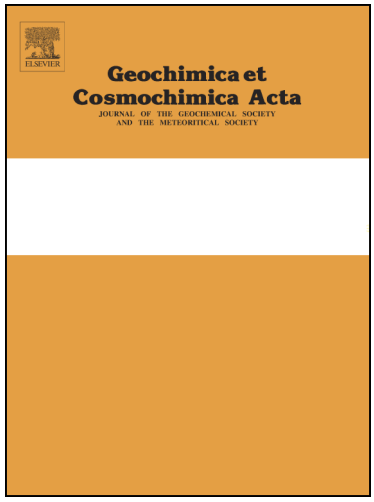
Contribution of NAD 2D-NMR in liquid crystals to the determination of hydrogen isotope profile of methyl groups in miliacin

Philippe Berdagué, Philippe Lesot, Jérémy Jacob, Valery T. Terwilliger, Claude Le Milbeau



Contribution of NAD 2D-NMR in liquid crystals to the determination of hydrogen isotope profile of methyl groups in miliacin

Philippe Berdagué, Philippe Lesot, Jérémy Jacob, Valery J. Terwilliger, Claude Le Milbeau

PII: S0016-7037(15)00586-4
DOI: http://dx.doi.org/10.1016/j.gca.2015.10.004
Reference: GCA 9458

To appear in: *Geochimica et Cosmochimica Acta*

Received Date: 20 February 2015
Accepted Date: 6 October 2015

Please cite this article as: Berdagué, P., Lesot, P., Jacob, J., Terwilliger, V.J., Milbeau, C.L., Contribution of NAD 2D-NMR in liquid crystals to the determination of hydrogen isotope profile of methyl groups in miliacin, *Geochimica et Cosmochimica Acta* (2015), doi: http://dx.doi.org/10.1016/j.gca.2015.10.004

Contribution of NAD 2D-NMR in liquid crystals to the determination of hydrogen isotope profile of methyl groups in miliacin

Philippe Berdagué[a,b], Philippe Lesot[a,b*], Jérémy Jacob[c,d,e], Valery J. Terwilliger[c,d,e,f,g] and Claude Le Milbeau[c,d,e]

[a]*RMN en Milieu Orienté, ICMMO, UMR CNRS 8182, Université de Paris-Sud, Université Paris-Saclay, Bât. 410, 91405 Orsay cedex, France.*
[b]*CNRS/INC, ICMMO, UMR 8182, 91405 Orsay cedex, France.*
[c]*Université d'Orléans, ISTO, UMR 7327, 45071, Orléans, France.*
[d]*CNRS/INSU, ISTO, UMR 7327, 45071 Orléans, France.*
[e]*BRGM, ISTO, UMR 7327, BP 36009, 45060 Orléans, France.*
[f]*LE STUDIUM®, Loire Valley Institute for Advanced Studies, France.*
[g]*Department of Geography, University of Kansas, Lawrence, KS 66045, USA.*

Abstract

The hydrogen isotopic composition (δD or (D/H) value) of molecular biomarkers preserved in sedimentary archives is increasingly used to provide clues about the evolution of past climatic conditions. The rationale is that intact biomarkers retain isotopic information related to the climatic conditions that prevailed at the time of their synthesis. Some of these biomarkers may be degraded during diagenesis, however. The extent to which these degradations alter the original δD value of the source biomarker is presently debated and the capacity to resolve the debate by determinations of compound-specific δD values alone is limited. The "bulk" or "global" δD value of any molecule is in fact a composite of δD values at each site within this molecule ((δD_i or $(D/H)_i$ with i= number of hydrogen/deuterium atoms in the considered molecule). Determination of this site-specific δD_i value in biomarkers could not only yield outstanding paleoenvironmental information but also help forecast the impacts of diagenesis and define essential steps in biosynthetic pathways.

This task is analytically challenging. Here, we examined the capabilities of natural abundance deuterium 2D-NMR (NAD 2D-NMR) using homopolypeptide liquid crystals as an NMR solvent to: i) analyze the NAD spectra of biomakers; ii) determine the site-specific distribution of hydrogen in the nine methyl groups (δD_{Mei} with i= 23 to 31) of miliacin, a pentacyclic triterpene of the amyrin family and key biomarker for broomcorn millet in sedimentary archives. Relative $(D/H)_{Mei}$ values–were established by anisotropic NAD 2D-NMR. Then absolute δD_{Mei} values were obtained by determining δD_{Mei} value of the methoxy group of miliacin using two independent approaches: isotropic NAD NMR (SNIF-NMR™) and GC-irMS. The resulting isotope profile for miliacin shows, for the first time, large variations in δD_{Mei} values that can directly be explained by biosynthetic processes. This approach has also the potential to permit predicting fractionations during pentacyclic triterpene diagenesis.

Corresponding author: Dr. Philippe Lesot

Address: RMN en Milieu Orienté, ICMMO, UMR CNRS 8182, Université de Paris-Sud, 91405 Orsay cedex, France, Tel.: +33 (0) 1 69 15 47 59, Fax: +33 (0) 1 69 15 81 05
E-mail address: philippe.lesot@u-psud.fr (Philippe Lesot)

Keywords: Deuterium NMR, Isotope profile, Site-specific, Natural abundance, Miliacin, Pentacyclic triterpene, Liquid crystal, Biosynthesis, Diagenesis
Type of article: Research article
Supplementary information: Yes

1. INTRODUCTION

Higher plant pentacyclic triterpenes are natural products that bear an oxygenated group at the C-3 position (**Fig. 1**). The large diversity of structures, isomerisms and functional groups resulting from specific enzymatic processes have led to the use of compounds in this family for establishing precise chemotaxonomic relationships (Das and Mahato, 1983; Mahato et al., 1992; Mahato and Sen, 1997). Although β - and α -amyrins are very common structures in angiosperms, several classes of compounds are more or less specific to given taxa. For example, pentacyclic triterpenes with an acetate function at C-3 are largely restricted to Asteraceae species (Lavrieux et al., 2011) whereas those bearing a methoxy group at the C-3 position are mostly synthetized by Poaceae (Ohmoto et al., 1970; Jacob et al., 2005). After the death of their source organisms, these compounds may be archived in geological deposits such as soils or sediments. If diagenesis does not strongly modify their structure, these compounds can be analyzed for presence/absence and concentrations in sedimentary records in order to reconstruct past vegetation dynamics (e.g., Cranwell et al., 1984; Killops et al., 1995; van Aarsen et al., 2000; Regnery et al., 2013).

3β-methoxyolean-18-ene, usually denoted miliacin or germanicol methyl ether (Fig. 1), is such a specific molecular biomarker. It is a pentacyclic triterpene bearing a methoxy group at position C-3 that is predominantly produced by Poaceae species. Furthermore, this compound is most abundant in broomcorn millet (*Panicum miliaceum* L) (Itô et al., 1934; Ohmoto et al., 1970; Bossard, 2013). Miliacin concentrations in lake sediments and soils have thus been used to reconstruct the history of millet cultivation (Jacob et al., 2008, 2008b, 2009; Simonneau et al., 2013; Motuzaite-Matuceviciute et al., 2013).

In addition to these crucial information on past vegetation, the hydrogen isotopic compositions (δD or D/H) of lipid compounds, such as miliacin, in geological deposits are providing superior insights about paleoclimate (Sachse et al., 2012; Terwilliger et al., 2013; Terwilliger and Jacob, 2013). This is because meteoric water is the original source of hydrogen atoms in biochemicals produced by autotrophs and the δD values of meteoric water are strongly influenced by climate (Dansgaard et al., 1964; Gat et al., 1996). Further studies are needed in order to fully exploit the paleoclimate potential of δD values of lipid compounds for providing clues about past climates, with distinct levels of understanding on the processes occurring at different steps (Sachse et al., 2012). The impacts of environmental factors such as soil evaporation and leaf transpiration are relatively well identified. Leaf transpiration can alter the δD value of meteoric water along its path to the sites of photosynthesis where its hydrogen atoms from leaf water will be used to produce organic compounds via biosynthetic pathways (Sessions et al., 1999; Chikaraishi and Narakoa, 2003; Hou et al., 2007b; McInerney et al., 2011). Additionally, it has been demonstrated that plant type can cause differences in δD values of the most common lipid compounds such as *n*-alkyl lipids (Sachse et al.,

2006; Liu et al., 2006; Smith and Freeman, 2006; Hou et al., 2007; Feakins and Sessions, 2010; Wang et al., 2013). Because it would not be affected by variations caused by differences in plant-source taxa, the δD value of more specific compounds such as sedimentary miliacin is expected to solve this issue and could provide even higher resolution information about climatic conditions at the time of their synthesis (Bossard et al., 2011; Bossard, 2013; Terwilliger et al., 2013; Schwab et al., 2015). A better understanding of the biosynthetic pathway that designs the δD values of a given compound is also necessary since distinct NADPH sources and enzymatic processes can lead to highly variable fractionations (i.e. Sauer et al., 2001; Chikaraishi et al., 2004, 2009). Although the biosynthetic pathway of pentacyclic triterpenes such as miliacin is rather well known (Xu et al., 2004), associated H isotopic fractionations are less understood.

Finally, it is crucial to determine whether diagenetic derivatives retain the original paleoclimate information acquired by biochemicals during biosynthesis. For instance, pentacyclic triterpenes can undergo several transformations during early diagenesis such as the loss of functional groups, the loss of rings, double bond migration, aromatization, or epimerization of methyl groups (Corbet, 1980; Trendel, 1985; Lohmann, 1988; Jacob et al., 2007). Only isotopic exchanges with water have been addressed on polycyclic compounds (Wang et al., 2013) but the impact of diagenetic transformations on δD values of diagenetic by-products compared to their precursors is yet unknown.

Our hypothesis is that a site-specific (or intramolecular) heterogeneous distribution of hydrogen isotopes in a biochemical precursor would explain any isotopic difference with its derivatives caused by the preferential loss of D-enriched or D-depleted groups during diagenesis. Identifying such isotopic difference between biochemicals and their diagenetic derivatives will allow better climate reconstructions from δD values of diagenetic derivatives. In addition, obtaining the site-specific distribution of hydrogen isotopes could provide useful information on biosynthetic pathways. As emphasized by Eglinton and Eglinton (2008), "*intramolecular compound-specific isotope measurements (i.e., analyses of the distribution of isotopes within a molecular structure) are analytically challenging and presently rarely attempted, but may ultimately prove of value*" and so addressing this hypothesis requires the development of new tools.

