ELSEVIER

Journal of Chromatography A, 794 (1998) 299–316

JOURNAL OF CHROMATOGRAPHY A

Liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance spectroscopy for the screening of plant constituents

J.-L. Wolfender, S. Rodriguez, K. Hostettmann*

Institut de Pharmacognosie et Phytochimie, Université de Lausanne, B.E.P., CH-1015 Lausanne, Switzerland

Abstract

Rapid detection of biologically active natural products plays a strategical role in the phytochemical investigation of crude plant extracts. In order to perform an efficient screening of the extracts, both biological assays and HPLC analysis with various detection methods are used. Hyphenated techniques such as HPLC coupled to UV photodiode array detection (LC–DAD-UV) and to mass spectrometry (LC–MS or LC–MS–MS) provide on-line numerous useful structural information on the metabolites prior to isolation. The recent introduction of HPLC coupled to nuclear magnetic resonance (LC–NMR) represents a powerful complement to the LC–UV–MS screening. Various plants belonging to the Gentianaceae and Leguminosae families have been analysed by LC–UV, LC–MS, LC–MS–MS and LC–NMR. The use of all these hyphenated techniques allows the rapid structural determination of known plant constituents with only a minute amount of plant material. With such an approach, the time-consuming isolation of common natural products is avoided and an efficient targeted isolation of compounds presenting interesting spectroscopical or biological features can be performed. In this paper several representative applications of this combined approach from our laboratory are discussed. © 1998 Elsevier Science B.V.

Keywords: Mass spectrometry; Nuclear magnetic resonance spectroscopy; *Gentianaceae*; *Leguminosae*; Polyphenols; Triterpene glycosides; Xanthones

1. Introduction

Efficient detection and rapid characterisation of natural products play an important role as an analytical support in the work of natural products chemists. Furthermore, the achievement of structural information on the constituents of an extract at the earliest stage of separation is a strategic element for guiding an efficient and selective isolation procedure.

A crude plant extract is a very complex mixture containing sometimes hundreds or thousands of different metabolites [1]. The chemical nature of these constituents differs considerably within a given extract and the variability of the physicochemical as well as the spectroscopic parameters of these compounds cause numerous detection problems. Although different types of LC detectors exist, such as UV, RI, fluorescence, electrochemical, evaporative laser light scattering, etc., none permits the detection of all the secondary metabolites encountered in a plant extract within a single analysis, each method having its own specificity. For example, a product having no important chromophore cannot be detected by a UV monitor. As every natural compound possesses a given molecular weight, detection of all these compounds using mass spectrometry (MS) can

*Corresponding author.

0021-9673/98/$19.00
PII S0021-9673(97)00939-4

be ideally considered as a universal detection technique [2]. At present, MS is the most sensitive method of molecular analysis and has the potential to yield information on the molecular weight as well as on the structure of the analytes. Furthermore, due to its high power of mass separation, very good selectivities can be obtained.

The main problem of the use of LC–MS in natural product chemistry resides in the ionisation of the very important variety of compounds found in a crude plant extract. Indeed, if many LC–MS interfaces exist on the market, none of them allows a real universal detection of all the constituents of a plant extract. Each of these interfaces has its own characteristics and range of applications, and several of them are suitable for the analysis of plant secondary metabolites [3]. In our approach to LC–MS, mainly used for the HPLC screening of crude plant extracts, three interfaces, thermospray (TSP) [4], continuous flow-fast atom bombardment (CF-FAB) [5] and electrospray (ESP) [6] have been investigated [3]. They cover the ionisation of relatively small non-polar products (aglycones, 200 u) to high polar molecules (glycosides, 2000 u). LC–TSP-MS allowed a satisfactory ionisation of moderately polar constituents such as polyphenols or terpenoids in the mass range of 200–800 u. For larger polar molecules such as saponins (MW>800 u), CF-FAB or ESP are the methods of choice [3]. In our laboratory, thermospray is the most widely used interface.

When screening extracts for unknowns, HPLC coupled with UV photodiode array detection (LC–UV) is an important complement to LC–MS, providing useful information on the class of natural products of interest. LC–UV has been used since more than a decade by phytochemists for screening extracts [7] and is now widespread in many laboratories. Recently LC–NMR [8] has been introduced as another powerful complementary technique for on-line structural identification. This technique which is much less sensitive than LC–MS has nevertheless shown to be of strategic importance for structural investigation of natural products.

The aim of this paper is to show how LC–MS and the other related hyphenated techniques such has LC–UV or LC–NMR have been used to solve phytochemical problems. Several analyses of plant constituents from our laboratory have been selected to cover various aspects of this analytical approach. While applications to the chemical screening of crude plant extracts are described, no attempt has been made here to provide an exhaustive review of the domain.

2. Experimental

2.1. Chemicals

HPLC-grade water was prepared by distillation on a Buchi Fontavapor 210 distillation instrument (Flawil, Switzerland) and passed through a 0.50-μm filter Millipore (Bedford, MA, USA). HPLC-grade MeCN and MeOH from Maechler AG (Reinach, Basel, Switzerland) were passed through a filter of 0.45 μm. D_2O (>99%) for LC–NMR experiments was purchased from Dr. Glaser AG (Basel, Switzerland). Ammonium acetate and trifluoroacetic acid (TFA) were obtained from Merck (Darmstadt, Germany) and glycerol from Fluka (Buchs, Switzerland).

2.2. Samples

Extracts were prepared from the dried plant material by maceration at room temperature during 24 h, with successively CH_2Cl_2 and MeOH. Solutions to be analysed were usually prepared by dissolving 30–100 mg of extract in 1 ml of a MeOH–H_2O mixture. The injected volumes varied from 10 to 20 μl.

2.3. HPLC conditions

All the separations were performed on a Nova-Pak C_{18} column, 4 μm (150×3.9 mm I.D.) from Waters (Bedford), equipped with a Nova-Pak Guard-Pak C_{18} precolumn. For the analysis of the Gentianaceae extracts, *Chironia krebsii*, *Halenia corniculata*, *Gentiana ottonis* and *Swertia calycina*, an MeCN–H_2O gradient (5:95→65:35 in 50 min; 1 ml/min) was used. For the separation of the saponins from *Swartzia madagascariensis* and *Phytolacca dodecandra* an MeCN–H_2O gradient (30:70→50:50 in 30 min; 0.9 ml/min) was used. In all cases, to avoid the tailing of phenolic compounds, 0.05% of trifluoro-

acetic acid was added to the solvents, giving a pH of 3. The eluent delivery was provided by an HPLC 600-MS pump (Waters-Millipore, Bedford, MA, USA) equipped with a gradient controller.

2.4. LC–UV (DAD) analyses

The UV traces were recorded on-line with an HP-1050 photodiode array detector from Hewlett-Packard (Palo Alto, CA, USA).

2.5. LC–TSP-MS analyses

A Finnigan MAT TSQ-700 (San Jose, CA, USA) triple quadrupole instrument equipped with a TSP 2 interface was used for the data acquisition and processing. The temperature of the TSP source block was maintained at 280°C and the vaporiser was set to 100°C giving aerosol temperatures varying between 305 and 330°C (according to the eluent composition). These temperature settings represent a good compromise for the analysis of both glycosides and aglycones within the same plant extract. The electron multiplier voltage was 1800 V, dynode −15 kV and the filament and discharge were off in all the cases. Usually full scan spectra from m/z 150–1500 in the PI and NI mode were recorded (scan time 3 s). Post-column addition of buffer (ammonium acetate 0.5 *M*) was achieved by a Waters 590-MS programmable HPLC pump (0.2 ml/min).

2.6. LC–CF-FAB-MS analyses

A continuous flow-FAB interface series-70 Bio-Probe from Finnigan MAT (San Jose, CA, USA) was used with the same mass spectrometer as described for LC–TSP-MS. Post-column addition of the matrix (glycerol 50% aqueous solution) was achieved by a Waters 590-MS programmable HPLC pump (0.15 ml/min) producing a matrix concentration of 7% in the eluent. In order to obtain an LC flow-rate compatible with CF-FAB operation, the total eluent flow-rate (1.05 ml/min) was reduced to 10.5 μl/min with a splitter (microflow processor splitter: 1/100 (Acurate®)) from LC-Packings (Zürich, Switzerland). For stable operation, the copper tip of the CF-FAB interface was maintained at 50°C and the source block was kept at 100°C. The FAB gun was set to 4 kV and 1.2 mA; xenon gas was used for bombardment.

2.7. LC–ESP-MS analyses

For the LC–ESP-MS analyses a Finnigan MAT TSQ-7000 (San Jose, CA, USA) triple quadrupole instrument equipped with the standard Finnigan API source was used. In ESP experiments, the temperature of the stainless steel capillary was set to 220°C for optimum performance (molecular ion intensity and sensitivity). This optimum temperature proved to be practically independent of the flow-rate of liquid into the API source. Nitrogen was used as nebulizing gas at a pressure of 60 p.s.i. (the resulting gas flow in the source was not calibrated). The flow-rate was kept at 1 ml/min for loop injection of pure standards and when extracts were analysed by liquid chromatography. TFA (0.05%) or NH_4OAc (0.01 *M*) was added to the water in order to increase the separation efficiency and to aid ion formation. These buffers were used in both ion polarities according to the type of information required. For the UV detection, about two-thirds of the eluent was split into the UV detector. In order to suppress solvent clustering, the r.f. potential of the source octapole was set to 10 V. When structural information was required, the potential of the octapole was raised up to 100 V, until about 50% of the precursor was lost due to collision-induced dissociation (up-front CID) and scatter. If available, CID conditions were optimised with pure standards. The standards were infused in these cases via a syringe pump at a liquid flow-rate of 5 μl/min.

