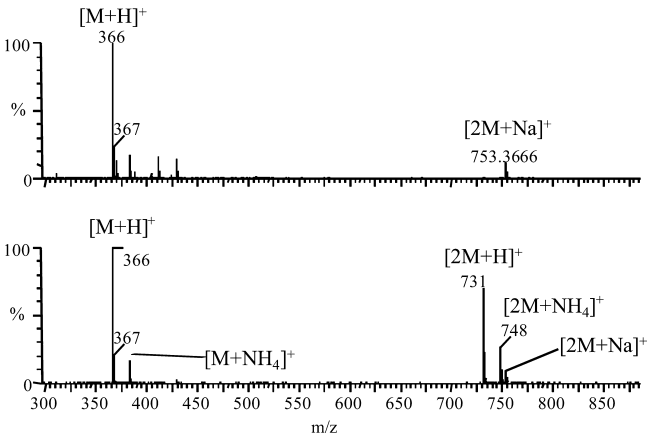


Figure 2. ESI^+ spectrum of brevianamide B from the tail (top) and apex (bottom) of the chromatographic peak. The peak apex has a higher prevalence of dimeric ions.

The most common adducts and fragments observed from the ESI^+ and ESI^- spectral data for the 719 database compounds are shown in Table 2, with their frequencies. Adducts and fragments are

Table 2. Common Mass "Jumps" (Δ) Observed in Electrospray MS and the Frequency of Adducts in 719 Reference Standards

positive electrospray	relative to $[M + H]^+$ Δ (amu)[a]	most intense (%)	observed (%)
$[M + H]^+$		60.2	88.7
$[M + Na]^+$, very stable adduct[b]	21.9820	0.8	21.0
$[M + NH_4]^+$ [c]	17.0265	8.6	20.0
$[M + H + MeCN]^+$, neutral adduct[d]	41.0265	7.8	26.7
$[M + Na + MeCN]^+$, very stable adduct[b,d]	63.0085	0.4	20.2
$[M - H + Na + HCOO]^+$ [g]			
$[M + K]^+$, very stable adduct[b]	37.9559	ND	1.0
$[M + K + MeCN]^+$, very stable adduct[b,d]	78.9824	ND	ND
$[M - H + 2Na]^+$ (acids, phenols, and enoles)	43.9640	ND	0.2
$[M - H + Fe]^+$, isotope m/z 2 lower ca. 4%[e]	53.9193	ND	0.2
Exchanging adducts			
$[M + Na]^+$ to $[M + NH_4]^+$	4.9555	8.9[f]	
$[M + Na]^+$ to $[M + K[^+$	15.9739	ND	
$[M - H_2O + H]^+$ to $[M + Na]^+$	39.9926	6.4	
$[M - H_2O + H]^+$ to $[M + MeCN + Na]^+$[d]	81.0191	7.1[d]	
$[M - H_2O + H]^+$ to $[M + NH_4]^+$	35.0371	7.5[d]	
$[M - 2H_2O + H]^+$ to $[M + Na]^+$	58.0032	1.4	
$[M + NH_4]^+$ to $[M + H - CH_3COOH]^+$[d]	77.0476	2.9[d]	
$[M + Na+MeCN]^+$ to $[M + H- CH_3COOH]^+$[d]	123.0296	2.6[d]	
Neutral losses			
H_2O	18.0106	4.9	33.0
$2H_2O$	36.0212	0.1	8.3
CO_2	43.9898	ND	4.9
CH_3CHO	44.0262	1.0	1.8
Neutral losses			
HCOOH	46.0055	1.0	5.3
CH_3COOH	60.0211	4.0	6.8
NH_3[c]	17.0265	ND	1.5
negative electrospray	relative to $[M - H]^-$ Δ (amu)[a]	most intense (%)	observed (%)
$[M - H]^-$		56.7	69.4
$[M + HCOO]^-$[g]	46.0055	17.2	36.4
$[M + CH_3COO]^-$[h]	60.0211	ND[h]	ND[h]
$[M - 2H + Na]^-$ (acids, phenols, and enoles)	21.9820	ND	4.5
$[M + Cl]^-$ (A + 2 isotope)	35.9767	ND	13.6
$[M - H + H_2O]^-$, e.g., opening of lactones	18.0106	ND	2.6
$[M - H + HCOONa]^-$	67.9874	0.10	9.9
$[M - H + CH_3COONa]^-$[h]	82.0031	ND[h]	ND[h]
Exchanging adducts			
$[M + HCOO]^-$ to $[M + Cl]^-$[g]	10.0288	12.2[g]	
H_2O	18.0106	ND	2.6
CO_2	43.9898	0.3	13.6
CH_3	15.0235	ND	2.6
HCOOH	46.0055	ND	0.6
CH_3COOH	60.0211	ND	2.2

[a]Mass shift. [b]Increased by high in-source fragmentation settings. [c]Loss or addition of NH_3 can usually be determined looking for $[M + Na]^+$. [d]If MeCN is used as solvent. MeCN forms NH_3 upon acidic hydrolysis. [e]Siderophores and artifact from Fe^{2+} liberation in ESI^+. [f]Often seen in peptides >1000 Da. [g]If HCOOH is added as buffer. [h]Not observed since acetate was not added.

dx.doi.org/10.1021/np200254t|J. Nat. Prod. 2011, 74, 2338−2348

recognized on the basis of their relative masses (Δ) compared to $[M + H]^+$ or $[M - H]^-$, creating characteristic patterns of ions. Figure 3 illustrates the annotation of Δ-values in ESI^+ and

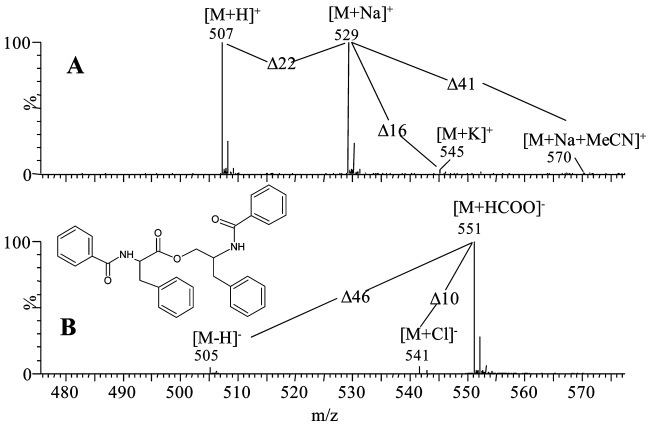


Figure 3. Asperphenamate ESI spectra (low in-source fragmentation): (A) ESI^+ and (B) ESI^- with annotations. The molecular mass of 506 was unambiguously determined in both polarities, e.g., by Δ22 in ESI+ and Δ10 in ESI^-.

ESI^- spectra of asperphenemate to achieve correct mass assignment.