The two main objectives of our study were to test the hypothesis that hydrogen isotopes are not randomly distributed in molecular biomarkers and to identify the means of analyzing the site-specific distribution of hydrogen isotopes necessary for this test. We use miliacin as a representative of pentacyclic triterpenes. Specifically, we determined the (D/H) ratios (δD) of hydrogen located in the methyl groups of miliacin by Natural Abundance Deuterium (NAD), two-dimensional Nuclear Magnetic Resonance (2D-NMR) spectroscopy using aligning organic media (anisotropic systems) as NMR solvents. This advanced technique modifies the well-known SNIF-NMR™ method pioneered

by Martin and co-workers in the 80's (Martin et al., 1980; 1986) by dissolving the analyte in a liquid crystal (LC) instead of a liquid solvent (isotropic system). The replacement of isotropic to oriented solvent is analytically highly valuable because we have access to ^{2}H residual quadrupolar couplings in solutes, leading to the presence on ^{2}H NMR spectra of quadrupolar doublets (QD) for each inequivalent deuteron instead of singlets. The range of splittings of each QD leads in turn to distribute the NAD signals on larger spectral windows, thus reducing the peak overlaps. Concomitantly the use of 2D-NMR experiments simplifies the analysis of entangled, overcrowded anisotropic NAD 1D-NMR spectra. As we will see, this approach makes it possible to distinguish several NAD signals from H in methyl groups of miliacin that resonated at the same (or very close) frequencies in isotropic solvents. This unusual strategy has been recently pioneered by Lesot and coworkers for analyzing the site-specific isotope ratios, $(D/H)_i$, in fatty acids and then triglycerides (Lesot et al., 2008, 2009, 2011, 2012; Serhan et al., 2010). We report here its first application for investigating biomarkers of interest such as miliacin.

2. ANISOTROPIC ^{2}H NMR SPECTROSCOPY

When a molecule in the spectrometer's magnetic field does not tumble isotropically, all orientation-dependent NMR interactions become visible on spectra. Thus ^{2}H NMR spectra of solutes in a mesophase contains new informative NMR interactions such as the chemical shift anisotropy (^{2}H CSA); ii) the dipolar interactions (D); iii) and the quadrupolar interaction ($I \geq 1$) (Sarfati et al., 2000; Lesot et al., 2015) (isotopically normal or enriched).

Assuming that ^{2}H CSA is generally negligible while the ^{2}H-^{1}H couplings can be removed by ^{1}H decoupling, the ^{2}H signal of aligned monodeuterated molecule is a simple QD (two transitions) centered on the anisotropic ^{2}H chemical shift, $\delta(^2H)^{aniso}$, (generally $\delta(^2H)^{aniso} \approx \delta(^2H)^{iso}$) (Lesot et al., 2009). The separation (in Hz) between components of doublet is referred to as ^{2}H quadrupolar splitting (or the residual quadrupolar coupling) and will be noted $\Delta\nu_{Q_i}(^2H)$ therein (**Fig. S-4, Appendix A.II**). For a C-D internuclear directions oriented in a weakly-aligning LC, the electric field gradient ($V_{C\text{-}D}$) is usually assumed to be axially symmetric along the C-D vector and the value of $\Delta\nu_{Qi}$ relative to the $\mathbf{B_o}$ can be expressed as (**Eq. 1**; Sarfati et al., 2000; Lesot et al., 2009):

$$\Delta\nu_Q = \frac{3}{2} C_Q^{D_i} \times S_{C-D_i} = \frac{3}{2}\left(\frac{eQ_{D_i}V_{C-D_i}}{h}\right) \times \left(\frac{1}{2}\left\langle 3\cos^2\left(\theta_{C-D_i}\right) - 1\right\rangle\right) \quad (1)$$

where $C_Q^{D_i}$ is the quadrupolar coupling constant at deuterium atom i, $S_{C\text{-}Di}$, is the order parameter of the C-D_i vector and θ_{C-D_i} is the angle between the $\mathbf{B_o}$ magnetic field axis and the C-D_i bond axis. The <⋯> denotes an ensemble average. As the trigonometric function in brackets varies between 1

(when θ_{C-D_i} // $\mathbf{B_o}$) and -0.5 (when $\theta_{C-D_i} \perp \mathbf{B_o}$), $\Delta\nu_{Qi}$ can be positive or negative. When θ_{C-D_i} is equal to 54.7° (the so-called 'magic angle'), $S_{C\text{-}Di}$ becomes null, and so the doublet collapses to a single line.

As the second stable isotope of hydrogen, deuterium atoms are naturally in all organic compounds, and hence they can be used as nuclear probe to study the site-specific natural isotope distribution of molecules. Even with low natural abundance of deuterons (around 1.55 10^{-2} % V-SMOW value; De Laeter et al., 2003), routine NMR spectrometers can detect quite easily (with reasonable experimental times) their NMR signals both in isotropic and anisotropic solvents, successful results depending also on the concentration of the analyte and its molecular weight (Lesot et al., 1998, 2009).

Proton-decoupled NAD NMR spectra of solutes dissolved in homogeneous mesophases consist of a sum of resolved independent quadrupolar doublets instead of single lines as in liquids, each of them corresponding to the various inequivalent ^{2}H sites in the molecule. For middle to large-sized molecules, the distribution of positions and the splitting of ^{2}H doublets over the spectral window can generate complex spectra (with possible overlaps of resonances), where the assignment of QD can be ambiguous. Various two and three-dimensional (2D/3D) NMR sequences belonging to NMR experiment class entitled "QUadrupole Ordered SpectroscopY" (QUOSY) have been specifically developed to achieve clearer analyses of overcrowded anisotropic NAD 1D spectra (Merlet et al, 1999; Lesot and Lafon, 2008b). All of them allow the spectral information (δ, $\Delta\nu_Q$) to be distributed on several (two or three) spectral dimensions (F_1, F_2 and F_3), thus facilitating the assignment of ^{2}H quadrupolar doublets.

3. ANALYTES, MATERIALS AND METHOD

3.1. Structure of miliacin and germanicol

Like other pentacyclic triterpenes, miliacin is formed from mevalonic acid *via* the mevalonate pathway through the cyclisation of squalene epoxide by an oxidosqualene cyclase (Xu et al., 2004; **see Section 6**). This scheme leads to the formation of germanicol, a triterpene alcohol constituted of five six-membered rings (**Fig. 1**) that bear eight methyl groups (Me23 to Me30). The final conversion of the alcohol function into a methyl ether function, which principally takes place in plants of the family Poaceae (Jacob et al., 2005), transforms germanicol into miliacin, by the addition of a methyl group (Me31) linked to the oxygen on the C-3 position. In miliacin, the nine methyl groups are numbered from Me23 to Me31 (**Fig. 1**). For simplicity, the five 6-membered rings of miliacin have been labeled A-E and contain two or three methyl groups each (ring A: Me23, Me24 and Me 25; ring C: Me26 and Me27; ring E: Me28, Me29 and Me30).

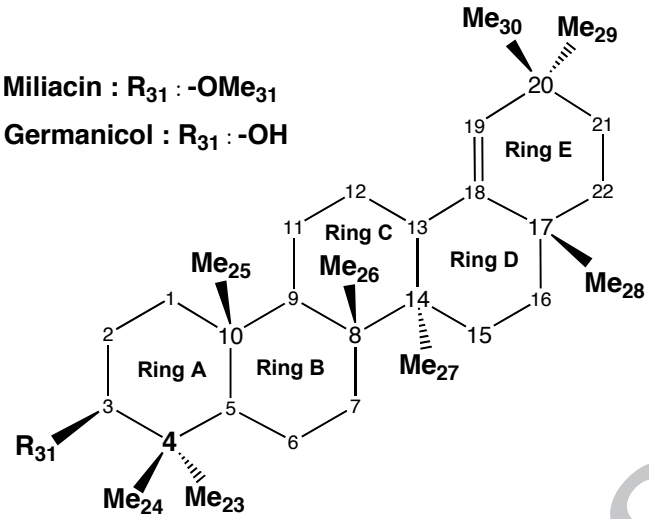


Figure 1: Structure of miliacin where R_{31} = -OMe and germanicol where R_{31} = -OH, together with the atomic numbering used and notation of rings (A to E) and methyl groups (23 to 31).

3.2. Miliacin and germanicol extraction and purification

Miliacin and germanicol were recovered from millet oil (provided by HITEX, Vannes, France) of > 98% purity as determined by Gas Chromatography/Mass Spectrometry (GC/MS). 6 g of miliacin were directly recovered from 750 ml of oil by repeated precipitation in cold acetone (Itô, 1934). From these 6 g of analyte, only about 100 mg of miliacin was necessary to prepare NMR samples used for NAD NMR analyses described in this work. Germanicol, which is far less abundant than miliacin in the oil, was recovered from a longer process of separating the oil into neutral, acidic and polar fractions by Solid Phase Extraction using aminopropyl bonded silica (Jacob et al., 2005). The neutral fraction was recovered by elution with CH_2Cl_2:isopropanol (9:1) and further separated on activated silica (120 °C during 24 h) deactivated with 5% weight water. After elution of aliphatic hydrocarbons, aromatics, esters and ether, and ketones, the alcohol fraction that contains germanicol was eluted with 2 ml of hexane:ethyl acetate (85:15) and then 2 ml of hexane:ethyl acetate (80:20) as described by Jacob et al. (2008b). To eliminate the large amounts of sitosterol that coelutes with germanicol, we then separated the alcohol fraction on activated silica by splitting the alcohol fraction into sub-fractions of 500 μl of hexane:ethyl acetate (85:15). After GC/MS analysis of an acetylated aliquot of these sub-fractions, we combined fractions containing germanicol.

3.3. Miliacin and germanicol δD values by GC-irMS

Germanicol was acetylated into germanicol acetate with acetic anhydride and pyridine (1:1) during 1h at 80°C prior to its GC-irMS analysis. Acetic anhydride with an offline-determined δD value of -133 ± 2 ‰ was acquired from A. Schimmelmann (Indiana University, USA). After drying under nitrogen, the fraction was dissolved in isooctane prior to analysis by Gas Chromatography-isotope ratio Mass Spectrometry (GC-irMS) in order to determine its δD value. Miliacin from the 6 g

in-house stock was also dissolved in isooctane prior to GC-irMS analysis.