2.8. LC–TSP-MS–MS analyses

The LC–TSP-MS–MS analysis of the flavone glycosides from *Halenia corniculata* were performed on the TSQ-700 MS instrument. Argon (TSQ) was used as collision gas. Collision conditions (gas pressure and collision energy) were optimised, to obtain maximum structural information for the different types of ions. For the analysis of flavone C-glycosides, the CID conditions were offset −30 V and pressure 0.6 mTorr.

2.9. LC–ESP-MS^n analyses

The LC–ESP-MS^n analyses of the saponins from *Phytolacca dodecandra* were performed on a prototype version of a Finnigan MAT LC–MS ion trap mass spectrometer (LCQ) (San Jose, CA, USA), equipped with the standard Finnigan API source. Helium was used for CID experiments, with the gas pressure optimised to obtain maximum structural information for the different types of ions. In the LCQ daughter experiments, a digital frequency synthesiser was used to isolate the parent ion of interest. The instrument then agitated the ions with the appropriate frequency and scanned out the products at a rate of 0.3 s/2000 u. If multiple stages MS–MS (MS^n) experiments were performed, only one of the product ions were kept in resonance, all other fragments were ejected from the trap without mass analysis. The ion was then agitated again and allowed to fragment by collision-induced dissociation (CID). The collision energy was adjusted experimentally to give >90% yield of fragmentation (0.5–1 V). Five to six cycles of MS^n could be performed with most compounds.

2.10. LC–NMR analyses

A Unity Innova 500-MHz NMR instrument (Varian, Palo Alto, CA, USA) equipped with a Varian ^1H [^{13}C] pulse field gradient indirect detection microflow LC–NMR probe (flow cell 60 μl, 3 mm O.D.) was used. Reversed-phase HPLC of the crude extracts was carried out on a Varian modular HPLC system, comprising a Varian 9012 pump, a Valco injection valve and a Varian 9050 UV detector. The Varian HPLC software was also equipped with the capability for programmable stop-flow experiments. The same HPLC conditions as those used for LC–MS were performed except that H_2O was replaced by D_2O.

3. Role of LC–UV and LC–MS in the early recognition of plant metabolites

The plant kingdom represents an extraordinary reservoir of novel molecules. Of the estimated 400 000–500 000 plant species around the globe, only a small percentage has been investigated phytochemically and the fraction submitted to biological or pharmacological screening is even lower. There is thus currently a new interest in the vegetable kingdom as a possible source of new leading compounds for introduction into therapeutical screening programmes. The rapid disappearance of tropical forests and other important areas of vegetation has meant that it is essential to have access to methods which lead to the rapid isolation and identification of bioactive natural products. The approach adopted, to obtain an exploitable pure plant constituent, involves interdisciplinary work in botany, pharmacognosy, pharmacology, chemistry, toxicology and can be formulated as follows (see Fig. 1) [1].

By following only a bioactivity-guided fractionation procedure, there is a risk of unnecessarily isolating known plant constituents with recognised activity. Furthermore, interesting leading compounds which do not exhibit the tested activity will simply

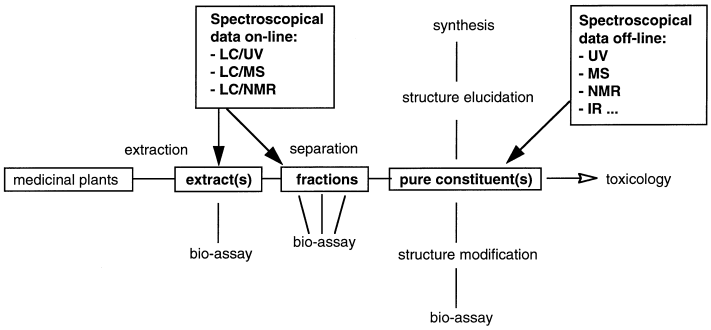

Fig. 1. Procedure for obtaining the active principles from plants and use of LC hyphenated techniques as strategic analytical screening tools during the isolation process of a plant extract.

be missed. In order to avoid the time-consuming isolation of known constituents, hyphenated techniques such as LC–MS or LC–UV are used at the earliest stage of separation (Fig. 1) [9]. This analysis is valuable to detect compounds with interesting structural features and to target their isolation.

For this chemical screening, the crude plant extracts are separated by HPLC using a reversed-phase gradient system. The different metabolites first pass through a UV photodiode array detector and then are transferred to the mass spectrometer via an adequate interface (Fig. 2). LC–MS should ideally provide molecular weight and structural information for the secondary metabolites of interest within the crude plant extract. As LC–MS, using TSP or ESP interfaces, is a soft ionisation technique, tandem MS experiments (LC–MS–MS) have also been performed in order to obtain additional structural information.

3.1. Complementary related hyphenated technique LC–NMR

LC–MS and LC–UV are not always sufficient for the on-line identification of natural products, information on the atom bonds are often necessary. Despite being known for over 15 years, HPLC coupled with nuclear magnetic resonance (LC–NMR) [8] has not yet been a widely accepted technique, mainly because of its lack of sensitivity. However, the recent progress in pulse-field gradients and solvent suppression, the improvement in probe technology and the construction of high-field magnets, have given a new impulse to this technique. LC–NMR has an important potential for on-line structure identification of natural products. Indeed, nuclear magnetic resonance ^{1}H NMR spectroscopy is by far the most powerful spectroscopic technique for obtaining detailed structural information about organic compounds in solution [10]. While the LC coupling itself was rather straightforward compared to LC–MS [2], the main problem of LC–NMR was the difficulty of observing analyte resonances in the presence of the much larger resonances of the mobile phase. This problem has even worsened in the case of typical LC reversed-phase operating conditions, where more than one protonated solvent was used and where the resonances changed frequencies during the analysis in gradient mode. Furthermore, the continuous flow of sample in the detector coil complicated solvent suppression. These problems have now been overcome thanks to the development of fast reliable and powerful solvent suppression techniques such as WET (water suppression enhanced through T_1 effects) [11], which produced high-quality spectra in both on-flow and stopped-flow modes. These techniques consist of a combination of pulsed-field gradients, shaped r.f. pulses, shifted laminar pulses and selective ^{13}C decoupling, and are much faster than classical presaturation techniques previously used in this field [11]. Thus, in typical reversed-phase HPLC conditions, non-deuterated solvents such as MeOH or MeCN can be used, while water is replaced by D_2O.

A general setup of the LC–UV–MS and the LC–

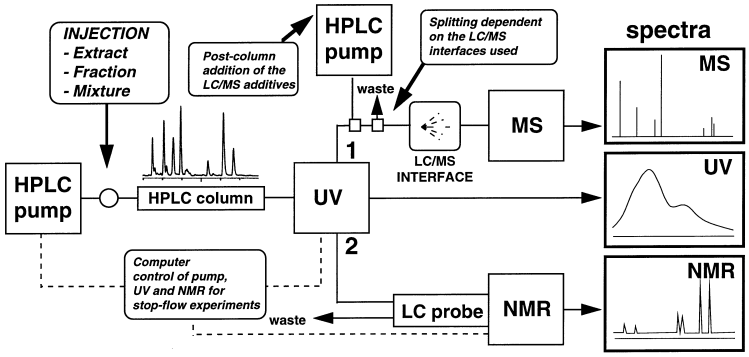

Fig. 2. Schematic representation of the experimental setup used for LC–UV–MS (1) and LC–UV–NMR (2) analyses.

NMR configuration used for screening chemically crude plant extracts is presented in Fig. 2.

4. The choice of an adequate LC–MS interface

As mentioned earlier, a crude plant extract is a very complex mixture containing sometimes hundreds of metabolites with various physicochemical and spectroscopic properties. In order to obtain a satisfactory ionisation of the compounds of interest within this mixture, the choice of a proper LC–MS interface and the optimisation of the ionisation conditions are thus factors of utmost importance. As all the different LC–MS interfaces have their own characteristics and ionisation mechanisms, an adequate interface and conditions have to be found for an optimum detection of a given analyte.

The important variability of the LC–MS response is illustrated by the analysis of two important classes of natural products, polyphenols and triterpene glycosides.

4.1. Comparison of TSP and ESP ionisation of polyphenols

The root MeOH extract of a Gentianaceae plant, *Chironia krebsii*, particularly rich in xanthones, which are a rare class of polyphenols, has been analysed by LC–TSP-MS and LC–ESP-MS in both positive and negative ion modes [3]. The crude MeOH extract (400 μg) was injected without prepurification on a C_{18} column using a gradient of MeCN and H_2O (Fig. 3). For comparison purposes, seven compounds (**1–7**) were selected in the different LC–MS traces of the extract.