In ESI^+, in 60% of the cases the molecular ion $[M + H]^+$ was the predominant ion in the spectrum, and it was detected from 88.7% of the analytes (Table 2). Of the 719 compounds listed in Supporting Information Table 1, 27.1% produced $[M + H]^+$ as the only pseudomolecular ion. Distinguishing among $[M + H]^+$, $[M + H - H_2O]^+$, $[M + Na]^+$, and $[M + NH_4]^+$ could usually be achieved by careful comparison of the spectra from the different scan functions (with increasing in-source fragmentation settings). Also, most compounds producing $[M + Na]^+$ or $[M + NH_4]^+$ produce both, causing a characteristic five mass unit difference that resulted in <3% risk of incorrectly assigning the monoisotopic mass. The fraction of molecules producing $[M + H]^+$ is significantly lower than for drug molecules,[49,50] since most of the latter contain amines added for enhanced solubility and bioavailability and therefore ionize better than non-amine-containing analogues. The frequency of fungal and bacterial secondary metabolites listed in Antibase2008 containing a basic amine functionality has been shown to be 8% and 28%, respectively.[41]

The acetonitrile (MeCN) adduct was the most frequently formed adduct in ESI^+ when using MeCN–water as the mobile phase. Small molecules (<250 Da) such as pyrones often displayed very prominent $[M + H + MeCN]^+$ ions ($\Delta m/z$ 41), frequently accompanied by a strong $[M + Na + MeCN]^+$ ion ($\Delta m/z$ 63), and in a few cases $[M + H - H_2O + MeCN]^+$ ($\Delta m/z$ 23 to $[M + H]^+$). Interestingly, the same compounds did not ionize well in negative mode. Because of their intense $[M + H + MeCN]^+$, they could easily be mistaken for $[M + H]^+$, and thus special attention should be paid to even low abundance $\Delta m/z$ 41. The $[M + H + MeCN]^+$ and $[M + Na + MeCN]^+$ pair can also be confused with $[M + H]^+$ and $[M + Na]^+$. In most cases, increasing the skimmer potential resulted in a lower abundance of $[M + H + MeCN]^+$, but for a few compounds it increased, possibly because the MeCN attached to $[M + H]^+$ upon collision. In a few cases, fragment ions with MeCN adducts were observed.

An abundant adduct with sodium $[M + Na]^+$ ($\Delta m/z$ 22) was formed by 21.0% of the compounds in ESI^+, but this ion was the most prominent ion in only a few cases. $[M + Na]^+$ was especially abundant for acids where ion exchange on the carboxylic acid group could produce $[M - H + 2Na]^+$. In alkaloids, the abundance of $[M + Na]^+$ was usually very low, presumably due to the low affinity of the lone pair on the nitrogen for Na^+ compared to H^+, and is probably already produced in the eluent.

An ammonium adduct $[M + NH_4]^+$ ($\Delta m/z$ 17) was observed for 20.0% of the compounds. The affinity for NH_4^+ is especially high for oxygenated molecules such as polyketides as well as highly oxygenated terpenes, where this adduct ion is often the most predominant. In contrast, this adduct is extremely rare with alkaloids. For some primary amides and amino acids, a loss of NH_3 was observed, $[M + H - NH_3]^+$, giving rise to the same mass difference of $\Delta m/z$ 17. However, as the $[M + NH_4]^+$ ion is often accompanied by $[M + Na]^+$, this set of ions can be identified by a characteristic mass difference of five mass units. The high abundance of $[M + NH_4]^+$ adducts was surprising since ammonia was not added to any of the solvents. The intensity of the $[M + NH_4]^+$ ion peaks usually increased during a sequence of samples, suggesting that NH_3 was formed from acidic hydrolysis of MeCN. This was confirmed by using pure nonacidified MeCN. Substitution of MeCN with MeOH had an even stronger effect and nearly eradicated the ammonium adducts.

Not surprisingly, loss of water $[M + H - H_2O]^+$ ($\Delta m/z$ 18) was the most common fragment ion formed. Inspection of the structures revealed that almost all lost H_2O was from an alcohol group in which a hydrogen on an α-carbon was available for elimination via double-bond formation.

ESI^+ was able to detect 91.5% (Table 3) of the 719 compounds in Supporting Information Table 1. However, 10.6% of the compounds ionized poorly and might be overlooked at low concentration and/or in complex extracts. Compounds that were unable to ionize in ESI^+ were often (i) anthraquinones (phenols and carbonyls in conjugated systems do not accept a positive charge) and (ii) small acids with the carbonyl bond carbon positioned next to a conjugated system and thus delocalized, whereas non-conjugated carboxylic acids seem to ionize well, probably by accepting a H^+ on the carbonyl lone pair. The 7-fold increase in nonionizing compounds

Table 3. Combined Positive and Negative Ionization Electrospray to Determine Correct Molecular Mass

ionization	ESI^+	ESI^-	both	pairs obsd (%)	most intense obsd			unambiguous[d]	unambig. or only $[M + H]^+/[M - H]^-$
poor (%)	10.6	15.3	1.5		37.8	59–4	ESI^+ alone (%)	56.1	83.2[e]
none (%)	8.5	17.0	2.4	$[M + H]^+$ and $[M-H]^-$	3.3	9.9	ESI^- alone (%)	36.4	77.9[f]
risk of wrong assignent	2.6[a]	7.1[b]	–[c]	$[M + NH_4]^+$ and $[M + HCOO]^-$			ESI^+ and ESI^- (%)	92.9	92.9

[a]2.5% as $[M + NH_4]^+$ (many saved by the m/z Δ5 mass difference of NH_4^+ and Na^+). [b]0.14% as $[M - H]^-$ as $[M + HCOO]^-$. [c]None if ions observed in both polarities. [d]Several adducts point toward same molecular mass. [e]If assuming most intense ion is $[M + H]^+$ in the case of only one ion. [f]If assuming most intense ion is $[M - H]^-$ in the case of only one ion.

compared to our last study[25] was predominantly caused by an increase in these types of small acids in our database. We found that the fraction of poorly ionizing and nonionizing compounds depended on the impurity levels in the solvents and how often the solvent bottles were changed. Besides contamination of the ion source with organic material, we observed that corrosion in the column oven (observed as corrosion at the inlet) introduced ions that caused a 10- to 100-fold drop in sensitivity for poorly ionizing compounds. We speculate that the poorly ionizing compound especially acids may bind to Na^+, Fe^{2+}, Fe^{3+}, and other metal and organic impurities and thus stay in the droplets and/or form many different low-intensity adduct ions. Strongly ionizing compounds such as alkaloids were not significantly affected by these factors, as they mainly accept an H^+. Na^+ efflux from the glass solvent bottles lowers sensitivity and causes ion-pair effects especially visible under HILIC separations (hydrophilic interaction chromatography).[51] This can be avoided by using special plastic bottles for the water (e.g., from Dionex, Sunnyvale, CA, USA), which we are now using on all our LC-MS systems. Solvent and water quality can also have a significant impact on the background. Occasionally the PTFE filters (from several manufactures) used for filtering extracts have been contaminated by polyethylene glycol (PEG) observed as numerous ions with $\Delta m/z$ 44 (CH_2CH_2O) in the 600 to 2000 Da region.