δD values were determined on a TraceGC chromatograph equipped with a Triplus Autosampler coupled to a DeltaV Advantage isotope ratio mass spectrometer through a GC-Isolink interface and a Conflo IV dilution system (all from Thermo Scientific, Bremen, Germany). The gas chromatograph was fitted with a TG5-MS capillary column (60 m, 0.25 mm i.d., 0.25 μm film thickness; ThermoScientific). The temperature of the column was held at 80°C for 1 min, and then increased from 80 to 300°C at 20°C.min^{-1}, with a final isothermal hold at 300°C for 30 min. Precision and accuracy were determined by using a mixture of 15 *n*-alkanes (*n*-C_{16} to *n*-C_{30}) with δD values ranging from -9 to -254 ‰ determined off-line (A5 mixture, A. Schimmelmann, Indiana University, USA) and allowed us to normalize the δD values of miliacin ($\delta D_{miliacin}$) and of germanicol ($\delta D_{germanicol}$) to the V-SMOW isotopic scale. The measured δD values for the standard *n*-alkanes are in good agreement with those measured off-line ($r^2 = 0.987$). The overall precision for the n-alkanes standards was ≤ 4‰. Samples were injected at least three times and were accurate to ± 5 ‰ for miliacin and germanicol. The δD value of germanicol acetate was mathematically corrected from the contribution of acetic anhydride to obtain the $\delta D_{germanicol}$ value. Our results are shown as both (D/H) ratios in ppm and δD values in ‰, calculated using a (D/H) value of V-SMOW of 155.76 ± 0.05 ppm, and converting from ppm values to ‰ values using Equation 2:

$$\delta D = \left(\frac{(D/H)_i}{(D/H)_{V\text{-}SMOW}} - 1 \right) \times 1000 \quad (2)$$

3.4. Isotropic ^{1}H and ^{13}C NMR of miliacin

The full assignment of the proton and carbon-13 signals recorded in liquid solvents ($CDCl_3$) was achieved using homonuclear ^{1}H-^{1}H 2D COSY, heteronuclear ^{13}C-^{1}H HMBC and HMQC 2D experiments combined with the analysis of ^{13}C-^{13}C ADEQUATE 2D spectrum. Several isotropic NMR spectra of miliacin are given in **Figs. S-1**, **S-2** and **S-3** in **Appendix A.II**. All isotropic NMR data were recorded on a Bruker 9.4 T (400 MHz) Advance I spectrometer equipped with 5 mm o.d. BBI probe.

3.5. Preparation of polypeptide oriented samples

The ability to solubilize a given analyte but also to form a polypeptidic homogeneous liquid-crystalline phase is an important factor that depends on what organic co-solvent is chosen. We were best able to solubilize miliacin using 30 mg of miliacin, 80 mg of poly-γ-benzyl-*L*-glutamate (PBLG) with a degree of polymerization of 743 (MW (viscosity) = 162 900 g/mol) and 810 mg of chloroform, corresponding to 8.7% w/w (PBLG) and 3.2 % w/w (miliacin). Over this latter mass ratio, miliacin is not totally dissolved in the mesophase, and solid particles are visible, even at higher temperatures. Other polar and apolar organic solvents that we experimented and that proved less

satisfactory include CH_2Cl_2, tetrachloroethane, THF, dioxane, pyridine. For all of them, the solubility of miliacin was found to be extremely low. Note that even DMSO (incompatible with PBLG) is not suitable.

As the polypeptide is chiral, oriented solutions of PBLG is able to discriminate between NMR signals of enantiomers of wide range of organic molecules but also enantiotopic directions in prochiral compounds (Lesot et al., 2000, 2009). Nevertheless, in this work, PBLG was selected not for its enantiodiscrimining properties, which are useless for miliacin (enantiopure biocompound), but for its ability to homogeneously orient a guest molecule with a rather low degree of alignment (Lesot et al., 2000).

Due to the high molecular weight (MW = 440.7 g/mol) and the small amount of miliacin that can be solubilized in the PBLG/$CHCl_3$ mesophase (only 30 mg), the molar amount in monodeuterated isotopomers is equal to 3.17×10^{-8} mol for methyl groups and 1.06×10^{-8} mol for methines, if (D/H) = 1.55×10^{-2}%. Assuming that density of the mixture is equal to 1, the concentration of isotopomers in the sample at methyl and methine sites is evaluated at about 3.4 µmol/l and 1.14 µmol/l, respectively. This isotopomeric concentration is weak (30 fold less) compared to the value of chloroform. However recording anisotropic NAD spectra of solute with rather short recycling delays (T_R = 1 sec) allows to significantly reduce the intensity of ^2H signal of $CDCl_3$ (and avoid a bad digitization of solute signals) whereas quantitative conditions ($T_R = 5T_1(^2H)$) were maintained to record the ^2H signals of miliacin (Lesot et al., 2008).

3.6. Anisotropic NAD 2D-NMR in LC

All anisotropic NAD 2D-NMR spectra were recorded with a 14.1 T Bruker Advance II spectrometer (92.1 MHz for deuterium nuclei) equipped with a 5 mm o.d. selective ^2H cryogenic probe. NAD *Q*-resolved Fz 2D experiments combined with proton Waltz-16 decoupling sequence was chosen to analyze of NAD NMR spectra of miliacin dissolved in PBLG/$CHCl_3$ (Lafon et al., 2009). In this experiment, the ^2H chemical shifts are refocused in the F_1 dimension. After a double FT, the 2D map can be tilted (see the principle on **Figure S-5** in **Appendix A.II**), and then symmetrized, thereby spreading the ^2H chemical shifts, $\delta(^2H)$, and quadrupolar splittings, $\Delta\nu_Q(^2H)$ information in the F_2 (horizontal axis) and F_1 (vertical axis) dimensions, respectively. Thus a series of ^2H singlets are displayed in F_2 whilst a series of independent ^2H doublets are observed in F_1.

The anisotropic *Q*-resolved Fz 2D spectra were recorded using a 2D matrix of 2652 (t_1) × 300 (t_2) data points, then zerofilled to 4096 (t_1) × 1024 (t_2) data points *prior* the double FT. 256 scans were added at each t_1 increment. A lorenzian filtering was applied in both dimensions (LB = 1.5 Hz). The recycling delay (T_R) of anisotropic NAD 2D experiments was set to 1 sec, leading to a total acquisition time of 24 hours. Such a delay is large enough to record the anisotropic NAD spectrum

of miliacin with quantitative conditions and reliable evaluation of peak integrals (on 1D spectra) or 2D volume on 2D spectra of miliacin is then possible (Lesot et al., 2008). "Integral of two components of NAD doublets measured on 1D traces extracted to 2D maps (see below and sup info) were determined by peak deconvolution (Top-spin 3.0 software)". To improve the accuracy of the measure of (D/H) ratios and evaluate the standard deviation, six identical NAD 2D-NMR experiments on the same NMR sample of miliacin were recorded and analysed.

3.7. SNIF-NMR™ protocol and experimental details

According to the SNIF-NMR protocol, isotropic NAD 1D-NMR spectra of analytes are recorded under quantitative conditions ($T_R = 5T_1$) with internal isotopic reference whose D/H ratio is calibrated. If the mass of compounds (analyte and reference) is accurately weighted, the site-specific $(D/H)_i$ ratios (expressed in ppm) can be calculated from evaluation of peaks surface on NMR spectra as follows (**Eq. 3**; Lesot et al., 2011; Martin et al., 1986):

$$\left(D/H\right)_i^{anal} = \left(\frac{A_i^{anal}}{A_i^{ref}}\right) \times \left(\frac{P_i^{ref} \times m^{ref} \times M^{anal}}{P_i^{anal} \times m^{anal} \times M^{ref}}\right) \times \left(D/H\right)_i^{ref} \qquad (3)$$

In **Eq. 3**, A_i^{anal} and A_i^{ref} are the areas of the signals at site i for the analyte and the reference, P_i^{anal} and P_i^{ref} are their stoichiometric numbers of hydrogen at site *i*, M^{anal}, M^{ref}, m^{anal} and m^{ref} are the molecular weights and masses of the analyte and the reference, respectively. Finally, $\left(D/H\right)_i^{ref}$ is the isotope ratio of reference (at site i).

From the experimental conditions viewpoint, the SNIF-NMR 1D spectrum of miliacin was recorded in chloroform in the presence of isotopically calibrated tetramethyl urea (TMU) as internal chemical reference ($(D/H)_{TMU}$ = 84.4 ± 2 ppm, Lot 0116XAC from Eurofins). The temperature of 305 K used provided a good compromise between a reasonable experimental time for the NAD 1D experiments (< 15 h) and a full solubilization of miliacin. At 305 K, the $T_1(^2H)$ of TMU was measured at 1.17 sec (average of three NAD inversion-recovery 1D-NMR experiments), leading to a recycling delay, T_R, of 6 sec (with 3.5 sec of acquisition time) and in turns to a total acquisition time of 15 hours for 9000 scans added. Six experiments were recorded with 6k data points (zerofilled to 16K), filtered by an exponential function (1 Hz) prior FT and peak integral for Me31 and TMU signal were quantified by peak deconvolution after a baseline correction.

The mixture was prepared in a small flask tube by adding 39.9 mg of miliacin and 18.6 mg of TMU (MW = 116.16 $g.mol^{-1}$) and 1.640 g of dry $CHCl_3$. The flask was then actively mixed (by strongly shaking) prior to a part of mixture (≈ 700 mg; this amount is optimized for the length of the coil in cryogenic probe) was transferred into the 5-mm NMR tube. The tube has been fire-sealed to avoid long-term solvent evaporation. Due to the low solubility of miliacin in $CHCl_3$, the sample was

prepared with larger amounts of components in order to reduce the error on the weightings of small masses. Also, the mass of TMU used (twelve deuterium equivalent sites) was calculated to reach a ratio in peak intensity of TMU and Me31 (three deuterium equivalent sites) equal to about 4, thus optimizing the S/N ratio of TMU.

4. RESULTS

4.1. Spectral assignment of isotropic ^{1}H and ^{13}C resonances

In order to avoid any misinterpretation of the isotopic profile of methyl groups on miliacin, we re-examined the assignment of ^{1}H signals in combination with ^{13}C signals before analyzing NAD 2D spectra in oriented solvent and then evaluating the site-specific isotope distribution. **Fig. 2** gives the 600 MHz ^{1}H spectrum of miliacin dissolved in chloroform at 295 K. Due to the presence of several stereogenic centers (asymmetric carbons), the geminal protons in CH_2 groups and the *gem*-dimethyl groups are diastereotopic, and hence yield distinct ^{1}H (and also ^{2}H) signals. The proton-decoupled carbon-13 1D spectrum (^{13}C-{^{1}H}) as well as $^{13}C/^{1}H$ heteronuclear HSQC and HMBC 2D spectra of miliacin dissolved in chloroform at 295 K and recorded at 9.4 T (100 MHz) (**Fig. S-1**) (**Figs. S-2** and **S-3**) are presented in **Appendix A.II**. The combined analysis of all spectral data allows an unambiguous assignment of all ^{13}C and ^{1}H resonances. The full assignment of ^{13}C and ^{1}H signals is listed in **Tables S-1**, **S-2** and **S-3** in **Appendix A.I**. Note here that our assignment agrees with Smetanina et al. (2001), but the two ambiguities pointed out by these authors have been eliminated (**Fig. S-5** in **Appendix A.II**).