The LC–TSP-MS analysis (Fig. 3c) showed an MS response for all the peaks detected by UV at 254 nm. Nevertheless, some highly hydroxylated xanthones such as **5** gave a weaker MS response than the other polyphenols when compared to the UV trace (254 nm). For the xanthone diglycosides **3** and **4**, clearly discernible $[M+H]^+$ ions were observed together with the aglycone fragments $[A+H]^+$, which were the base peaks. In the case of the xanthone monoglycoside (**2**), the $[M+H]^+$ ion was the main ion of the spectrum (Fig. 4a). In the NI mode (Fig. 3d), 1,3,7,8-tetrahydroxyxanthone (**5**) gave rise to $[M-H]^-$ ion but the less polar aglycones such as **7** did not give any response. The glycosides with hydroxyl groups on the aglycone moiety (**2** and **3**) showed weak $[M-H]^-$ and $[A-H]^-$ ions but fully methoxylated xanthones such as **4** were not detectable.

On the other hand, in LC–ESP-MS (Fig. 3a and 3b), detection in both PI and NI modes is necessary for the achievement of a general picture of the extract composition and the spectra obtained often presented different types of ions (formation of doubly charged ions or dimers). The LC–ESP-MS analysis with NH_4OAc as buffer in the PI mode (Fig. 3a) showed more selectivity towards the UV active compounds of *C. krebsii* than LC–TSP-MS. Indeed only compounds **1**, **4** and **7** were detected. The fully methoxylated xanthones (**4** and **7**) exhibited $[M+H]^+$ and $[A+H]^+$ ions. The xanthone monoglycoside (**2**) was not detectable in the extract but the spectrum recorded by loop injection of the pure standard (10 μg) exhibited an $[M+H]^+$ ion at m/z 421 and a dimer $[2\cdot M+NH_4^+]$ at m/z 858 (Fig. 4c). In the NI mode, as with TSP, xanthones **2**, **3**, **5** and **6** gave clearly discernible ions (Fig. 3b). The ESP spectrum of **2** for example showed an acetate adduct $[M+CH_3COO^-]$ at m/z 478 together with the deprotonated ion $[M-H]^-$ at m/z 419, an ion at m/z 839 corresponded to the deprotonated dimer (Fig. 4d). In this case, no aglycone ion was observed. However, by applying an r.f. potential (50 V) on the source octapole (up-front CID), the aglycone ion of **2** $[A-H]^-$ was observable at m/z 257 (Fig. 4e). When TFA was used as buffer, TFA adduct ion $[M+CF_3COO^-]$ was detected as base peaks at m/z 533 (Fig. 4b).

Compound **1** (sweroside) is a monoterpene glycoside, one of the bitter principles of *C. krebsii*, and is common to many Gentianaceae species. Sweroside, contrary to the polyphenols, was detected in all the ionisation modes except in ESP NI (Fig. 3). In the PI mode, a $[M+H]^+$ ion at m/z 359 was observed in TSP and ESP; in the NI mode, an acetate anion adduct $[M+CH_3COO^-]$ was recorded at m/z 417 in TSP.

4.2. Comparison of TSP, ESP and CF-FAB ionisation of triterpene glycosides

Triterpene glycosides such as saponins are frequent constituents of higher plants [12]. Mono-

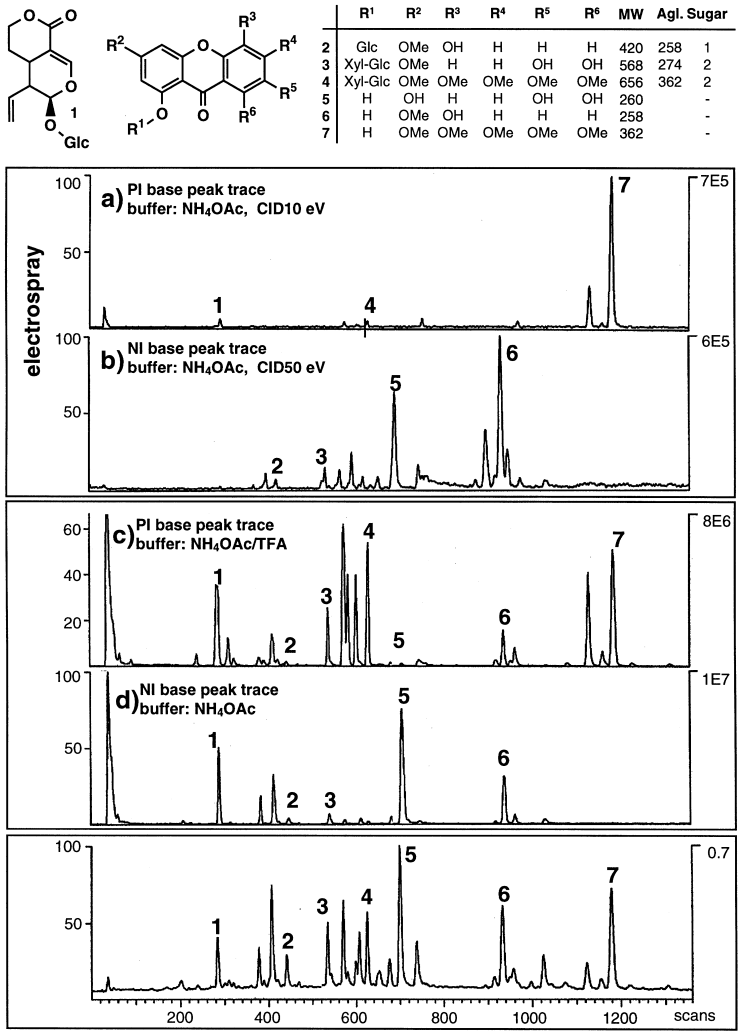

Fig. 3. LC–TSP-MS and LC–ES-MS of the MeOH extract of *Chironia krebsii* (Gentianaceae) [3]. HPLC: C_{18} Nova-Pak (4 μm, 150×3.9 mm I.D.); gradient, CH_3CN–H_2O (5:95→65:35 in 50 min; 1 ml/min). Injection 400 μg. The analyses were performed in the full scan mode (150–1500 u).

desmosidic saponins (glycosylated at position C_3, with a free carboxylic group at position C_{28}) are known to exhibit important molluscicidal activities [12,13] (Fig. 5).

LC–ESP-MS was compared to LC–TSP-MS and LC–CF-FAB-MS [3] for the analysis of saponins in the crude MeOH extracts of the fruits of *Swartzia madagascariensis* (Leguminosae) [14], a plant used in field trials for the local control of schistosomiasis in Africa [13]. The crude MeOH extract (600 μg) was injected without prepurification on a C_{18} column (MeCN–H_2O gradient (20→50 in 30 min; 0.9 ml/min). In order to compare the LC–MS results obtained with the three different interfaces, the on-line spectra of an oleanolic acid saponin (**8**) bearing four sugar units will be discussed here (Fig. 6).

On the LC–TSP-MS (PI, NH_4OAc buffer) spectrum of the tetraglycosylated saponin (**8**, Fig. 6a), the expected molecular ions at m/z 1103 $[M+H]^+$ or m/z 1120 $[M+NH_4^+]$ were not visible. Only the dehydrated oleanolic acid moiety at m/z 439 was recorded, together with weak fragments. However, under these TSP conditions, the saponin diglycosides, also present in the extract, exhibited