When working with poorly ionizing compounds, we highly recommend trying to purify the extract on small solid-phase extraction columns of an orthogonal chromatographic principle as the analytical HPLC, as, for example, outlined by the E-SPE method.[41] This can both remove coeluting impurities and improve chromatographic peaks of the remaining compounds, facilitating faster and more accurate data interpretation.

On the basis of ESI^+ alone, only 56% of the compounds in our database could have their monoisotopic masses unambiguously assigned (Table 3) via at least two adduct ions with the same molecular mass. In an additional 27% of the compounds, only the $[M + H]^+$ ion was detected. As previously observed,[35] ESI^- produced fewer adducts than ESI^+ (Table 2), and we also observed that in ESI^- 7.1% of the compounds (Table 3) could have their molecular mass incorrectly assigned compared to 2.6% in ESI^+, but none if both ESI^+ and ESI^- were used.

In ESI^- the most commonly formed adduct ion besides $[M - H]^-$ was $[M + HCOO]^-$ ($\Delta m/z$ 46), and whereas the loss of 46 Da was rare (0.5%), an observed Δ46 was almost certain to be from a formate adduct. We hypothesize that in some instances formate delivers the charge by forming $[M + HCOO]^-$, leaving $[M - H]^-$ by evaporation of HCOOH. The $[M + HCOO]^-$ ion can often be confirmed by the presence of $[M + Cl]^-$. A relative mass difference of Δ10 (and a Cl isotope pattern) is therefore characteristic of this pair. In this study, $[M + HCOO]^-$ and $[M + Cl]^-$ were mainly formed from neutral compounds. Therefore, the detection of these adducts was a strong indicator of molecules that did not contain carboxylic acid or acidic phenol groups.

Loss of CO_2 $[M - H - CO_2]^-$ ($\Delta m/z$ 44) was the most common (13.6%) fragment ion formed by ESI^-. Among the carboxylic acids, 37% lost CO_2, which was a 2.5-fold higher prevalence than for the entire data set, but still lower than expected.[36,52] Of the compounds losing CO_2 in negative mode, one-third lost HCOOH in positive mode (5% of the data set). Interestingly, loss of H_2O in negative mode was very rare (Table 2) and observed in only 2.6% of the molecules investigated.

Interestingly, alkali metal ion adducts were also observed in negative mode often as $[M - 2H + Na]^-$ and $[M - H + HCOONa]^-$, which fits well with the Na^+ exchange with a H^+ on the acidic group. This was observed for 4.5% of all compounds investigated, usually for compounds containing a carboxylic acid or another acidic functionality, although these adducts are also reported from phospholipids.[53] Thus, when an unknown compound is being dereplicated, $[M - 2H + Na]^-$ is a strong indication of an acidic functionality. The positive adduct $[M + 2Na - H]^+$ was occasionally observed, but usually at less than 2% abundance compared to the base peak.

Notably, 17 compounds including statins, citrinin, oosporein, and mitorubrins produced $[M - H + H_2O]^-$ in ESI^- and $[M + H]^+$ in ESI^+. We speculate that this is due to equilibrium between (i) open/closed lactones; (ii) anhydrides; or (iii) an internal aldehyde and alcohol forming a hemiacetal. On-column and in-source reactions can be differentiated by looking at the peak width, since on-column reactions produce a broad peak, which was observed for rubratoxins A and B, glauconic acid, byssochlamic acid, and terrestric acid.

For 36.4% of the compounds, the molecular mass could be unambiguously assigned directly from ESI^- (Table 3). In cases where a single ion was observed, it was most likely to be $[M - H]^-$ (41.5%); in only 7.1% of the cases was $[M + HCOO]^-$ misassigned as $[M - H]^-$. This can be clarified by using CH_3COOH in the solvent or looking carefully for a possible $[M + Cl]^-$ ion. The lower fraction of $[M - H]^-$ compared to $[M + H]^+$ is likely due to fundamental differences in the ionization mechanism. In ESI^+, charged species (H^+, Na^+, K^+, NH_4^+) are added to the native molecule due to the high ion strength on the surface originating from HCOOH, Fe^{2+} liberated from the steel capillary, and oxidation of water.[54,55] In negative mode, $[M - H]^-$ is formed through loss of H^+, whereas all other ions in both polarities are formed by addition of a charged species.

One-third of the compounds in our database did not ionize or ionized poorly in ESI^- (Table 3), making ESI^+ the first choice of ionization mode, unless *a priori* knowledge on the target compound(s) suggests the use of ESI^-. We note that a high concentration of formic acid suppresses ionization in ESI^- because of low pH, which makes deprotonation more difficult, and because of competition with formate ions. We obtained better ESI^- sensitivity by reducing the concentration of formic acid; however, to obtain the best possible peak shape and the same RTs between ESI^+ and ESI^- runs we maintained 20 mM formic acid for both ESI^+ and ESI^-.

A comparison of compounds producing $[M + H]^+$ and/or $[M - H]^-$ ions is shown in Table 3. Both ions were observed for 59.4% of the compounds investigated, yet in only 37.8% of the cases did they represent the most abundant ions in both polarities. As mentioned previously, unambiguous assignment of molecular masses was possible for 56.1% of the components tested using ESI^+ and 36.4% using ESI^-. Combining the two ionization modes resulted in 92.9% correctly assigned masses. The remaining 7.1% were poorly/nonionizing compounds or compounds ionizing strongly in only one mode. For example, highly polar alkaloids, such as roquefortines A and B, as well as simple ergot alkaloids ionize strongly as $[M + H]^+$, but are not detected in negative mode. Presumably they are positively charged in the high aqueous droplets formed early in the gradient and thus require elimination of two positive charges to be detected. Poorly ionizing compounds were often small molecules (<200 Da), many having structures that appeared to be volatile (few polar groups and no charge at pH 3.0−3.3).