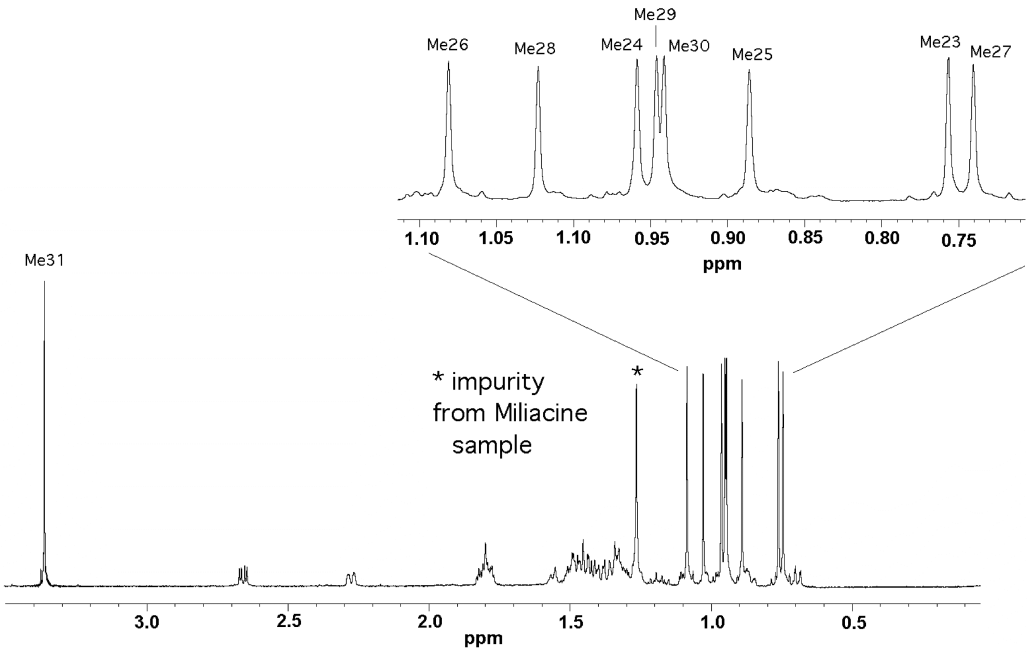


Figure 2: 600.1 MHz proton 1D-NMR spectrum of miliacin in $CDCl_3$ (295 K). The assignment of resonances of methyl groups is given in **Table S-1**, **S-2** and **S-3** in **Appendix A.I** of **Supp. Info.**). The peak marked by an asterisk (δ = 1.24 ppm) is an unidentified impurity.

4.2. Analysis of anisotropic NAD signals of methyl groups

Miliacin is a very illustrative example of the analytical performance of anisotropic NAD 2D-NMR to determine the intramolecular site-specific isotope distribution. Indeed, on the isotropic NAD 1D-NMR spectrum of miliacin, the resonances of methyl groups, Me24, Me29 and Me30 as well as Me23 and Me27 are not spectrally resolved (even at 92.1 MHz) (see **Fig. S-4** in **Appendix A.I of Supp. Info.**). As a consequence, only two broad ^{2}H resonances are observed, thus preventing a specific determination of their respective peak areas, and consequently the associated (D/H) ratios at each individual site.

Fig. 3a shows the shielded region of methyl groups (Me23- Me30) of the tilted *Q*-resolved Fz 2D map of miliacin recorded in PBLG/$CHCl_3$ at 295 K (see **Fig. S-7, Appendix A.II of Supp. Info.** for the full map). As expected, all methyl signals are now separated on the basis of their d their $\delta(^2H)$ and $\Delta\nu_Q(^2H)$. The sixteen ^{2}H resonances (8 quadrupolar doublets) detected on the 2D map correspond to eight inequivalent methyl sites, Me23-Me30; each of them giving one integrated signal of three ^{2}H equivalent sites. As expected, almost all QDs of methylene and methine ethylene groups of miliacin do not clearly emerge from the spectral noise (baseline) and cannot be used reliably for quantitative purposes. In contrast, the doublets of all methyl groups are detected on 2D map with S/N ratios varying between 8 (Me30) and 25 (Me26) whereas a value of 35 is found for the methyl of methoxy group (Me31) (**Table 1**). Due to the small amount of miliacin solubilized in the mesophase and its high MW, the S/N ratios are unusually low. However a reliable (acceptable) evaluation of peak areas leading to reliable (D/H) ratios is still possible.

Table 1: Spectral data associated to anisotropic NAD signals of methyl groups of miliacin at 295 K.

Methyl group	Ring	$\delta(^2H)^a$ (ppm)	$\Delta_{1/2}$ (Hz)	S/N	$\lvert\Delta\nu_Q\rvert$ (Hz)
Me_{27}	C	0.771 (0.746)	2.4	24	309
Me_{23}	A	0.774 (0.763)	2.8	22	36
Me_{25}	A	0.905 (0.891)	3.3	12	275
Me_{30}	E	0.978 (0.948)	3.6	8	272
Me_{29}	E	0.999 (0.952)	5.6	9	152
Me_{24}	A	0.982 (0.965)	3.3	9	296
Me_{28}	E	1.057 (1.029)	3.3	17	292
Me_{26}	C	1.105 (1.086)	2.3	25	314
Me_{31}	A	3.399 (3.368)	2.1	35	193
$CHCl_3$	–	7.270 (7.270)	1.8	3000[(b)]	273

[(a)] $\delta(^2H)$ measured on the F_2 projection of anisotropic NAD 2D spectrum ($\delta_{chloroform}$ = 7.27 ppm). In parenthesis are reported the $\delta(^1H)$ measured on the isotropic ^{1}H 1D-NMR spectrum ($\delta_{chloroform}$ = 7.27 ppm).
[(b)] The large difference between the S/N ratios of Me groups and $CHCl_3$ originates from the difference of deuterated isotopomer concentrations and the linewidths of ^{2}H signals.

On the tilted 2D map, it becomes possible to extract a single or several columns (projection of columns) associated to a single or several ^{2}H site(s) (Lesot et al., 2008). Four NAD 1D sub-spectra (containing the signal associated to one to four ^{2}H sites) were selected (**Fig. 4**) in order to separate reliably all quadrupolar doublets (regions selected are shown in **Fig. S-8** in **Appendix A.II. of Supp. Info.**) **Table 1** lists the ^{2}H quadrupolar splittings in Hz (distance in Hz between two components of each ^{2}H doublet). As seen, the $\Delta\nu_Q$'s (in absolute value) varies from 36 to 314 Hz, but the majority of doublets (6 / 9) range from 272 to 314 Hz. As expected by using 2D NMR approach, the NAD signals associated to each methyl group of miliacin are spectrally distributed (on the 2D map) on the basis of their ^{2}H chemical shifts and quadrupolar splittings. Hence it becomes possible to evaluate the areas of each doublet by peak deconvolution. For the same molecule, these areas are proportional to the number of monodeuterated isotopomers contributing to the NAD signal (three for a methyl group) and the $\left(D/H\right)_i^{Me}$ ratio at methyl site i (results, discussion in **Section 4.4**).

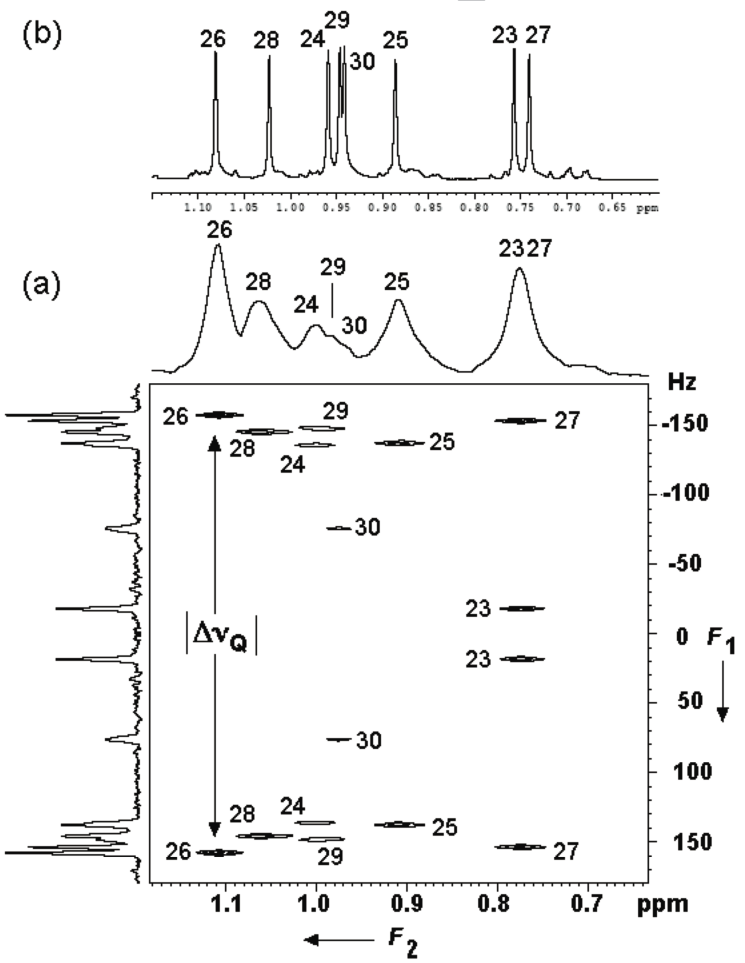


Figure 3: (a) Methyl region (Me31 not shown) of the 92.1 MHz NAD tilted *Q*-resolved Fz 2D map of miliacin dissolved in PBLG/$CHCl_3$ and recorded at 295 K. (b) 600 MHz proton 1D spectrum of miliacin in $CHCl_3$ plot at the horizontal same scale (in ppm). A lorenzian filtering (LB = 1.5 Hz) has been applied in both spectral dimensions. On the F_1 and F_2 axes are shown the spectral 2D projections. Note the significant variations of peak intensity of ^{2}H doublets in F_1 dimension illustrating the enrichment/depletion in deuterons.