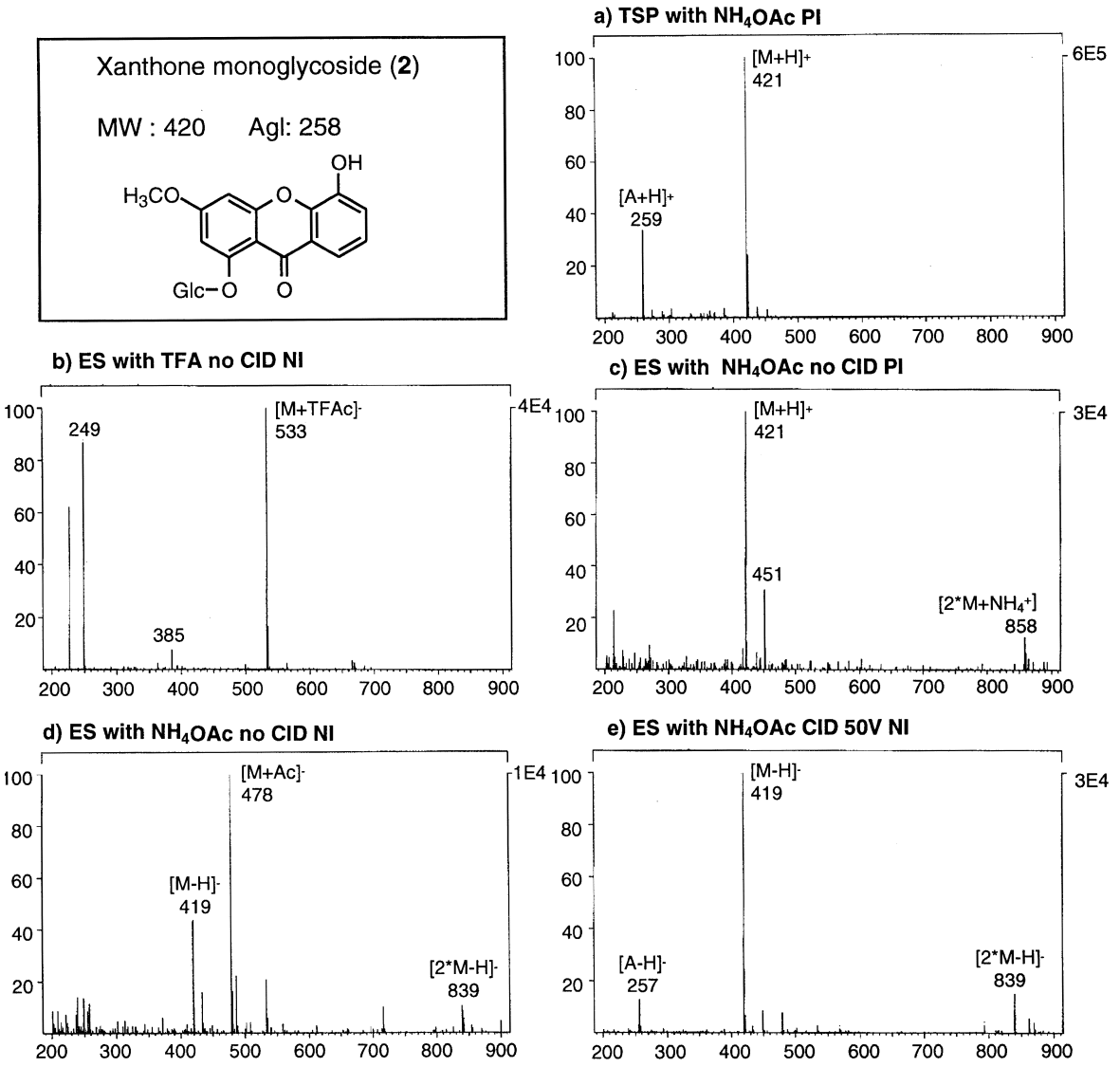

Fig. 4. LC–TSP-MS and LC–ES-MS spectra of the xanthone monoglycoside (**2**) of the MeOH extract of *Chironia krebsii* (Gentianaceae).

clear $[M+NH_4^+]$ pseudomolecular ions [3]. The LC–TSP-MS analysis alone did not provide molecular weight information on-line for large saponins.

The LC–CF-FAB-MS (NI, glycerol matrix) spectrum of **8** (Fig. 6b) exhibited a molecular ion at m/z 1101 $[M-H]^-$ as base peak. The different fragment ions observed (m/z 939 $[M-H\text{-}162]^-$ and m/z 955 $[M-H\text{-}146]^-$) confirmed the presence of terminal hexosyl and desoxyhexosyl units in this saponin. The other ions at m/z 777 and 793 indicated an additional hexosyl unit. Calculation by difference with the aglycone moiety ion $[A-H]^-$ (m/z 455) suggested the presence of a glucuronic acid moiety, probably directly linked at C_3. The LC–CF-FAB spectrum was thus very informative and proved that saponin (**8**) was a tetraglycosylated triterpene with two hexosyl, one deoxyhexosyl and one glucuronic acid moieties.

The LC–ESP-MS (NI, NH_4OAc buffer, CID 50 eV) of **8** (Fig. 6c) gave, as in the case of CF-FAB,

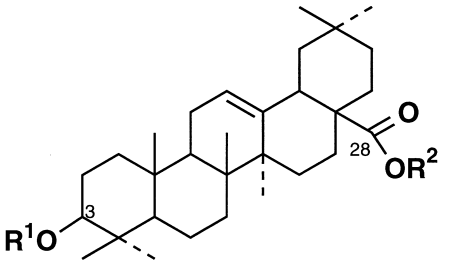

Fig. 5. General structure of the triterpene glycosides (saponins), the aglycone part here is oleanolic acid. R^1 and/or R^2 are sugar chains.

the $[M-H]^-$ ion as base peak. With the aid of up-front CID (50 eV) in the octapole of the ESP source, ions due to the loss of different sugar moieties were also observable. Contrary to CF-FAB, the aglycone ion at m/z 455 was not visible.

It is also important to mention that, in ESP, almost no LC-peak broadening was observed while, in CF-FAB, a large diminution of the chromatographic resolution was due to the important splitting and to the post-column addition of the glycerol matrix.

These different examples have shown that no universal interface allows an optimum ionisation of all the metabolites within a single crude plant extract. Often, different ionisation modes or different interfaces are necessary to obtain a complete picture of the extract composition and to ensure molecular weight determination. This point is essential to have in mind especially when screening unknowns.

5. LC–UV and LC–TSP-MS for the screening of xanthones in Gentianaceae species

Numerous phytochemical studies have been performed on species of the Gentianaceae family in our

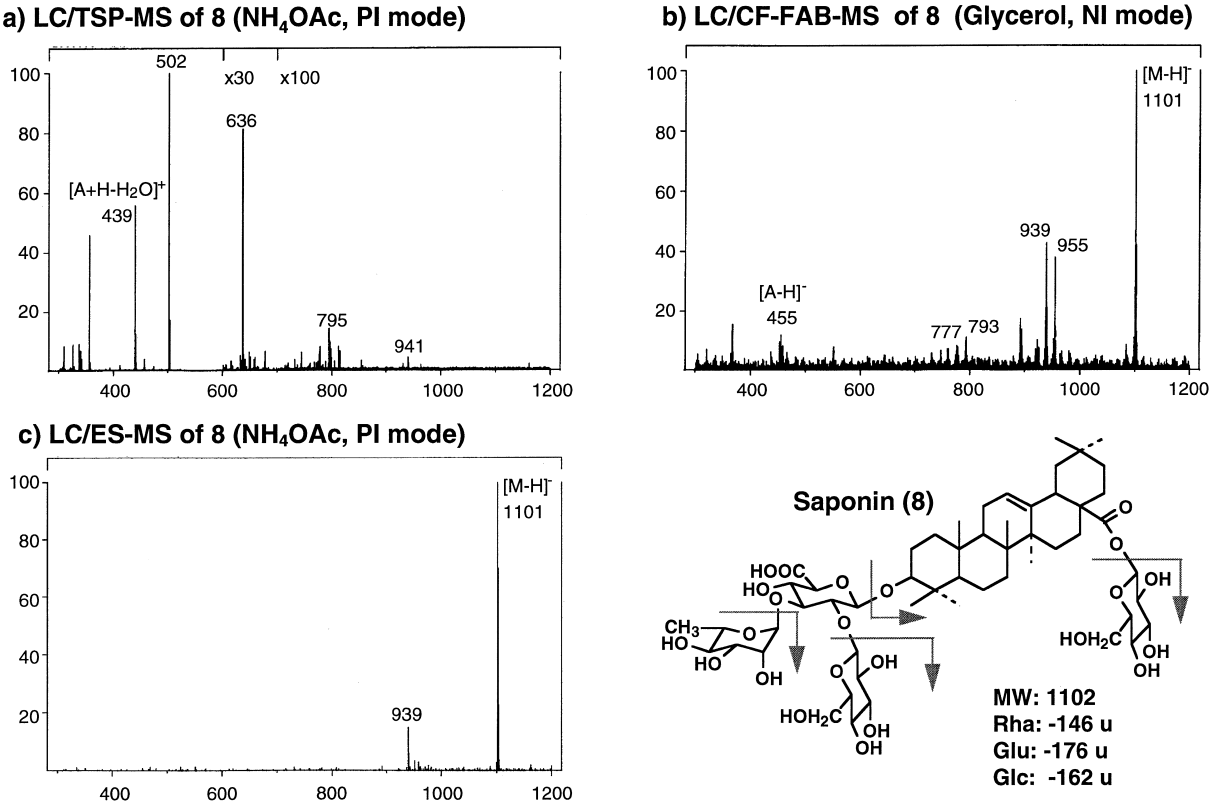

Fig. 6. LC–TSP-MS, LC–CF-FAB-MS and LC–ES-MS spectra of the saponin tetraglycoside (**8**) of the MeOH extract of *Swartzia madagascariensis* (Leguminosae).