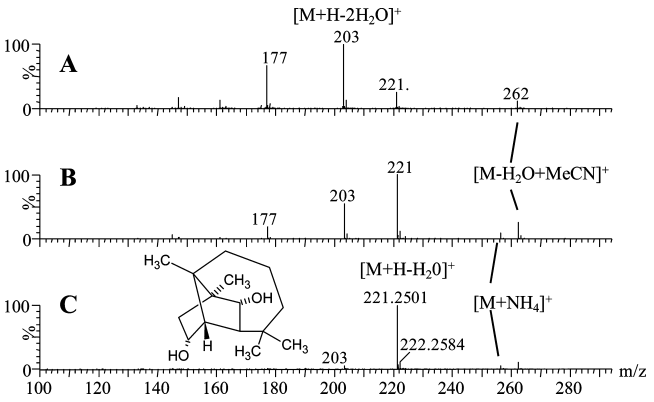


Figure 4. ESI$^+$ spectra of culmorin at cone potential difference of 50 V (A), 30 V (B), and 18 V (C), showing extensive water losses, which hamper the correct assignment unless the $[M + NH_4]^+$ ion is identified and data from ESI$^-$ are used where $[M - HCOO]^-$ is observed.

We propose that these are lost as neutral volatiles because of the high source temperature (350 °C). Presumably, these compounds require APCI ionization to be detected. As expected, ionization efficiency increased during the gradient run, because of the increasing concentration of MeCN, which increased the volatility of the eluent going into the ESI. This phenomenon has been known since the early days of ESI.[56,57] The sensitivity of poorly ionizing polar compounds can therefore be enhanced by using columns improved for polar analytes, such as pentafluorophenyl or HILIC columns.[58,59] Perhaps the most important observation from comparing ESI$^+$ and ESI$^-$ was that the polarity providing the most intense ionization always resulted in the most simple and easily interpreted spectrum and the most reliable adduct pattern.

In-Source Fragmentation versus Ion Transmission. In the ESI source, efficient transport of ions is achieved by a high potential difference between the skimmers. However, for labile compounds, this leads to fragmentation and loss of H_2O, CO_2, HCOOH, CH_3COOH, and other compounds, which can lead to incorrect molecular mass assignment. Therefore, we recommend a relatively labile compound for ESI source tuning (instead of the very stable reserpine, which is used by many manufacturers) especially for adjustment of the potential between the skimmers (down to 12−20 V). However, tuning the source on a small labile molecule (200−300 Da) leads to very poor sensitivity for large m/z ions (>m/z 800). This originates, in part, from the higher velocity (skimmer potential difference ~$1/2 \times m \times V^2$) of low m/z ions between the skimmers (cones). An m/z 60 ion has nearly six times the velocity between the skimmers than an m/z 2000 ion. To obtain a broad m/z range of ions for transfer from the ESI source to the MS, full scans at two or three alternating skimmer potentials need to be used (or one scan where the potential is ramped from low to high). Furthermore the scan time needs to reflect the chromatographic peak width, so one gets enough scans over a peak (eight scans over a peak for nonquantitative purposes). Thus, a 12 s peak width requires a scan time of 0.5 s if three scans at three different cone potentials are desired. If the chromatographic peaks are narrower, the number of different cone potentials or the scan time should be reduced. On the LCT used we have found that scan times below 0.5 s yield an unsatisfactory mass accuracy (≫5 ppm).

Another advantage of alternating cone potentials is that it enhances fragment ions and adducts with stable alkali metal ions.[31,60] In addition, it allows collection of pseudo MS/MS data,[61] as seen in Figure 4. This concept is now marketed as MS^E in QTOF instruments from one manufacture (fragmentation moved to the collision cell).[62]

Chromatography. The inherent nature of reversed-phase chromatography makes it tempting to link retention and log D, which has been shown to correlate well for acylhomoserine lactones (R^2 = 0.993).[63] As shown in Figure 5A, a similar correlation was obtained for the fumonisins, which contain five ionizable groups. However, for the neutral aflatoxins and sterigmatocystins (Figure 5B) the correlation was less convincing (R^2 = 0.65) with an average predicted error of 55 and a max error of 95 for the retention index. Figure 6 shows that linking retention and log D for all 719 compounds in our database gave an unsatisfactory correlation (R^2 = 0.38). However, when several groups of compounds were analyzed individually, compounds

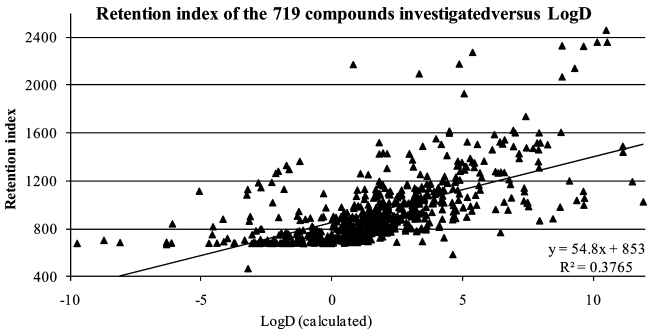


Figure 6. Plot of the retention index (relative to alkylphenones) of 719 secondary metabolites (collected over two years), with calculated log D at the pH (3.0) used in the chromatographic run.

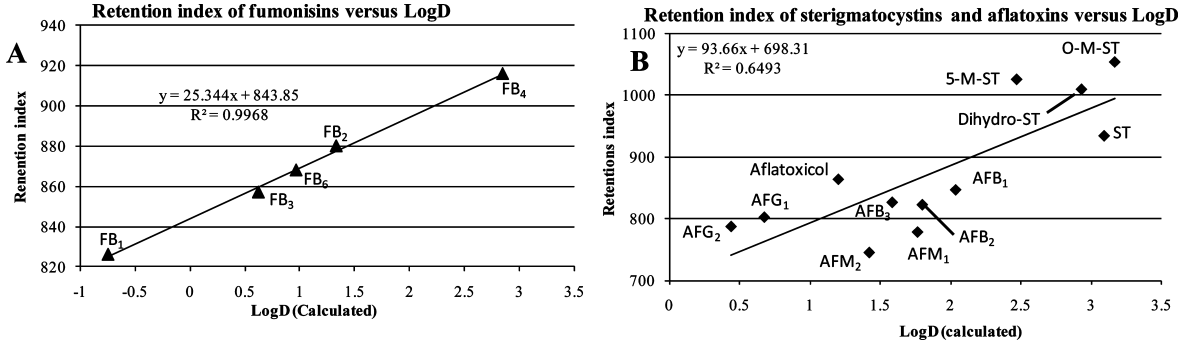


Figure 5. Correlation between calculated log D (pH 3.0) and retention index for two groups of mycotoxins: (A) fumonisins (FB); (B) sterigmatocystins (ST) and aflatoxins (AF).