As far as possible, the assignment of ^{2}H resonances in F_2 dimension on the *Q*-resolved Fz map of miliacin can be derived from the position of ^{1}H lines of methyl groups observed the isotropic ^{1}H spectrum (see **Fig. 3**). However when ambiguities exist (as for sites 23 and 27), further arguments based on the orientation of the C-CH_3 axis (see **Fig. 5**) can help to assign the quadrupolar doublets on the basis of their magnitude. Indeed as the quadrupolar splittings are proportional to the order parameter, $S_{C\text{-}D}$, of the C-D bond (see **Eq. 1**), they can be related to the relative positions of associated C-D vectors in a (rigid) molecular structure (Aroulanda et al., 2001; Lesot et al., 2015). We can therefore correlate the magnitude of ^{2}H doublets to geometrical data, thus providing complementary information for their assignment. This approach was typically useful to assign the quadrupolar doublets associated to Me23 and Me27, for instance. Detailed explanations and further discussions on the assignment of doublets are given in **Appendix A.III** of **Supp. Info.**

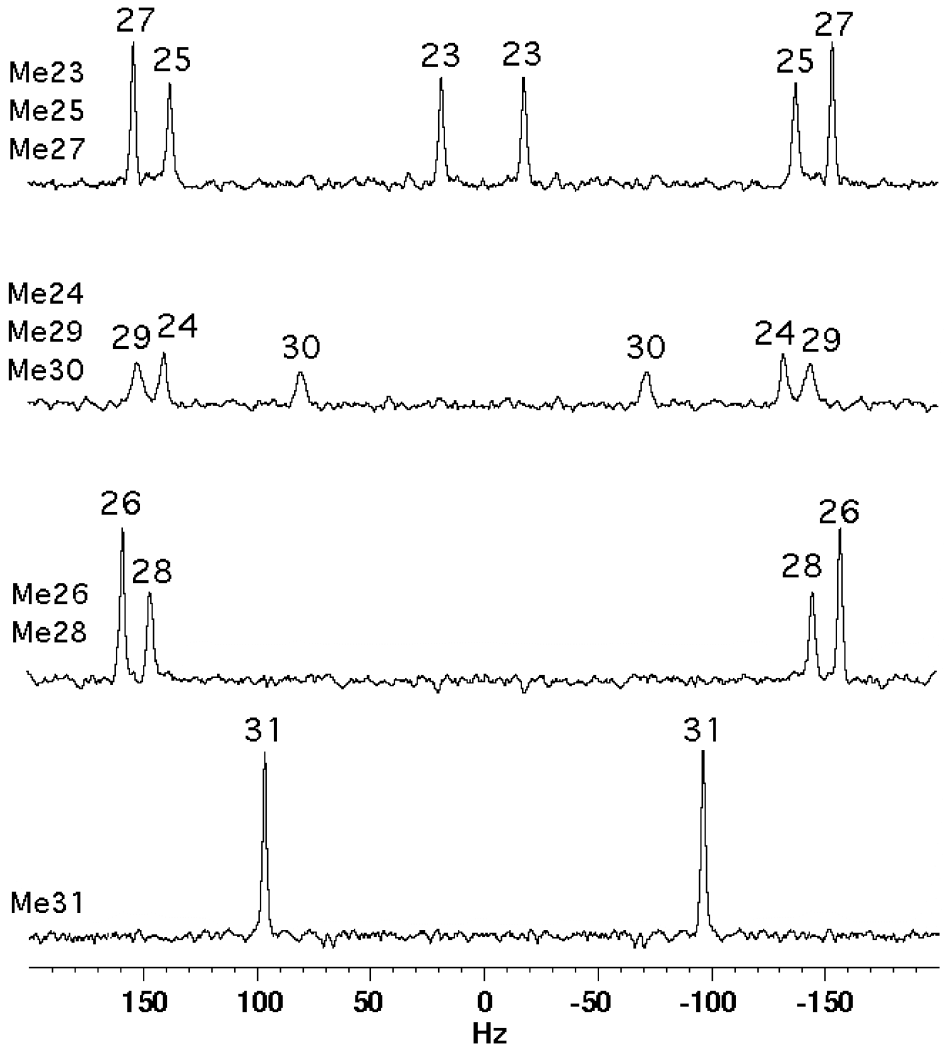


Figure 4: Series of four NAD 1D sub-spectra extracted from the tilted *Q*-resolved Fz 2D map showing the nine methyl groups of miliacin along with their assignments (see also **Fig. S-9**). Each sub-spectrum is plotted with the same amplitude of noise. Note the variation of peak intensity of ^{2}H doublets. The variations observed reveal the differences of (D/H) for each site.

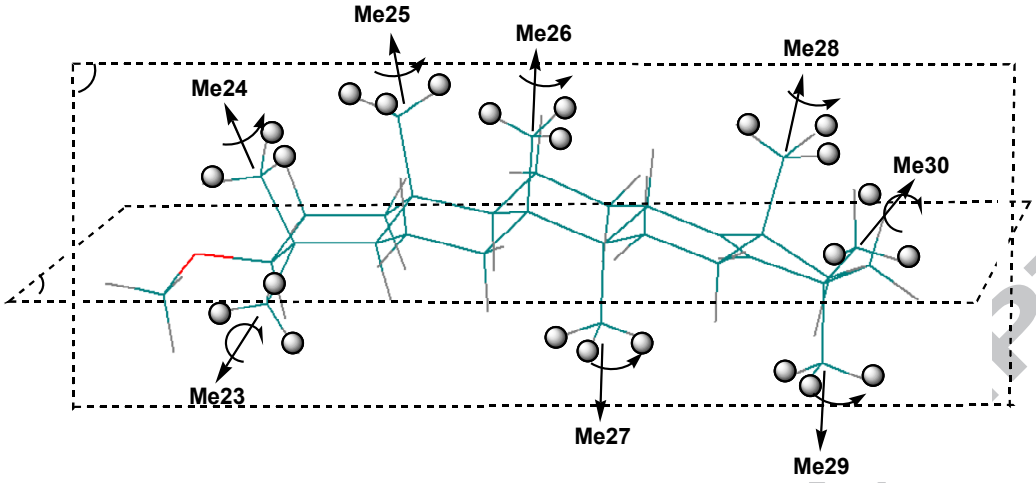


Figure 5: 3D view of the minimal energy conformation of miliacin calculated with Hyperchem 3.0 software. A semi-empirical quantum chemistry methods based on the Hartree–Fock formalism (Austin Model 1 or AM1) has been used for modeling the molecule. The minimal energy was found at -8177.99 kJ/mol. Except for ring E distorted by the double bond, all other cycles have adopted a chair conformation. Note the difference of orientation axes of C-CH_3 vectors for Me23 and Me30 (close to the median horizontal molecular plan) compared to Me24 to Me29 (see also **Fig. S-9** in **Appendix A.II** of **Supp. Info.**).

4.3. Absolute (D/H) and δD values of methyl groups

In contrast to their ^{1}H signals (**Fig. 2**), significant variations in the intensity of ^{2}H doublets for the nine methyl groups are observed on the 2D map or the associated F_1 and F_2 projections (**Figs. 3** and **4**). As NAD 2D-NMR spectra have been recorded with quantitative experimental conditions, the peak areas on projections (or peak volume on the 2D map) are in principle proportional to the number of isotopomers contributing to the NAD signal (three for a methyl group) and the isotope ratio (D/H) at each methyl site.

4.3.1. Evaluation of relative ^{2}H peak surface by anisotropic NAD 2D-NMR

When the direct quantification of peak surfaces on anisotropic 2D spectra (using an isotopically calibrated, internal chemical reference) is experimentally difficult as in case of miliacin (problem with mesophase, interference between the ^{2}H solute and reference signals), other strategies must be proposed. Thus the use of an electronic reference signal (ERETIC method) is a possible alternative but required special equipment not implemented in the majority of spectrometers (Akoka et al., 1999). For the study of miliacin, we have adopted a two-step protocol for quantification of relative D concentrations in methyl groups (or any other sites) consisting into: i) evaluating the relative variation of ^{2}H surfaces of each doublet using the NAD signal of methyl group of methoxy (Me31) as internal reference value; ii) determining the (D/H) and δD values of Me31 using either GC-irMS (see **Section 4.3.2**) or isotropic quantitative NAD NMR in presence of TMU (see **Section 4.3.3**). Note that the choice of the Me31 as internal reference is advantageous for both methods. From GC-irMS viewpoint, the $(D/H)_{Me31}$ ratio can be easily quantified by comparing the global δD of miliacin and the one of germanicol. From an NMR viewpoint, the isotropic ^{2}H signal of Me31 is

443 isolated from other signals of miliacin, reducing errors on the measure of peak surface by
444 deconvolution, and does not interfere with the signal of TMU.

445 The measurement of relative surface of peaks can be either directly performed on the 2D map
446 or using 1D projections of selected slices extracted from the map (Massou et al., 2007; Lesot et al.,
447 2008). Adopting this second approach, the evaluation of relative surface of doublet of methyl groups
448 was performed as following: i) selection of 10 columns centered on the NAD signal of each
449 inequivalent ^{2}H site and calculation of their respective F_1 projection, ii) addition of NAD signal of
450 Me31 (10 columns) to each NAD 1D sub-spectrum; iii) determination (by signal deconvolution) of
451 the peak surface for each methyl group, iv) calculation of the average value obtained for the six
452 different NAD 2D spectra recorded with the same experimental conditions (acquisition and
453 processing), assuming a reference value for the Me31 signal. In order to determine the reliability of
454 this approach, we have compared the peak area of Me31 for the six experiments. The variation
455 between each experiment is below 2%. In column 3 of **Table 2** are listed the averaged surfaces
456 measured for the six NMR experiments, when the value for the Me31 is arbitrary set at 100. Values
457 for each experiment are given in **Table S-4** and plotted in **Fig. S-10** whereas **Fig. S-11** in **Appendix**
458 **A.II** of **Supp. Info.** displays the associated average values together with the SD.
459
460 ***4.3.2. δD_{Me31} determined by GC-irMS and SNIF-NMR™***

461 δD_{Me31} was first estimated by comparing the GC-irMS δD values of miliacin ($\delta D_{miliacin}$) and
462 of germanicol ($\delta D_{germanicol}$) extracted from millet oil according to the following rationale and
463 equations. We assumed no other hydrogen isotopic difference between miliacin and germanicol (see
464 **Fig. 1**) than the addition of three hydrogen atoms from the Me31.