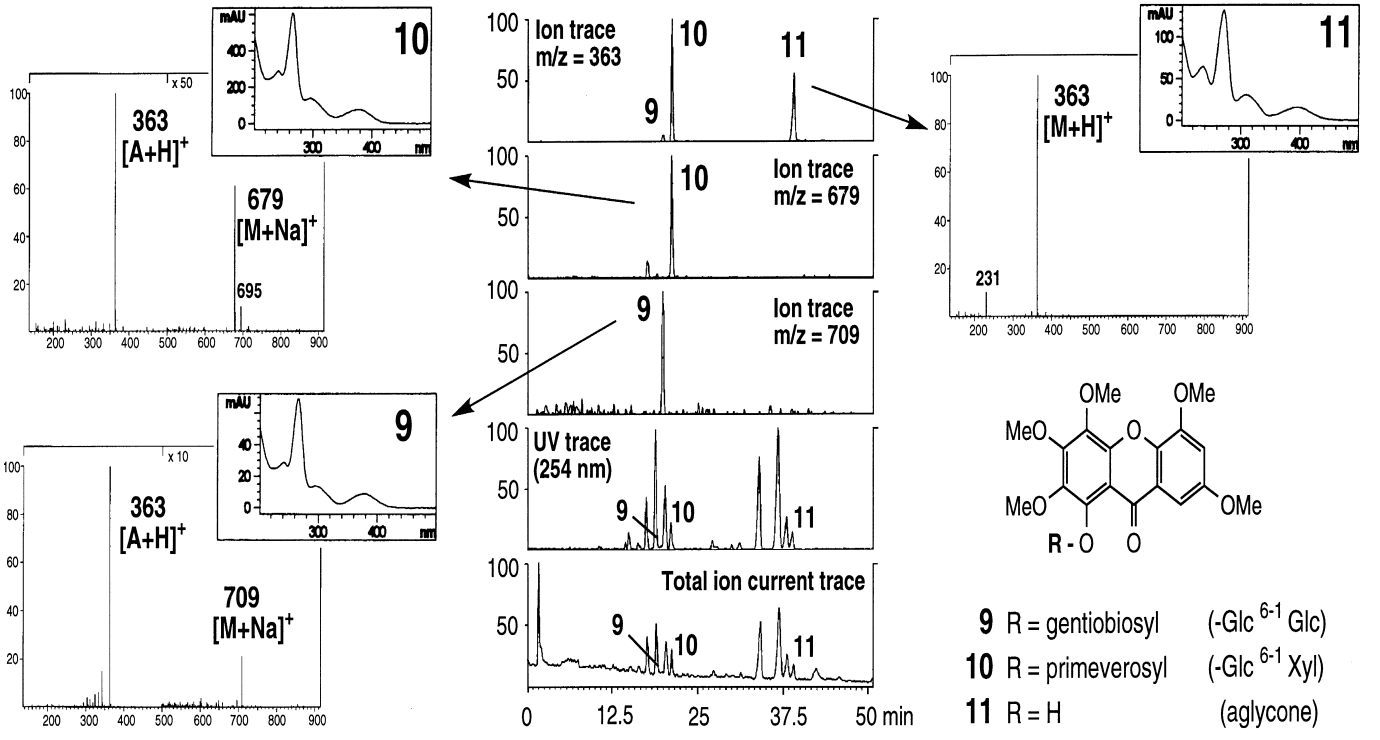

Fig. 7. LC–UV–MS analysis of the crude dichloromethane extract of *Halenia corniculata* [16]. The UV and MS spectra of three related xanthones (aglycone (**11**) and its two glycoside derivatives (**9**) and (**10**)) are displayed. The ion m/z 363 is the main fragment and corresponds to the aglycone moiety $[A+H]^+$. The ion traces displayed at m/z 679 and 709 correspond to the molecular adduct ions $[M+Na]^+$ of gentiobioside (**9**) and primeveroside (**10**), respectively. Same LC conditions as in Fig. 3.

laboratories [15,16]. These plants are indeed very interesting for their potential as new antidepressive drugs. In Europe, gentians and especially the yellow gentian (*Gentiana lutea*) are well-known for their bitterness and are used to cure digestive problems. This bitter character is mainly due to their content of monoterpene glycosides of the secoiridoid class. In traditional Indian medicine, these plants are used to cure light depressions. Research started in 1972 by Indian scientists [17] showed that the gentian extracts, especially those of *Canscora decussata*, stimulated the central nervous system. Our group, together with that of Suzuki [18], demonstrated that xanthones presented a remarkable in vitro inhibition of monoamine oxidase (MAO), an enzyme which plays an important role in the regulation of the neurotransmitters in the central nervous system. As it is known that patients suffering from depression lack some neurotransmitters, xanthones are potentially interesting as new antidepressive drugs.

In order to find new xanthones, numerous gentians have been screened chemically by both LC–UV and LC–MS [15,16]. Without prepurification, the crude extracts (ca. 200 μg) were directly separated on a 4-mm I.D. reversed-phase column with an acidic acetonitrile–water gradient (1 ml/min). Under these conditions, it was possible to obtain a satisfactory separation of the main constituents. The UV spectra recorded on-line allowed the distinction of xanthones from other constituents. These polyphenols show characteristic UV spectra usually with four bands of decreasing intensity between 200 and 400 nm [19]. The other metabolites were generally secoiridoids (one UV maximum at ca. 240 nm) or flavones (two UV maxima at 200–400 nm).

The LC–UV chromatogram of the dichloromethane extract of *Halenia corniculata*, a Gentianaceae from Mongolia, presented more than 19 peaks having UV spectra characteristic for xanthones [16]. Two types of xanthone *O*-glycosides as well as xanthone aglycones were detected in this extract. A display of the selective ion traces of the molecular ions of the aglycones and the glycosides of these xanthones was used for their specific detection. As shown in Fig. 7, for example, three xanthones of *H. corniculata* (**9**, **10** and **11**) appeared in the single ion trace m/z 363, which indicated the presence of a common hexasubstituted aglycone (one hydroxyl and five methoxyl groups) in each case. This was confirmed by similar UV spectra for the three xanthones. The glycosides **9** and **10** were detected separately at m/z 709 and 679 u, respectively. These ions corresponded to weak molecular sodium adducts $[M+Na]^+$ of the xanthone glycosides. Important fragments corresponding to the aglycone moiety $[A+H]^+$ were also observed in the LC–TSP-MS spectra (Fig. 7). The difference of 316 u between the ions $[M+Na]^+$ and $[A+H]^+$ in the TSP-MS spectrum of **10** was characteristic of the loss of a disaccharide moiety constituted of hexosyl and pentosyl units (-pentosyl (132)–hexosyl (162)–Na (23)+H (1)). Likewise, a difference of 346 u (-hexosyl (162)–hexosyl (162)–Na (23)+H (1)) for **9** indicated the presence of a dihexosyl derivative. As primeverosyl and gentiobiosyl are the only disaccharide residues corresponding to these masses which have been found to date in the Gentianaceae family, **9** and **10** were, respectively, the gentiobiosyl and the primeverosyl derivative of the corresponding free aglycone (**11**). The UV spectra of the three compounds were very similar but did not correspond to any of the available data. Thus, a rare oxidation pattern was indicated by these on-line data. Following this approach, other series of aglycones and corresponding glycosides (**12-25**) were detected on-line (Fig. 8). The isolation of the most interesting xanthones was subsequently undertaken [16]. The analysis of the isolated compounds confirmed the results obtained on-line. The xanthones (**9–13**, **15**, **17**, **22** and **23**) were found to be new natural products [16].

This example of LC–UV–MS analysis showed that, for simple known compounds like xanthones, a precise identification can be performed on-line, provided that some information about the type of the constituents and their occurrence in given plant families is already available. When needed, more structural information on the position of the hydroxyl groups on the polyphenol skeleton can also be obtained by LC–UV with the aid of post-column addition of UV shift reagents [15].

5.1. LC–MS-MS and LC–MS^n as an important complement

The LC–MS interfaces generally produce a soft

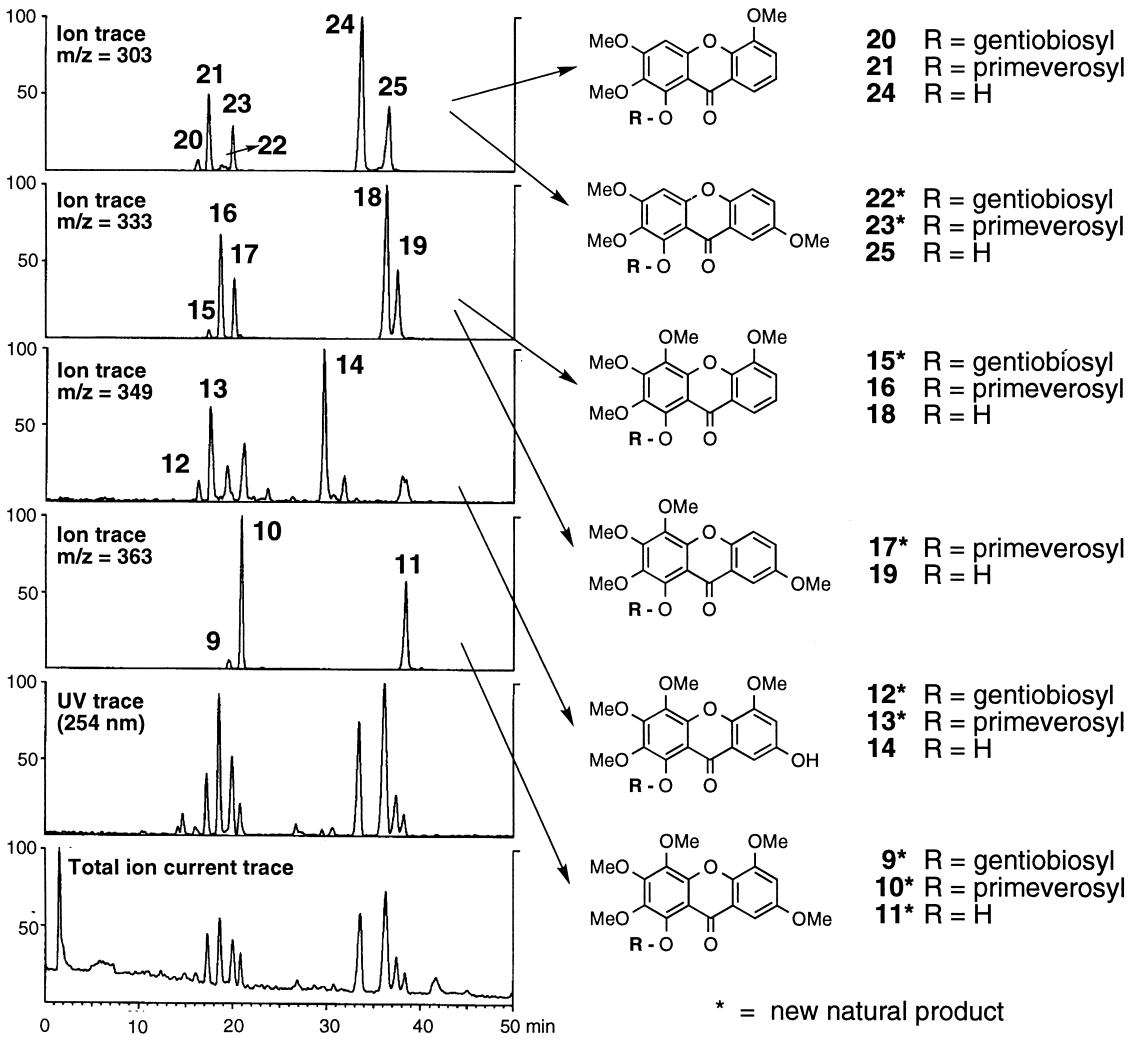

Fig. 8. LC–UV–MS analysis of the crude dichloromethane extract and structures of all the xanthones identified in *Halenia corniculata* [16]. For each oxidation pattern, the aglycone moiety ion trace is displayed. Thus, series of xanthones with the same aglycone can be observed. Same LC conditions as in Fig. 3.

ionisation of the plant metabolites and gave mainly molecular weight information [3]. Some structurally significant fragments such as sugar losses, for example, can be observed with LC–TSP-MS, but usually only limited structural information is obtained. In order to record more characteristic fragments, collision-induced dissociation (CID) spectra of the metabolites can be performed using tandem mass spectrometers (MS–MS) or ion-trap MS systems.