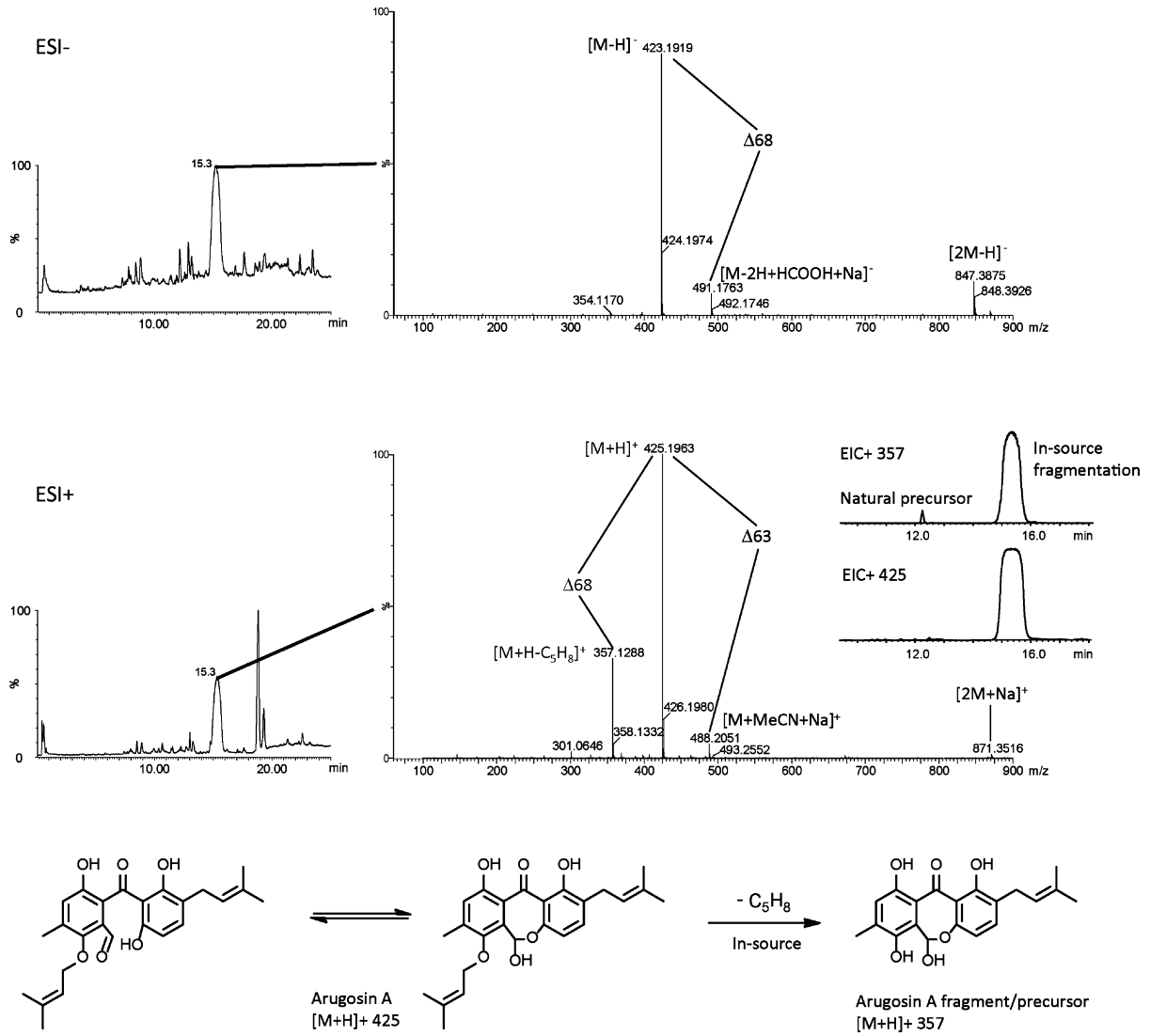


Figure 7. Dereplication of arugosin A in an *Aspergillus nidulans* extract, showing the extremely broad peak and loss of *o*-prenyl, as well as a nonprenylated precursor (arugosin H).

such as acylhomoserine lactones correlated well (R^2 = 0.99) under our conditions. In general, compounds with log *D* below −1 were not retained, so a lower log *D* did not impact the retention index. This shows that calculated log *D* values are of limited use for exclusion of candidates in dereplication unless sufficiently related compounds are available as reference standards to verify the regression model. The generally poor correlation seen in Figure 6 is likely due to factors such as inaccurate pK_a calculations, ionic interactions with the underlying silica, steric effects of molecules, and ion-pairing effects with $HCOO^-$ and H_3O^+.[64−68]

Examples of Dereplication. *Aspergillus nidulans* is an important filamentous fungus known to produce a wide range of secondary metabolites. One such compound is arugosin A, also known from *Emericella rugulosa*,[69] which is a major component in *A. nidulans* crude extracts (Figure 7). This compound ionizes well in both ionization modes, with several adducts verifying the molecular ion ($[M + H]^+$ *m/z* 425.1963, calcd 425.1964). Even at low in-source fragmentation settings (25 V), the fragment ion representing $[M + H - prenyl]^+$ (*m/z* 357) occurs at 30% of the intensity of the $[M + H]^+$ ion. This characteristic fragment can be used to selectively exclude compounds from an AntiBase search, as only four of 34 potential (±0.05 Da) candidates have a labile prenyl group. The four possible candidates, arugosin A, B, and C and variecoxanthone C, have the molecular formula $C_{25}H_{28}O_6$, and all have been isolated from *Aspergillus*. The three arugosins have the same chromophore, which can be matched with the obtained UV data.[69] Also, the appearance of the very broad chromatographic peak is consistent with the structure of the arugosins and coincides with their ability to form a hemiacetal (pronounced at pH 3). The three arugosin isomers cannot be distinguished by LC, MS, or UV characteristics, but ultimately NMR is needed for absolute identification of arugosin A.

The antifungal metabolite aspirochlorine was tentatively identified from an extract of *Aspergillus oryzae* (Figure 8). This compound did not ionize at all in ESI^+, so determination of the molecular mass required negative mode data. $[M - H]^-$ (*m/z* 358.9601, calcd 358.9563) was recognized as the base peak supported by a strong $[M - H - HCOONa]^+$ ion (Δ *m/z* 68) and the $[2M - 2H + Na]^+$ dimeric adduct, in agreement with the presence of an acidic phenol group. Interfering ions from coeluting compounds were deconvoluted based on the high mass defect. Of seven possible candidates in AntiBase, five