465 Experimentally, we obtained $\delta D_{miliacin}$ = -100 ± 5 ‰ and $\delta D_{germanicolacetate}$ = -110 ± 5 ‰, which
466 gives $\delta D_{germanicol}$ = -108.6 ± 5 ‰, after mathematically removing the contribution from acetic
467 anhydride (δD_{acanh} = -133 ± 2 ‰) according to **Eq. 4**:
468

469 $$\delta D_{germanicol} = \left(\frac{52 \times \delta D_{germanicolacetate} - 3 \times \delta D_{acanh}}{49} \right). \quad (4)$$

470 The error on $\delta D_{germanicol}$ was classically calculated as $\sqrt{(52/49)^2 \times (\delta D_{germanicolacetate})^2 + (3/49)^2 \times (\delta D_{acanh})^2}$, thus
471 taking into account the multiplicative factors of **Eq. 4** and assuming that the correlative term is null
472 (independent variables). Note here that this $\delta D_{germanicol}$ value does not take into account the H atom of
473 the alcohol function that is highly exchangeable and was lost during acetylation. Hence, when
474 calculating δD_{Me31} in mass balance equations **Eq. 4** and **Eq. 5**, we consider 49 H atoms in germanicol
475 (instead of 50) and 52 H atoms in miliacin.
476

$$\delta D_{Me31} = \left(\frac{52 \times \delta D_{miliacin} - 49 \times \delta D_{germanicol}}{3} \right). \quad (5)$$

The calculation in **Eq. 5** gives δD_{Me31} equal to +41 ‰ ± 119 ‰. Propagation of uncertainties on the δD_{Me31} was calculated as for **Eq. 4** (see **Table 2**, footnote i). Using **Eq. 2**, this corresponds to a $(D/H)_{Me31}^{GC-irMS}$ value of 162 ± 19 ppm. Details for the conversion from δD to (D/H) are given in footnote g of **Table 2**.

A possible experimental alternative to GC-irMS consists to determine δD_{Me31} value applying the so-called SNIF-NMR™ protocol using NAD NMR in isotropic solvent in presence of calibrated TMU (see **Section 3.7**). The analysis of experiments recorded indicates that standard deviation on the area ratio ($R = A_{Me31}/A_{TMU}$) was found to be equal to 0.28 (e.g. R = 25.2 ± 0.6), thus indicating a very good repeatability of isotropic NAD 1D NMR experiments (see Table S-5). Knowing the $(D/H)_{TMU}$ and the masses of TMU and miliacin (purity of 98%) used, the $(D/H)_{Me31}$ was obtained using **Eq. 3**; the average value was found to be equal to 154 ppm. The uncertainty value on $(D/H)_{Me31}$ was classically calculated as $\Delta X = X \times \sqrt{\sum_i (\Delta g_i / g_i)^2}$ where g_i are the key variables of **Eq. 3** (see **Table 2** and **S-5**) and found to be equal to ± 10 ppm. Expressed in δD_{Me31} (from the **Eq. 2**), the value is equal to -11 ± 64 ‰ (see **Table 2**, footnote e).

Conclusively, it is worthwhile noting that, although they were determined by two very distinct analytical methods, the $(D/H)_{Me31}$ / (δD_{Me31}) values are consistent within the uncertainty range (162 ± 19 ppm *versus* 154 ± 10 ppm (+41 ± 119 ‰ *versus* -11 ± 64 ‰). The difference on the nominal $(D/H)_{Me31}$ value (relative discrepancy of 5%) could be explained by the lower sensitivity of NAD NMR spectroscopy (compared to GC-irMS) in particular when samples cannot be enough concentrated in analyte.

4.3.3. Absolute (D/H) and δD values of methyl groups 23 to 30

From $(D/H)_{Me31}$ value determined by the SNIF-NMR™ protocol or GC-irMS, it becomes possible to evaluate the (D/H) value for each methyl group $(D/H)_{Mei}$ and corresponding δD_{Mei} with i = 23 to 30) in miliacin, using relative (D/H) values previously determined using the anisotropic NAD NMR experiment (**Section 4.3.1**). All data and details of the calculation are given in **Table 2** and footnotes. A graphical comparison of the $(D/H)_{Mei}$ and δD_{Mei} variations *versus* the methyl groups is proposed in **Figs 6a** and **6b**. Large variations of $(D/H)_{Mei}$ values (up to a factor ≈ 2) exist from one site to another one. Except for Me31, the $(D/H)_{Mei}$ values are significantly below the V-SMOW value. This deuterium depletion in lipids produced by autotrophs is expected from biosynthetic fractionation (Sessions et al., 1999). Besides for Me31, Me26 and Me27 appear to be the less depleted methyl

511 **Table 2:** Values of $(D/H)_{Mei}$ (in ppm) and δD_{Mei} (in ‰) of the methyl groups of miliacin.

			Anisotropic NMR / SNIF-NMR™		Anisotropic NMR / GC-irMS	
Ring	**Methyl[a] group**	**Peak area[b]** no unit	**$(D/H)_{Me_i}$[c]** ppm	**δD_{Mei}[e]** ‰	**$(D/H)_{Me_i}$[f]** ppm	**δD_{Mei}[h]** ‰
A	***Me_{31}***	**100 (0.0)**	**154 ± 10[d]**	**-11 ± 64**	**162 ± 19[g]**	**+41 ± 119[j]**
A	Me_{23}	74.1 ± 13.8 (6.9)	114 ± 23	-268 ± 148	120 ± 25	-229 ± 162
	Me_{24}	54.6 ± 17.6 (8.8)	84 ± 28	-461 ± 180	88 ± 30	-432 ± 193
	Me_{25}	73.5 ± 14.8 (7.4)	113 ± 24	-274 ± 154	119 ± 28	-235 ± 177
C	Me_{26}	90.3 ± 10.2 (5.1)	139 ± 18	-107 ± 122	146 ± 23	-60 ± 151
	Me_{27}	82.4 ± 9.6 (4.8)	127 ± 17	-186 ± 116	133 ± 22	-143 ± 139
E	Me_{28}	76.8 ± 14.2 (7.1)	118 ± 23	-242 ± 154	124 ± 27	-201 ± 173
	Me_{29}	66.8 ± 12.0 (6.0)	103 ± 20	-340 ± 128	108 ± 23	-305 ± 148
	Me_{30}	41.5 ± 10.4 (5.4)	64 ± 17	-590 ± 109	67 ± 18	-568 ± 116

512 a) The methyl group Me31 is used as internal reference.

513 b) Average value of peak areas of NAD signals of methyl groups measured on anisotropic NAD 2D spectra (six NMR
514 experiments) when the methoxy group is arbitrary calibrated at 100. In parenthesis is given the value of the simple
515 standard deviation (σ_n) on peak areas calculated as: $\sigma_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ where n = 6. The uncertainty, noted ΔX, on value X (such

516 as X = x ± ΔX) is calculated (in ppm) as $\Delta X = k \times \sigma_n$ where k = 2, namely leading to a confidence interval of 95% [x -
517 1.96σ; x + 1.96σ].
518
519 c) The $(D/H)_{Me23\text{-}Me30}$ values in ppm (methyl groups 23 to 30) are calculated as $(D/H)_{Me23-Me30} = (\%A^{Aniso\ NMR}_{Me23-Me30}) \times (D/H)^{SNIF-NMR}_{Me31}$,
520 where $(D/H)^{NAD\ NMR}_{Me31}$ is determined by isotropic NAD 1D-NMR (*i.e.* $(D/H)^{SNIF-NMR}_{Me31}$ = 154 ± 10 ppm, see Table S-4). The
521 uncertainty on $(D/H)_{Me23\text{-}Me30}$ values is evaluated (in ppm) as: $\Delta(D/H)_i = (D/H)_i \times \sqrt{\left(\frac{\Delta(Area)_i}{Area_i}\right)^2 + \left(\frac{\Delta(D/H)^{SNIF-NMR}_{Me31}}{(D/H)^{SNIF-NMR}_{Me31}}\right)^2}$.

522 d) The $(D/H)^{SNIF-NMR}_{Me31}$ value in ‰ (methyl group 31) is experimentally determined by SNIF-NMR™ protocol in isotropic
523 solvent. See Table S-5 for the calculation of uncertainty on $(D/H)^{NAD\ NMR}_{Me31}$.

524 e) The $\delta D_{Me23\text{-}Me31}$ values in ‰ are calculated using **Eq. 2**. Considering that the $(D/H)_{V\text{-}SMOW}$ value is a constant
525 (international standard value) and not a variable, the uncertainty on $\delta D_{Me23\text{-}Me31}$ values is simply evaluated as
526 $\Delta\delta_{Mei} = \sqrt{\left(1000 \times \frac{\Delta(D/H)_{Mei}}{(D/H)_{V-SMOW}}\right)^2}$, assuming that correlation term is null.

527 f) The $(D/H)_{Me23\text{-}Me30}$ values in ppm are calculated as $(D/H)_{Me23-Me30} = (\%A^{Aniso\ NMR}_{Me23-Me30}) \times (D/H)^{GC-irMS}_{Me31}$, namely from the relative peak
528 surfaces (%) measured on anisotropic NMR spectra (see Table S-4) and using the $(D/H)^{GC-irMS}_{Me31}$ value (ppm) as reference.
529 The uncertainty on $(D/H)_{Me23\text{-}Me30}$ is evaluated (in ppm) as: $\Delta(D/H)_i = (D/H)_i \times \sqrt{\left(\frac{\Delta(Area)_i}{Area_i}\right)^2 + \left(\frac{\Delta(D/H)^{GC-irMS}_{Me31}}{(D/H)^{GC-irMS}_{Me31}}\right)^2}$.