5.2. LC–MS^n analysis of triterpene glycosides

As shown earlier, the LC–ESP–MS (NI, TFA buffer) analysis of saponins provided molecular weight information, an MS response in the ng range, but almost no fragment ions. Thus, in order to obtain more structure information, MS-MS experiments have to be performed. Multiple stage MS-MS experiments (MS^n) can be performed with ion-trap mass

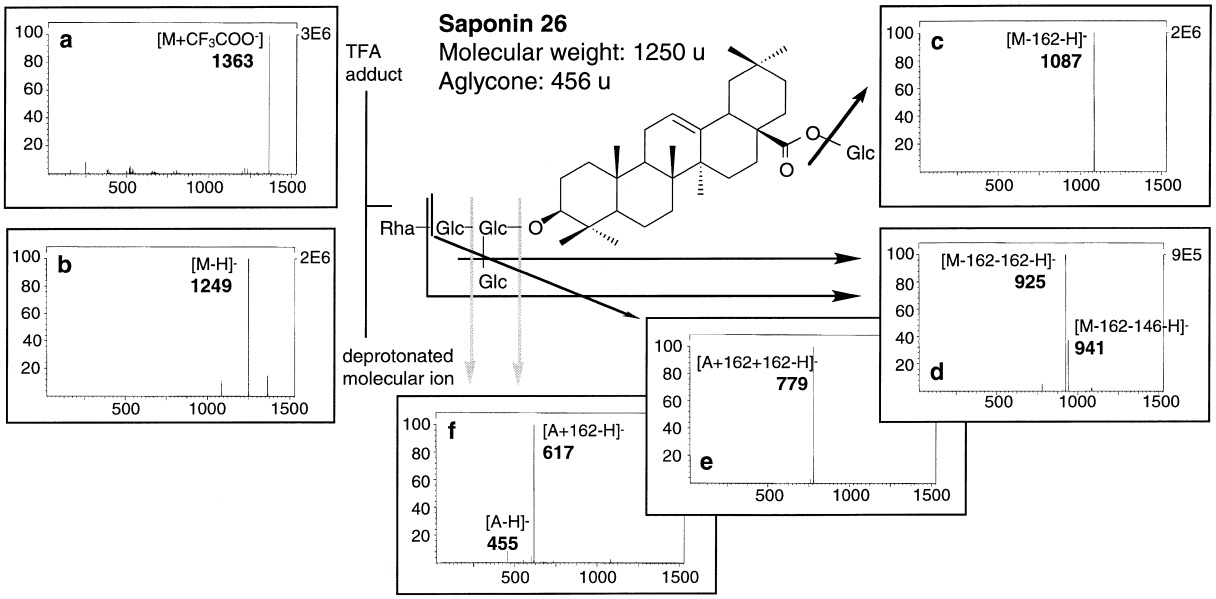

Fig. 9. ES-MS^n spectra of the pentaglycosylated saponin (**26**) from the methanolic fruit extract of *Phytolacca dodecandra* (Phytolaccaceae) [3]. Sample (1 mg/ml) injected by a syringe pump (5 μl/min) (infusion experiment). This experiment allowed a sequential fragmentation of the saponin sugar chain (cleavage of only one sugar at each MS-MS step), clarifying the structure determination.

analyser on compounds such as glycosides to determine their sugar sequence [3]. With such an instrument only one ion of interest can be isolated and excited. At the same time, the other consecutive reactions usually observed in a high-pressure collision cell are reduced. By adjusting the collision energy, it is possible to cleave only one sugar at a time, making the interpretation of the spectra very simple.

For example, an MS^n experiment was performed on the pentaglycosylated saponin (**26**) from *Phytolacca dodecandra*, a plant used as *Swartzia madagascariensis* for the local control of schistosomiasis (Fig. 9). The TFA anion adduct at *m/z* 1363 (Fig. 9a) was easily fragmented in the ion trap giving the deprotonated ion *m/z* 1249 (Fig. 9b). This latter $[M-H]^-$ ion yielded a first fragment at *m/z* 1087 (-Glc) (Fig. 9c) which then cleaved into ions at *m/z* 925 or 941 showing the simultaneous loss of a glucosyl or a rhamnosyl unit, respectively. These losses were characteristic of a branched sugar chain. The ion *m/z* 779 (Fig. 9e) issued from the fragmentation of *m/z* 925 (-Rha) or 941 (-Glc). Finally, the *m/z* 617 ion and the aglycone ion at *m/z* 455 were observed from cleavage of *m/z* 779.

This type of MS^n experiment was found to be very helpful for clarifying the sugar sequence of saponins.

6. Complementary detection with LC–NMR

As shown, the LC–UV–MS and LC–MS–MS analyses of plant extracts provides numerous useful structural information, but often this information is not sufficient enough for a full on-line structural identification of plant metabolites. Other hyphenated techniques such as LC–NMR are thus needed for deeper structural investigations.

6.1. On-flow LC–NMR analysis of Swertia calycina

The dichloromethane extract of another plant from the Gentianaceae family, *Swertia calycina*, has been screened by LC–UV, LC–MS and LC–NMR. This extract presented a strong antifungal activity against

Cladosporium cucumerinum and *Candida albicans* [20]. Various hyphenated LC techniques were used in order to identify the antifungal agent directly in the extract, prior to isolation.

The LC–UV chromatogram of the dichloromethane extract of *S. calycina* was rather simple and only three main peaks (**1**, **27** and **28**) were detected (see *y*-axis chromatogram in Fig. 10). The LC–UV and LC–MS data, together with chemotaxonomical considerations, allowed the identification of most xanthones and secoiridoids of this plant. Compound **28** presented a UV spectrum with four absorption bands characteristic of a xanthone. Its TSP-MS spectrum exhibited a strong protonated ion at *m/z* 303, indicating a xanthone with a molecular weight of 302, thus substituted by one hydroxyl and three methoxyl groups. This information, together with the comparison with a home-made UV spectral library, permitted the identification of **28** as decussatin, a widespread xanthone in the Gentianaceae family. The on-line data obtained for compound **1** indicated the presence of a secoiridoid-type molecule with a molecular weight of 358 u. The loss of 162 u observable in the TSP-MS spectrum was characteristic of the presence of a hexosyl moiety. As in the case of *C. krebsii*, these data suggested that **1** was most probably sweroside. The UV spectrum of **27**, however, was not attributable to a common polyphenol of the Gentianaceae, such as flavones or xanthones. It was very weakly ionised in LC–TSP-MS, but a protonated molecular ion was nevertheless found at *m/z* 189. This small molecular weight (188 u) and the UV spectrum (maxima at 243, 248, 277, 330 nm) suggested that **27** could be a quinonic compound, but as no metabolite of this type was previously found in the Gentianaceae family, it was not possible to identify it unequivocally.

In order to confirm these attributes and to obtain more structure information on-line, the same extract of *S. calycina* was submitted to an on-flow LC–^{1}H NMR analysis on a 500-MHz instrument [21]. The same LC conditions as for the LC–UV–MS analysis were used except that the water of the LC gradient system was replaced by D_2O. However the quantity of extract injected onto the column was increased to 1 mg to obtain at least 20 μg for each peak of interest. For the suppression of the solvent signal of MeCN and its two ^{13}C satellites, as well as the residual HOD peak, the sequence WET, mentioned before, was run before each acquisition [21].

Under these conditions, the on-line LC–NMR analysis of *S. calycina* provided ^{1}H NMR spectra for all the major constituents. A plot of the retention time (*y*-axis) versus the NMR shifts (*x*-axis) permitted the localisation of the resonances of compounds **1**, **27** and **28** (Fig. 10). The important trace starting from 4.8 ppm (at 0 min) and ending at 4 ppm (at 30 min) was due to the change of the chemical shift of the residual negative water (HOD) signal during the LC gradient. The traces between 1 and 2.6 ppm were due to residual MeCN signal and solvent impurities.