dx.doi.org/10.1021/np200254t|J. Nat. Prod. 2011, 74, 2338−2348

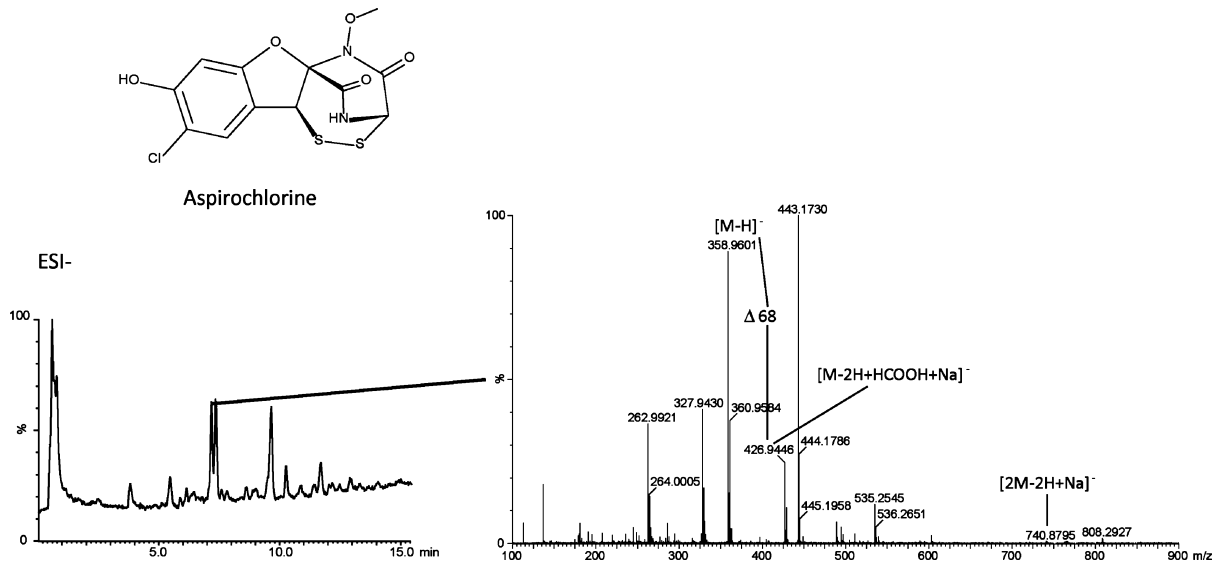


Figure 8. Dereplication of aspirochlorine in an *Aspergillus oryzae* extract in which the compound ionizes only in negative mode and produces many adducts. The mass accuracy is slightly off (10 ppm) due to peak overloading, which is a known problem on this instrument type.[25].

could be excluded based on isotope pattern. The remaining candidates, aspirochlorine and A-30641, both have the molecular formula $C_{12}H_9ClN_2O_5S_2$, consistent with the observed A + 2 ion intensity. Characteristic fragment ions (25 V) at *m/z* 327 ($[M - H-CH_3O]^-$) and *m/z* 263 ($[M - H - CH_3O - 2S]^-$) matched the structural moieties of both compounds. Again NMR was used to prove the identity of the peak as aspirochlorine (data not shown).

CONCLUSION

Thorough investigation of adduct formation from 719 natural products demonstrated that ESI^+ was the most versatile detection method for 93% of the compounds. Unambiguous assignment using adduct patterns was possible for 56% of compounds by ESI^+ alone and 37% by ESI^- alone. Of the 719 compounds, 41% produced $[M - H]^-$ with no validating ions. No perfect ion-source settings were found to cover the full range of compounds with *m/z* 60−2000. However, ultrafast alternation between two or three settings during the chromatographic run allowed both small labile molecules and large peptides into the MS and provided pseudo MS/MS data. When combining ESI^+ and ESI^- data, the ionization polarity with the most intense ionization always provided both the simplest spectrum and the most reliable adduct pattern for interpretation.

EXPERIMENTAL SECTION

General Experimental Procedures. Solvents were HPLC grade, and all other chemicals were analytical grade. All were from Sigma-Aldrich (Steinheim, Germany) unless stated otherwise. The MeCN used in the LC gradient system was Sigma catalogue #5485, and the formic acid used was Fluka catalogue #56302 (for LC-MS). Water was purified using a Milli-Q system (Millipore, Bedford, MA, USA). Alkylphenone standards for retention index determination were diluted in MeOH to 2−3 mM as previously described.[25,47] Extracts of *Aspergillus nidulans* and *A. oryzae* were available from a previous study.[70]

Theoretical pK_a calculations for compounds in Antibase2008[14,24] were made by converting the ChemFinder versions to ACD ChemFolder (Advanced Chemistry Development, ACD, Toronto, Canada) then batch-calculating the log *D* values at pH 2.7, 3.0, and 3.3 using the ACD 2008 PhysChem suite[41] by Dr. Shahriar Jahanbakht (Chemacad, Obernai, France).

Reference Standards. Metabolite standards have been collected over the last 30 years,[25,47] either from commercial sources, as gifts from other research groups, or from our own projects; they are therefore available only in micro- to milligram quantities, and some are only about 50% pure. About 95% of the compounds are fungal metabolites, while 5% are bacterial metabolites. About one-third of the standards were purchased from Sigma-Aldrich, Axxora (Bingham, UK), Cayman (Ann Arbor, MI, USA), TebuBio (Le-Perray-en-Yvelines, France), Biopure (Tulln, Austria), Calbiochem, (San Diego, CA, USA), and ICN (Irvine, CA, USA). Generally, reference standards were powders transferred to an autosampler vial using a Pasteur pipet (1−2 mm of powder in the tip), and, unless stated otherwise, 1.5 mL of MeCN was added to the vial. For highly polar substances not soluble in MeCN, a few drops of water were added or 2-propanol was used. Standard solutions were kept at −20 °C.

LC-DAD-TOFMS. Tuning and calibration in ESI^+ was performed as previously described.[25] In negative mode, the MS was tuned to a resolution of approximately 5000 FWHM for $[M - H]^-$ of leucine enkephaline and calibrated using a solution of PEG-di-acids (average MW 200 and 600), polyalanine, 1,6-dihydroxybenzoic acid, and chloramphenicol, diluted in MeCN−water (1:1 v/v).

The lock spray was connected to a custom-made system equipped with a 0.5 L bottle to which a positive pressure of 2−3 bar N_2 was applied for a constant flow of approximately 20 μL/min leucine enkephaline solution (0.1 μg/mL in MeCN−H_2O−HCOOH at 50:50:0.1). The $[M + H]^+$ and $[M - H]^-$ ions were used as lock mass ions in ESI^+ and ESI^-, respectively.

All analyses were performed on an Agilent 1100 LC system with a DAD coupled to a LCT orthogonal TOF mass spectrometer (Micromass, Manchester, UK), with a Z-spray electrospray source and a LockSpray probe. The system was controlled with MassLynx 4.0 software. Separation was at 40 °C on a 50 × 2 mm i.d., 3 μm, Luna C_{18} II column (Phenomenex, Torrance, CA, USA) equipped with a Security Guard precolumn. A linear water−MeCN gradient of 15% MeCN−water to 100% MeCN in 20 min was applied at a constant flow of 0.3 mL/min; then 100% MeCN was maintained for 5 min before returning to the starting conditions in 2 min and equilibrating for 5 min. Formic acid was added to all solvents to 20 mM. UV spectra were collected by a DAD every 0.4 s from 200 to 700 nm with a resolution of 2 nm.