530 g) The $(D/H)^{GC-irMS}_{Me31}$ value in ppm is calculated as $(D/H)_{Me31} = \frac{\delta D^{GC-irMS}_{Me31} \times (D/H)_{V-SMOW}}{1000} + (D/H)_{V-SMOW}$. From the same reason evoked in
531 footnote e, the uncertainty on $(D/H)^{GC-irMS}_{Me31}$ is evaluated as $\Delta(D/H)_{Mei} = \sqrt{\left(\frac{(D/H)_{V-SMOW} \times \Delta\delta_{Mei}}{1000}\right)^2}$, leading to a value of ± 19 ppm.

h) The $\delta D_{Me23\text{-}Me30}$ values in ‰ are calculated as $\delta D_{Me23-Me30} = (\%A^{Aniso\ NMR}_{Me23-30}) \times \frac{(D/H)^{Aniso\ NMR}_{Me23-Me30}}{(D/H)_{V\text{-}SMOW}}$. Note that these values cannot directly be obtained from $\%A^{Aniso\ NMR}_{Me23-30}$ and $\delta D^{GC-irMS}_{Me31}$ values due to the non-linearity of **Eq. 2**. For the same reason evoked in footnote e, the uncertainty on δD_{Mei} is -simply evaluated as $\Delta\delta_{Mei} = 1000 \times \frac{\Delta(D/H)_{Mei}}{(D/H)_{V\text{-}SMOW}}$.

i) The $\delta D^{GC-irMS}_{Me31}$ value in ‰ was experimentally determined by GC-irMS. The calculation of uncertainty on $\delta D^{GC-irMS}_{Me31}$ is evaluated as $\Delta\delta D^{GC-irMS}_{Me31} = \sqrt{(52/3)^2 \times (\delta D_{germanicolacetate})^2 + (49/3)^2 \times (\delta D_{acanh})^2}$ and gives ± 119 ‰ (see main text).

groups (δD_{Mei} ~ -100 ‰), followed by Me23, Me25, Me28 and Me29 (δD_{Mei} ~ -250 ‰), whereas Me30 and Me24 are strongly depleted (δD_{Mei} ~ -500 ‰).

More interesting is actually the global variation of $(D/H)_{Me23\text{-}Me30}$ values (the exogeneous methyl 31 group is disregarded), that can be modeled as a quasi symmetrical, bell curve, the biosynthetic implication of which is discussed below. Note also that these conclusions may still hold based on the relative variation of peak areas measured on anisotropic NAD 2D map of miliacin (see **Table S-4** and **Figure S-12**)."

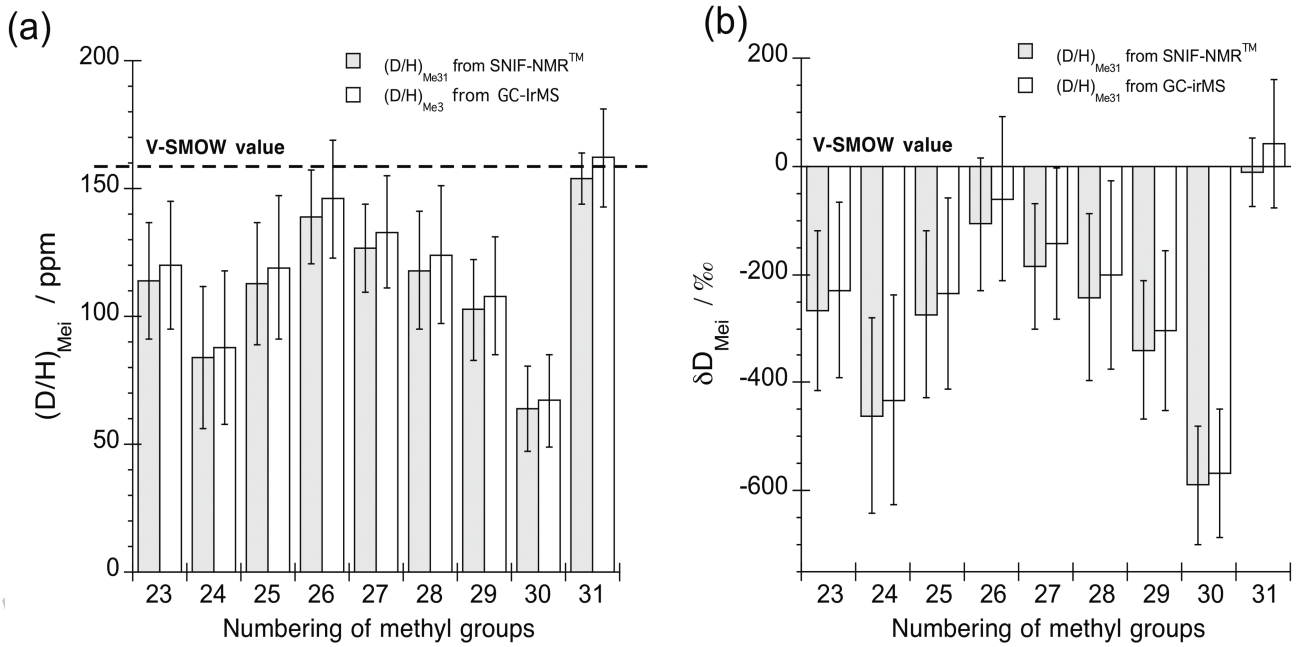


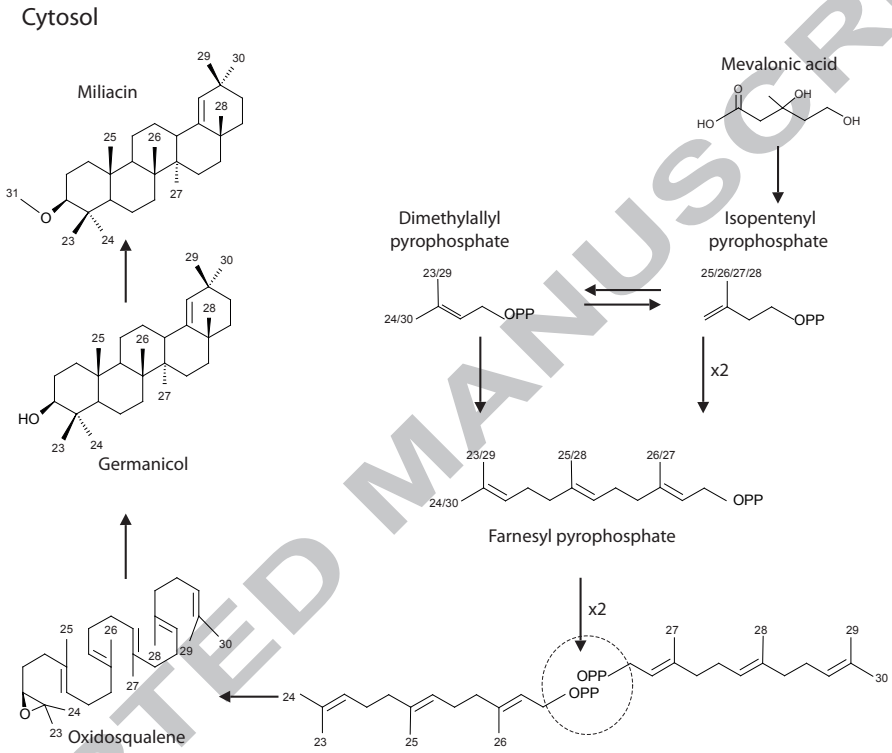
Figure 6: Variation of (a) (D/H) ratios in ppm and (b) δD in ‰ for each methyl group of miliacin when $(D/H)_{Me31}$ is evaluated by the SNIF-NMR™ (154 ± 10 ppm or -11 ± 64 ‰) and by GC-irMS (162 ± 19 ppm or 41 ± 119 ‰). Excluding the exogenous methyl group 31, note the (inverted) quasi-symmetrical bell shape of both plots.

5. DISCUSSION

5.1. Implications for biosynthesis

The heterogeneous distribution of δD_{Mei} values in miliacin can be explained by the different steps involved in the biosynthesis of this molecule, by tracking back the origin of methyl groups (**Fig.**

556 **7**). As for any C-3 oxygenated pentacyclic triterpenes in other vascular plants, miliacin is produced in
557 the cytosol from mevalonic acid, through the mevalonate pathway (i.e. Sachse et al., 2012).
558 Mevalonic acid is converted into isopentenyl pyrophosphate that can be isomerized into dimethylallyl
559 pyrophosphate. Then, farnesyl pyrophosphate is synthesized by the addition of two isopentenyl
560 pyrophosphate units and a dimethylallyl pyrophosphate unit. Oxidosqualene is synthesized by the
561 tail-to-tail combination of two farnesyl pyrophosphate units and is then cyclized by an oxidosqualene
562 cyclase.



563

564 **Figure 7**: Synthetic pathway of miliacin in the cytosol. Modified from Sachse et al. (2012).
565

566 Hence, from a biosynthetic point of view, Me24-Me30, Me23-Me29, Me25-Me28 and Me26-Me27
567 are isotopically equivalent because they derive from the same methyl groups in the farnesyl
568 pyrophosphate unit. This biosynthetic equivalence explains reasonably the symmetry and the "bell-
569 curve" variation of $(D/H)_{Mei}$ / (δD_{Mei}) values of Me23 to Me30 as observed in **Fig. 6**. Farnesyl
570 pyrophosphate is produced by head-to-tail combination of two isopentenyl pyrophosphate units and a
571 dimethylallyl unit. Thus, Me25, Me26, Me27 and Me28 in miliacin originate from the same methyl
572 groups in isopentenyl pyrophosphate, whereas Me23, Me29, Me24 and Me30 originate from the
573 dimethylallyl pyrophosphate. The dimethylallyl pyrophosphate itself results from the isomerization
574 of isopentenyl pyrophosphate by an isopentenyl diphosphate isomerase (Sachse et al., 2012).

575 During this process, the formation of a carbocation affects hydrogen atoms of Me24/Me30.
576 This could explain their D depletion when compared to other methyl groups. As already noted, Me31
577 is highly enriched when compared to other methyl groups. We have currently no information on the

origin of this methyl group that would explain such a relative D enrichment. These results for miliacin can be directly extrapolated to any compound with an oleanane, and potentially an ursane or lupane structure. For compounds that originate from oleanane after successive migrations of methyl groups (friedo rearrangements leading to structures such as taraxerane, glutinane, multiflorane and friedelane; i.e Xu et al., 2004), additional fractionations of hydrogen can be expected.

5.2. Implications for diagenesis

Diagenesis of vascular plant C-3 oxygenated pentacyclic triterpenes can potentially consist of several structural modifications that comprise loss of functional groups, ring-A degradation, epimerisations, double bond migrations, loss or migration of methyl groups (Rullkötter et al., 1994) and aromatization (*i.e.* **Fig. 8** and Wolff et al., 1989). No discernible difference in $\delta^{13}C$ values has been found during aromatization between biochemical precursors and diagenetic byproducts (Freeman et al., 1994). Preliminary results in organic archaeological layers showed significant D depletion of diagenetic derivatives compared to their potential biochemical precursors whereas no carbon isotopic shift was recorded (Jacob et al., 2011).