A slicing of this bidimensional plot in single on-line LC–^{1}H NMR spectra for each constituent allowed a precise assignment of their specific resonances. The ^{1}H NMR data of **1** and **28** confirmed their on-line identification by LC–UV–MS as sweroside and decussatin, respectively. On the LC–^{1}H NMR spectrum of **27**, two signals (2H, δ 8.11, *m*, H-5,8 and 2H, δ 7.89, *m*, H-6,7) were characteristic of four adjacent protons of an aromatic ring with two equivalent substituents. The low field shift of the H-5,8 signal indicated that these two protons were in the *peri* position of carbonyl functions suggesting, most probably, the presence of a naphthoquinone nucleus. The singlet at 4.35 ppm was attributed to H_3 and the remaining methoxyl group was thus at position C_2. By combining these on-line data and the molecular weight information deduced from the LC–TSP-MS spectrum (MW 188), **27** was finally identified as 2-methoxy-1,4-naphthoquinone. As this was the first naphthoquinone to be reported in Gentianaceae, **27** was specifically isolated, by a combination of column chromatography on silica gel and HPLC in reversed-phase mode, and was found to be the compound responsible for the strong antifungal activity of the extract of *S. calycina* [20].

6.2. Stopped-flow LC–NMR analysis of Gentiana ottonis

The whole plant MeOH extract of another Gentianaceae, *Gentiana ottonis*, was more complex than those of *S. calycina* and contained metabolites in smaller amounts. In order to obtain satisfactory LC–^{1}H NMR spectra, this extract was analysed in the stopped-flow mode. In this case, as soon as a

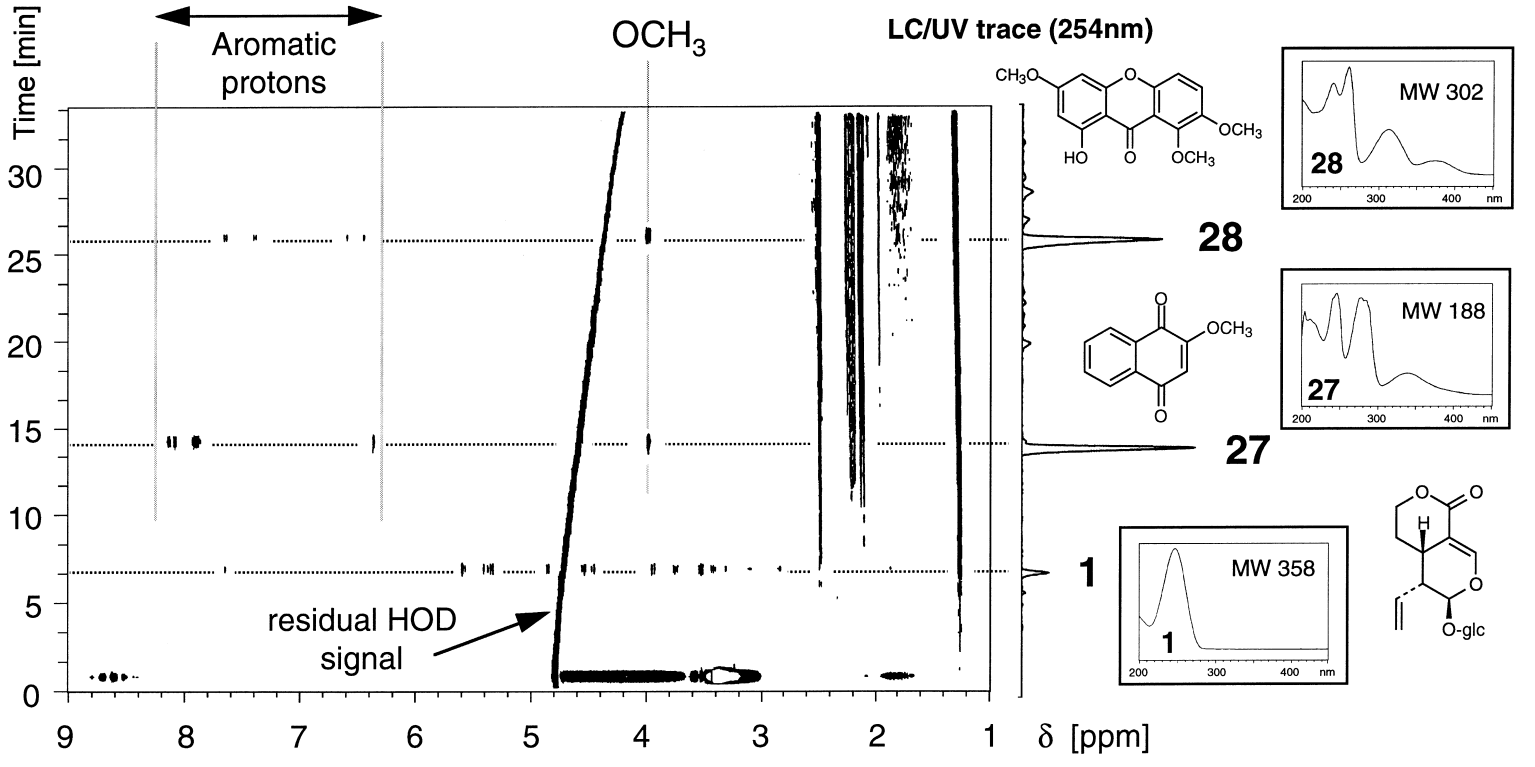

Fig. 10. Bidimensional LC–^{1}H NMR chromatogram of the crude CH_2Cl_2 extract of *Swertia calycina* [21]. Methoxyl groups and aromatic proton signals of **27** and **28** are clearly visible together with all the resonances of the monoterpene glycoside (**1**). The signal of HOD is negative and was continually shifted during the LC gradient.

constituent of the extract was passing through the UV detector, the HPLC flow was stopped after a given delay, the sample was thus kept in the NMR cell for longer acquisition time (see setup in Fig. 2).

The LC–UV analysis of this extract presented peaks with UV spectra characteristic of secoiridoids (**29**), flavones (**31** and **33**) and xanthones (**30**, **32** and **34**–**36**) (Fig. 11). The LC–TSP-MS analysis of this extract allowed the attribution of molecular weights of all these compounds. Compounds **30**, **31** and **33** presented all fragments characteristic for C-glycosides (losses of 90 and 120 u). According to their

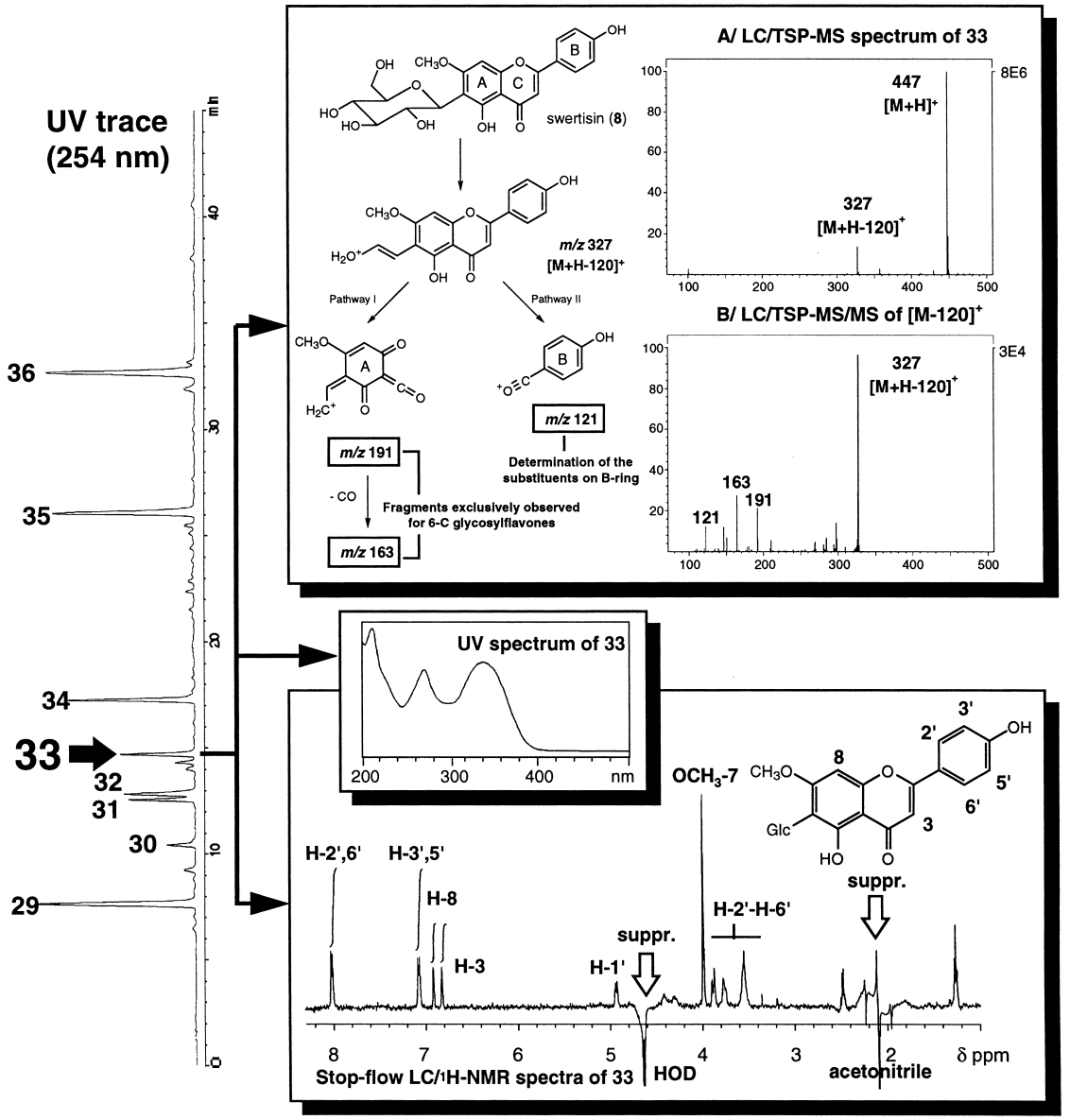

Fig. 11. LC–UV chromatogram of *Gentiana ottonis* (vertical display) with the stop-flow LC–^{1}H NMR, LC–UV, LC–MS and LC–MS–MS spectra of the flavone C-glycoside (**33**) [21]. On these LC–^{1}H NMR spectra, the coupling constants and the integration of the aromatic signals were easily measured.