The source was kept at 120 °C, and the desolvation temperature was 400 °C. The sample cone was continuously flipped among 18, 30, and 50 V (scan functions 1, 2, and 3, respectively), for 0.5 s each, with scan ranges of *m/z* 60−2000 for the last. The lock mass scan was performed every 6 s at a sample cone of 30 V. The capillary was held a

3000 V for ESI^+ and 2000 V for ESI^-, and the desolvation flow held at 400−450 L/h (ESI^+) or 675−725 L/h (ESI^-).

Data Analysis. Data from 719 metabolites are in Supporting Information Table 1, which lists the metabolite, formula, and retention index calculated as described by Frisvad and Thrane.[47]

ESI^+ and ESI^- data were separated by (i) how well the compound ionized; (ii) ions that represented M adducts (H^+, NH_4^+, Na^+, K^+, Cl^-, MeCN, H_2O, and formate, individually or in combination), or simple losses (H_2O, NH_3, CH_3, CH_3COOH, HCOOH, CO_2, CH_3CHO); or (iii) significant diagnostic fragments (usually from scan function 2 or 3). Isotope data included Cl loss only for significant fragments. All ions of the individual components were confirmed by manual mass deconvolution. Adduct data were obtained from the peak apex or averaged over the peak, while accurate mass measurements were calculated from spectra obtained from the front or tail of the peak. This reduced the influence of detector dead time by keeping ion counts below 1000 counts/s (overloading of the detector on the LCT can lead to a mass error of up to m/z 0.05).

Data Management. MS and UV spectra data were stored in a custom-made database made in ACD v. 12 Chemfolder. The structures of most compounds were imported from Antibase 2007 or 2008.[14,24] Table 1 in the Supporting Information is an Excel version of the database. Log D at pH 2.7, 3.0, and 3.3 was batch-calculated and added to the database. UV data were subjected to background subtraction and confirmed to be not saturated.

■ ASSOCIATED CONTENT

Supporting Information

Supporting Information Table 1 with data on the 719 compounds is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +45 45252602. E-mail: kfn@bio.dtu.dk.

■ ACKNOWLEDGMENTS

We thank Dr. S. Jahanbakht (ACD Laboratories, Obernai, France) for assistance with pK_a and log D calculations. Funding was from the Programme Committee for Food, Health, and Welfare under the Danish Strategic Research Council, Danish Research Council for Technology and Production Sciences (247-07-0513), and the EEC project MycoRed (KBBE-2007-222690-2). The authors are grateful to the two anonymous reviewers for their comments and recommendations.

■ REFERENCES

(1) Zengler, K.; Paradkar, A.; Keller, M. New Methods to Access Microbial Diversity for Small Molecule Discovery. In *Natural Products: Drug Discovery and Therapeutic Medicine*; Zhang, L., Demain, A. L., Eds.; Humana Press Inc.: Totowa, 2009; pp 275−293.
(2) Jensen, P. A.; Mincer, T. J.; Williams, P. G.; Fenical, W. *Ant. van Leeuwenhoek* **2005**, *87*, 43−48.
(3) Firn, R. D.; Jones, C. G. *Nat. Prod. Rep.* **2003**, *20*, 382−391.
(4) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. *ChemBioChem* **2002**, *3*, 619−627.
(5) Bills, G. F. *Can. J. Bot.* **1995**, *73*, s33−s41.
(6) Butler, M. S. *J. Nat. Prod.* **2004**, *67*, 2141−2153.
(7) Bitzer, J.; Kopcke, B.; Stadler, M.; Heilwig, V.; Ju, Y. M.; Seip, S.; Henkel, T. *Chimia* **2007**, *61*, 332−338.
(8) Bobzin, S. C.; Yang, S.; Kasten, T. P. *J. Ind. Microbiol. Biotechnol.* **2000**, *25*, 342−345.
(9) Cordell, G. A.; Shin, Y. G. *Pure Appl. Chem.* **1999**, *71*, 1089−1094.
(10) Feng, X.; Siegel, M. M. *Anal. Bioanal. Chem.* **2007**, *389*, 1341−1363.
(11) Dinan, L. Dereplication and Partial Identification of Compounds. In *Methods in Biotechnology, Vol. 20, Natural Products Isolation*, 2nd ed.; Humana Press: New York, 2005; pp 297−321.
(12) Zhang, L. Integrated Approaches for Discovering Novel Drugs From Microbial Natural Products. In *Natural Products: Drug Discovery and Therapeutic Medicine*; Zhang, L., Demain, A. L., Eds.; 2005; pp 33−55.
(13) Stadler, M.; Ticky, H.-V.; Katsiou, E.; Hellwig, V. *Mycol. Progress* **2009**, *2*, 95−122.
(14) Laatsch, H. AntiBase 2010; Wiley-VCH: Weinheim, Germany; http://www.wiley-vch.de/stmdata/antibase2010.php. 2010.
(15) Larsen, T. O.; Smedsgaard, J.; Nielsen, K. F.; Hansen, M. E.; Frisvad, J. C. *Nat. Prod. Rep.* **2005**, *22*, 672−695.
(16) Julian, R. K. Jr.; Higgs, R. E.; Gygi, J. D.; Hilton, M. D. *Anal. Chem.* **1998**, *70*, 3249−3254.
(17) Strege, M. A. *Anal. Chem.* **1998**, *70*, 2439−2445.
(18) Constant, H. L.; Beecher, C. W. W. *Nat. Prod. Lett.* **1995**, *6*, 193−196.
(19) Corley, D. G.; Durley, R. C. *J. Nat. Prod.* **1994**, *57*, 1484−1490.
(20) Shingematsu, N. *J. Mass Spectrom. Soc. Jpn.* **1997**, *45*, 295−300.
(21) Waridel, P.; Ndjoko, K.; Hobby, K. R.; Major, H. J.; Hostettmann, K. *Analusis* **2000**, *28*, 895.
(22) Eldridge, G. R.; Vervoort, H. C.; Lee, C. M.; Cremin, P. A.; Williams, C. T.; Hart, S. M.; Goering, M. G.; O'Neil-Johnson, M.; Zeng, L. *Anal. Chem.* **2002**, *74*, 3963−3971.
(23) Molinari, G. *Adv. Exp. Med. Biol.* **2009**, *655*, 13−27.
(24) 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.; Munro, M. H. G. *J. Nat. Prod.* **2008**, *71*, 1595−1599.
(25) Nielsen, K. F.; Smedsgaard, J. *J. Chromatogr. A* **2003**, *1002*, 111−136.
(26) Nielsen, K. F.; Smedsgaard, J.; Larsen, T. O.; Lund, F.; Thrane, U.; Frisvad, J. C. Chemical Identification of Fungi: Metabolite Profiling and Metabolomics. In *Fungal Biotechnology in Agricultural, Food, and Environmental Applications*, Arora, D. K., Ed.; Marcel Dekker: New York, 2004; pp 19−35.
(27) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2006**, 7.
(28) Sharma, A.; Jansen, R.; Nimtz, M.; Johri, B. N.; Wray, V. *J. Nat. Prod.* **2007**, *70*, 941−947.
(29) Bartok, T.; Tolgyesi, L.; Szekeres, A.; Varga, M.; Bartha, R.; Szecsi, A.; Bartok, M.; Mesterhazy, A. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 35−42.
(30) Huang, N.; Siegel, M. M.; Kruppa, G. H.; Laukien, F. H. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1166−1173.
(31) Nielsen, K. F.; Graefenhan, T.; Zafari, D.; Thrane, U. *J. Agric. Food Chem.* **2005**, *53*, 8190−8196.
(32) Schug, K.; Mcnair, H. M. *J. Chromatogr. A* **2003**, *985*, 531−539.
(33) Fredenhagen, A.; Derrien, C.; Gassmann, E. *J. Nat. Prod.* **2005**, *68*, 385−391.
(34) Gorlach, E.; Richmond, R. *Anal. Chem.* **1999**, *71*, 5557−5562.
(35) Overy, D. P.; Enot, D. P.; Tailliart, K.; Jenkins, H.; Parker, D.; Beckmann, M.; Draper, J. *Nat. Protoc.* **2008**, *3*, 471−485.
(36) Levsel, K.; Schiebel, H. M.; Terlouw, J. K.; Jobst, K. J.; Elend, M.; Preib, A.; Thiele, H.; Ingendoh, A. *J. Mass Spectrom.* **2007**, *42*, 1024−1044.
(37) Sulyok, M.; Berthiller, F.; Krska, R.; Schuhmacher, R. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2649−2659.
(38) Wan, E. C. H.; Yu, J. Z. *Env. Sci. Tox.* **2007**, *41*, 2459−2466.
(39) Larsen, T. O.; Petersen, B. O.; Duus, J. O.; Sørensen, D.; Frisvad, J. C.; Hansen, M. E. *J. Nat. Prod.* **2005**, *68*, 871−874.
(40) Wietz, M.; Månsson, M.; Godtfredsen, C. H.; Larsen, T. O.; Gram, L. *Marine Drugs* **2010**, *8*, 2946−2960.
(41) Månsson, M.; Phipps, R. K.; Gram, L.; Munro, M. H.; Larsen, T. O.; Nielsen, K. F. *J. Nat. Prod.* **2010**, *73*, 1126−1132.
(42) Valko, K.; Bevan, C.; Reynolds, D. *Anal. Chem.* **1997**, *69*, 2022−2029.
(43) Valko, K.; Slegel, P. *J. Chromatogr.* **1993**, *631*, 49−61.
(44) Valko, K.; Snyder, L. R.; Glajch, J. L. *J. Chromatogr. A* **1993**, *656*, 501−520.