UV spectra, **31** and **33** (MW 448 and 446) were, respectively, tri- and tetraoxygenated C-glycoside flavones.

In order to show the type of information obtained on-line, the identification of the flavone C-glycoside **33** is described in more detail here. The stopped-flow LC–^{1}H NMR of **33** (MW 446) presented signals of six aromatic protons and one methoxyl group together with those of the C-glycoside moiety. These data allowed a full structure determination of the aglycone moiety of **33** but did not permit an unambiguous location of the C-glycoside moiety.

In order to ascertain the position of C-glycosylation, an LC–MS–MS experiment was performed. As shown in Fig. 11, the LC–TSP-MS-MS spectrum of **33** obtained by choosing the $[M+H\text{-}120]^+$ as parent ion exhibited fragments at m/z 191 and 163 characteristic for 6-C-glycosylated flavones. Indeed it is known that TSP-MS-MS spectra of the $[M+H\text{-}120]^+$ fragments of isomeric C-glycosylflavones show different specific daughter ions [22]. The A–C-ring fragments issued from a retro Diels-Alder cleavage are only observable for the C_6 position isomers, as was the case for **33**. Furthermore the fragment observed at m/z 121 was indicative of a monohydroxylated B-ring, confirming the position of the methoxyl group on the A-ring.

These UV, MS, ^{1}H NMR and MS–MS data allowed the identification of **33** as the known 5,4′-dihydroxy-7-methoxy-6-C-glucosylflavone (swertisin). Similarly, the flavone C-glycoside **31** (MW 448) was identified as the common 5,7,3′,4′-tetrahydroxy-6-C-glucosylflavone (isoorientin).

Following a similar approach, the constituents of the extract were fully identified on-line. The secoiridoid **29** was swertiamarin. The xanthones **30**, **32**, **34**, **35** and **36** were respectively mangiferin 1,3,5-trihydroxy-8-O-glucosylxanthone (8-O-glucosyldesmethybellidifolin), 1,5,8-trihydroxy-3-methoxyxanthone (bellidifolin), 1,3,5,8-tetrahydroxyxanthone (desmethylbellidifolin), 1,5-dihydroxy-3-methoxy-8-O-glucosylxanthone (8-O-glucosylbellidifolin) [21].

These two examples of LC–UV–MS and LC–NMR analyses have shown that it was possible to identify on-line the major metabolites of Gentianaceae crude extracts. In both cases, UV and MS detection was sensitive enough for detecting the compounds of interest, in amounts compatible with an optimum chromatographic resolution. However the sensitivity of LC–NMR, did not permit experiments in the same range of concentrations. With the 500-MHz instrument used, the LC–^{1}H NMR spectra were obtainable in the on-flow mode with approximately 0.05 μmol per peak. These detection limits could be lowered by a factor of ca. 100 in the stopped-flow mode, but long acquisition times were then needed. In this type of experiment a given compound can be maintained in the LC–NMR cell for several hours without appreciable dispersion. It permitted a significant increase in the signal-to-noise ratio of all the LC–^{1}H NMR spectra measured and the realisation of important complementary bidimensional experiments such as ^{1}H–^{1}H and ^{1}H–^{13}C inverse correlations.

7. Conclusions

Today the work of phytochemists lies mainly on bioassay-guided fractionation schemes of crude plant extracts. This type of approach has led to the isolation of numerous compounds with interesting activities, since plants contain thousands of constituents with considerably different biological properties. Obviously for any successful investigation of biological material with such a wide range of properties, the future lies in the ability to have as large a number of biological screens as possible.

Chemical screening of crude plant extracts which allows the localisation and the efficient targeted isolation of new types of constituents with potential activities, can be considered as an efficient complementary approach to the biological screening. As shown, the use of hyphenated techniques such as LC–UV, LC–MS and LC–NMR enables rapid initial screening of crude plant extracts, providing a great deal of preliminary information about the content and nature of constituents of these extracts. According to these structural data, once the novelty or utility of a given constituent is established, it is then important to process the plant extracts in the usual manner, to obtain samples for full structure elucidation and biological or pharmacological testing. In such a way, the unnecessary isolation of common compounds of minor interest is avoided.

The examples discussed have shown that the LC–UV–MS or LC–MS-MS approach alone permits the identification of small metabolites such as xanthones or flavones, provided that some information about the chemotaxonomy of the plant families of interest is already available. For screening more complicated constituents or unknowns, LC–NMR data are of crucial importance for a deeper structural investigation. The recent introduction of LC–NMR for the crude plant extract screening will probably make another breakthrough in the on-line structural determination of natural products.

The chemical screening of extracts with such complex hyphenated techniques generates a huge amount of information. In order to rationalise this approach and use it efficiently with a high sample throughput, the next challenge will be to find a way to centralise all these data for rapid pattern recognition by reference to standard compounds. With such an analytical system, natural product chemists will then be able to concentrate their efforts on finding new biological targets. This aspect still remains the more difficult problem to solve when searching for leading compounds.

Acknowledgements

Financial support was provided by the Swiss National Science Foundation. Thanks are due to Dr. Wolf Hiller, Varian AG, Darmstadt, for the LC–NMR measurements and to Dr. Winfried Wagner-Redeker, Spectronex AG, Basel, for the LC–ESP-MS experiments.

References

[1] M. Hamburger, K. Hostettmann, Phytochemistry 30 (1991) 3864.
[2] W.M.A. Niessen, J. van der Greef, Liquid Chromatography–Mass Spectrometry, Principles and Applications, Marcel Dekker, New York, 1992.
[3] J.-L. Wolfender, S. Rodriguez, K. Hostettmann, W. Wagner-Redeker, J. Mass Spectrom. Rapid. Commun. Mass Spectrom. (1995) S35.
[4] C.R. Blakley, M.L. Vestal, Anal. Chem. 55 (1983) 750.
[5] R.M. Caprioli, F. Tan, J.S. Cotrell, Anal. Chem. 58 (1986) 2949.
[6] R.C. Whitehouse, R.N. Dreyer, M. Yamashita, J.B. Fenn, Anal. Chem. 57 (1985) 675.
[7] K. Hostettmann, B. Domon, D. Schaufelberger, M. Hostettmann, J. Chromatogr. 283 (1984) 137.
[8] N. Watanabe, E. Niki, S. Shimizu, Jeol News 15A (1979) 2.
[9] J.-L. Wolfender, K. Hostettmann, in: J.T. Arnason, R. Mata, J. Romeo (Eds.), Phytochemistry of Medicinal Plants, Plenum Press, New York, 1995, p. 189.
[10] K. Albert, J. Chromatogr. A 703 (1995) 123.
[11] S.H. Smallcombe, S.L. Patt, P.A. Keiffer, J. Magn. Reson. A Ser. A 117 (1995) 295.
[12] K. Hostettmann, A. Marston, Saponins, Cambridge University Press, Cambridge, 1995.
[13] K. Hostettmann, in: H. Wagner, H. Hikino, N.R. Farnsworth (Eds.), Economic and Medicinal Plant Research, vol. 2, Academic Press, London, 1989, p. 73.
[14] C. Borel, K. Hostettmann, Helv. Chim. Acta 70 (1987) 570.
[15] J.-L. Wolfender, K. Hostettmann, J. Chromatogr. 647 (1993) 191.
[16] S. Rodriguez, J.-L. Wolfender, G. Odontuya, O. Purev, K. Hostettmann, Phytochemistry 40 (1995) 1265.
[17] S.K. Bhattacharya, S. Ghosal, R.K. Chaudhuri, A.K. Sanyal, J. Pharm. Sci. 61 (1972) 1838.
[18] O. Suzuki, Y. Katsumata, M. Oya, V.M. Chari, R. Klapfenberger, H. Wagner, K. Hostettmann, Planta Med. 42 (1981) 17.
[19] K. Hostettmann, M. Hostettmann, in: J.B. Harborne (Ed.), Methods in Plant Biochemistry, Academic Press, London, 1989, p. 493.
[20] S. Rodriguez, J.-L. Wolfender, E. Hakizamungu, K. Hostettmann, Planta Med. 61 (1995) 362.
[21] J.-L. Wolfender, S. Rodriguez, W. Hiller, K. Hostettmann, Phytochem. Anal. 8 (1997) 97.
[22] G. Rath, A. Touré, J.-L. Wolfender, K. Hostettmann, Chromatographia 41 (1995) 332.