(45) Mogensen, J. M.; Frisvad, J. C.; Thrane, U.; Nielsen, K. F. *J. Agric. Food Chem.* **2010**, *58*, 954−958.
(46) Nielsen, K. F.; Mogensen, J. M.; Johansen, M.; Larsen, T. O.; Frisvad, J. C. *Anal. Bioanal. Chem.* **2009**, *395*, 1225−1242.
(47) Frisvad, J. C.; Thrane, U. *J. Chromatogr.* **1987**, *404*, 195−214.
(48) Ding, J. M.; Anderegg, R. J. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 159−164.
(49) Smyth, W. F. *Anal. Chim. Acta* **2003**, *492*, 1−16.
(50) Colombo, M.; Sirtori, F. R.; Rizzo, V. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 511−517.
(51) Matyska, M. T.; Pesek, J.; Fischer, S. M.; Sana, T. R. Method Development Strategies for the Analysis of Hydrophilic Metabolites using a Silica Hydride-Based Stationary Phase. Poster 2A-048. Metabolomics2010. Amsterdam.
(52) Holcapek, M.; Jirasko, R.; Lisa, M. *J. Chromatogr. A* **2010**, *1217*, 3908−3921.
(53) Hsu, F. F.; Turk, J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1146−1157.
(54) Cole, R. B. *J. Mass Spectrom.* **2000**, *35*, 763−772.
(55) de la Mora, J. F.; Van Berkel, G. J.; Enke, C. G.; Cole, R. B.; Martinez-Sanchez, M.; Fenn, J. B. *J. Mass Spectrom.* **2000**, *35*, 939−952.
(56) Amad, M.; Cech, N. B.; Jackson, G. S.; Enke, C. G. *J. Mass Spectrom.* **2000**, *35*, 784−789.
(57) Enke, C. G. *Anal. Chem.* **1997**, *69*, 4885−4893.
(58) Ikegami, T.; Tomomatsu, K.; Takubo, H.; Horie, K.; Tanaka, N. *J. Chromatogr. A* **2008**, *1184*, 474−503.
(59) Sørensen, J. L.; Nielsen, K. F.; Thrane, U. *J. Agric. Food Chem.* **2007**, *55*, 9764−9768.
(60) Bruhn, J. B.; Nielsen, K. F.; Hansen, M.; Bresciani, J.; Hjelm, M.; Shulz, S.; Gram, L. *Appl. Environ. Microbiol.* **2005**, *71*, 7263−7270.
(61) Ding, J. M.; Burkhart, W.; Kassel, D. B. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 94−98.
(62) Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1989−1994.
(63) Fekete, A.; Frommberger, M.; Rothballer, M.; Li, X. J.; Englmann, M.; Fekete, J.; Hartmann, A.; Eberl, L.; Schmitt-Kopplin, P. *Anal. Bioanal. Chem.* **2007**, *387*, 455−467.
(64) Snyder, L. R.; Dolan, J. W.; Carr, P. W. *J. Chromatogr. A* **2004**, *1060*, 77−116.
(65) Dolan, J. W.; Snyder, L. R.; Blanc, T. *J. Chromatogr. A* **2000**, *897*, 51−63.
(66) Dolan, J. W.; Snyder, L. R.; Blanc, T.; Van Heukelem, L. *J. Chromatogr. A* **2000**, *897*, 37−50.
(67) Wilson, N. S.; Nelson, M. D.; Dolan, J. W.; Snyder, L. R.; Wolcott, R. G.; Carr, P. W. *J. Chromatogr. A* **2002**, *961*, 171−193.
(68) Snyder, L. R.; Dolan, J. W. *LCGC North Am.* **2002**, *20*, 1016−1026.
(69) Ballanti, J. A.; Francis, D. J.; Hassall, C. H.; Wright, J. L. C. *J. Chem. Soc., Chem. Commun.* **1970**, 1175−1182.
(70) Rank, C.; Nielsen, K. F.; Larsen, T. O.; Varga, J.; Samson, R. A.; Frisvad, J. C. *Fungal Biol.* **2011**, *115*, 406−420.