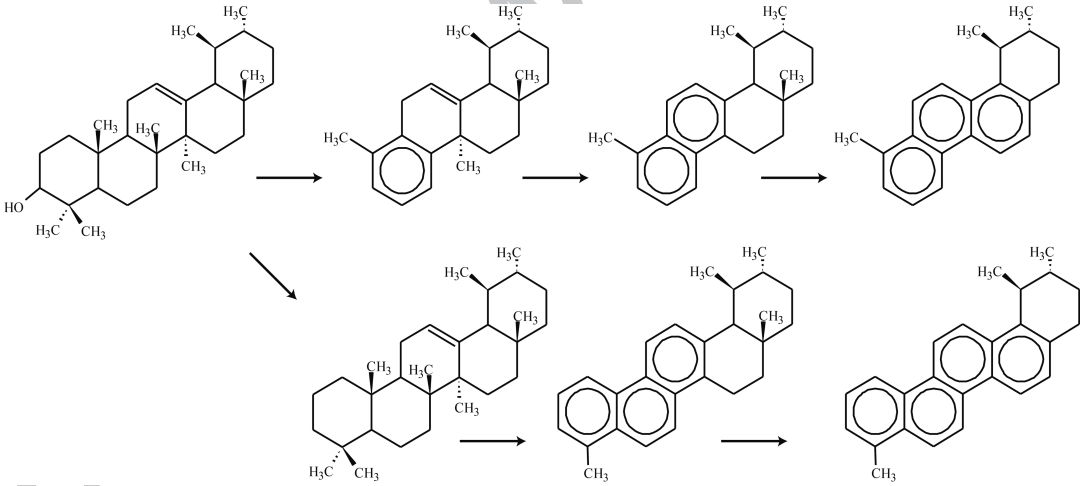


Figure 8: Potential diagenetic pathways of pentacyclic triterpenes (here starting from α-amyrin) illustrating the loss of hydrogen atoms related to methyl groups, methylene/methine. Modified from Wolff et al. (1989).

Our results have shown significant variations in δD_{Mei} values of methyl groups in miliacin with a range > 500 ‰. Thus, removing of methyl groups during diagenesis could significantly enrich diagenetic by-products compared to their biochemical precursor. But, as depicted in **Fig. 8**, the loss of methyl groups during diagenesis is always associated with the loss of hydrogen atoms in methylene and methine groups associated to rings. Thus, in order to fully understand any isotopic shift during diagenesis, it is necessary to take into account δD values of methylene and methine (δD_{meth}). Miliacin is composed by 52 hydrogen atoms of which 25 are in methine or methylene groups and 3*9 = 27 compose methyl groups. The average δD_{meth} value can thus be accessed by mass

balance according to equation **Eq. 5**.

$$\delta D_{Meth} = \left(\frac{52 \times \delta D_{miliacin} - 3 \times \sum \delta D_{Mei}}{25} \right). \quad (6)$$

Using δD_{Mei} values of methyl groups from anisotropic NMR / SNIF-NMR^{TM} combined approach (**Table 2, column 5**) and the $\delta D_{miliacin}$ = -100 ‰ in Eq. 6, we obtain an average δD_{meth} = 89 ‰. Using δD_{Mei} values of methyl groups from GC-irMS, we obtain δD_{meth} = 48 ‰. From these results it appears that methylene and methine are, on average, largely enriched in D compared to methyl groups. One could thus expect that, during diagenesis, any D-enrichment of the diagenetic byproduct compared to its precursor due to the loss of depleted methyl groups could be counterbalanced by D-depletion due to the associated loss of D-enriched methylene and methine. Nevertheless, in the absence of information on the heterogeneous distribution of D in methylene and methine, it is yet impossible to fully forecast the impacts of pentacyclic triterpenes degradation on the hydrogen isotopic composition of their by-products. In addition, only removal of hydrogen atoms is here considered but fractionation processes can be expected from the preferential biotic or abiotic degradation of biochemical precursors.

6. CONCLUSION

Obtaining the site-specific H isotope distribution in biomarkers by a reliable method is crucial but has been analytically challenging. GC-irMS allows δD values determination at a molecular scale but does not inform on any preferential position of H isotopes on molecular sites. In contrast, anisotropic NAD 2D-NMR spectroscopy provides more detailed information on isotope location within molecules. This study explains how and why NAD 2D-NMR using weakly aligning LC is a fruitful analytical tool able to overcome difficulties associated to isotropic NAD 1D-NMR (SNIF-NMR), with the final aim of studying the site-specific distribution of hydrogen isotopes. By coupling anisotropic and isotropic NAD NMR techniques or anisotropic NAD NMR and GC-irMS measurements, we were, for the first time, able not only to establish the relative distribution of hydrogen isotopes in methyl groups of miliacin, but also to afford quantitative (D/H) and δD values for each methyl group.

The resulting distribution of (D/H) values of methyl groups in miliacin provides detailed information that is consistent with its biosynthetic pathways; i.e. miliacin results from the addition of a methyl group on germanicol that is synthesized through the cyclisation of oxidosqualene. This later is produced by the condensation of two farnesyl pyrophosphate units, themselves built from two isopentenyl pyrophosphate units and a dimethylallyl pyrophosphate unit. The heterogeneous distribution of site-specific δD values revealed in this study could have major consequences on the

δD values of diagenetic derivatives of pentacyclic triterpenes since removal of methyl groups is often observed in diagenetic pathways. Although this work paves the way to exciting prospects in the frame of a better understanding of hydrogen isotopic changes during pentacyclic triterpene diagenesis, thus improving any climate reconstruction based on these derivatives, the next challenge will consist of experimentally evaluating the δD values of all hydrogen sites in biomarkers in order to have a complete picture of their heterogeneous distribution. By achieving this task, we expect being able to forecast δD values of diagenetic derivatives from their precursors or to calculate the δD values of biochemical precursors through δD values of their diagenetic derivatives.

Various analytical difficulties (assignment of the anisotropic NAD signals, low S/N ratios, …) met and discussed in this study are both related to the structure and the low solubility of miliacin. They are therefore highly specific to this structurally-complex geological biomarker, and will not necessarily come up gain when other (simpler) organosoluble biomakers will be studied. As seen here, possible protocols/strategies/solutions were proposed to overcome them, thus showing that NAD 2D-NMR in LC remains an adaptable method able to face the majority of problems associated to given analyte. In all cases, the possibility to significantly increase the concentration of deutero-isotopomers in NMR samples and/or record the NAD experiments by using very high-field NMR spectrometers (B_o > 14 T) still equipped by a ^2H cryoprobe are the most direct way to overcome the sensitivity problem, rendering the accuracy of measurements higher.

ACKNOWLEDGMENTS

French authors thank CNRS and University of Paris-Sud for their continuous financial support. P.L. thanks also Dr. Isabelle Billault for the gift of calibrated TMU and J.-P. Baltaze for his help. J.J. is grateful to R. Boscardin for her technical support and I. Lamour (HITEX, Vannes, France) for providing millet oil. This work is also a part of the ANR Project PalHydroMil project, supported by Agence Nationale de la Recherche grant (ANR JCJC, 2011-2013). V.T. was supported by PalHydroMil and Le STUDIUM Loire Valley Institute for Advanced Studies for her participation in this study. The authors wish to thank Pr. John Hayes and the two anonymous reviewers for their constructive comments on an earlier version of the manuscript.

REFERENCES

van Aarssen B.G.K., Alexander R. and Kagi R.I. (2000) Higher plant biomarkers reflect palaeovegetation changes during Jurassic times. *Geochim. et Cosmochim. Acta* **64**, 1417-1424.

Akoka. S., Barantin L., Trierweiler M. (1999) Concentration measurement by proton NMR using the

678 ERETIC method. *Anal. Chem*. 13, 2554-2557.
679 Aroulanda C., Lesot P., Merlet D., and Courtieu J. (2001) Structural ambiguities in bridged ring
680 systems resolved using natural abundance deuterium NMR in chiral liquid crystals. *J. Phys*.
681 *Chem. A*., **107**, 10911-10918.
682 Bossard N., Jacob J., Le Milbeau C., Lallier-Vergès E., Terwilliger V.J. and Boscardin R. (2011)
683 Variation in δD values of a single, species-specific molecular biomarker: a study of miliacin
684 throughout a field of broomcorn millet (*Panicum miliaceum* L.). *Rapid Com. Mass Spec*. **25**,
685 1-9.
686 Bossard, N. (2013). Pertinence et calibration d'un nouveau marqueur paléohydrologique : Le rapport
687 isotopique de l'hydrogène mesuré sur la miliacine. Ph.D. thesis, Université d'Orléans,
688 Orléans, France.
689 Bossard N., Jacob, J. LeMilbeau C., Sauze J., Terwilliger V.J., Poissonnier B., and Vergès E. (2013).
690 Distribution of miliacin (olean-18-en-3β-ol methyl ether) in broomcorn millet (Panicum
691 miliaceum), and other reputed potential sources. Consequences on the use of sedimentary
692 miliacin as a tracer of millet. *Org. Geochem*. **63**, 48-55.
693 Chikaraishi Y. and Narakoa H. (2003) Compound-specific δD-δ^{13}C analyses of n-alkanes extracted
694 from terrestrial and aquatic plants. *Phytochem*. **63**, 361-371.
695 Chikaraishi, Y., Naraoka H. and Poulson S.R. (2004) Hydrogen and carbon isotopic fractionations of
696 lipid biosynthesis among terre (C3, C4 and CAM) and aquatic plants. *Phytochem*. **65**, 1369-
697 1381.
698 Chikaraishi Y., Tanaka R., Tanaka A. and Ohkouchi N. (2009) Fractionation of hydrogen isotopes
699 during phytol biosynthesis. *Org. Geochem*. **40**, 569-573.
700 Corbet B. (1980) Origine et transformation des triterpènes dans les sédiments récents. Ph.D. thesis,
701 Université Louis Pasteur, Strasbourg, France, 106 p.
702 Cranwell P.A. (1984) Organic geochemistry of lacustrine sediments: triterpenoids of higher plant
703 origin reflecting post-glacial vegetational succession. In *Lakes Sediments and Environmental*
704 *History* (Eds. E.Y., Haworth and J.W.G. Lund). University Press, Leicester, pp. 69-92.
705 Das M. and Mahato S. (1983) Triterpenoids. *Phytochem*. **22**, 1071-1085.
706 Dansgaard W. (1964) Stable isotopes in precipitation. *Tellus* **16**, 436-468.
707 de Laeter J.R., Böhlke J.K., De Bièvre P., Hidaka H., Pieser H.S., Rosman K.J.R. and Taylo, P.D.P.,
708 (2003) Atomic weights of the elements. *Pure Appl. Chem.* **75**, (6), 683-800.
709 Eglinton I.E. and Eglinton G. (2008) Molecular proxies for paleoclimatology. *Earth and Planet. Sci.*
710 *Let*. **275**, 1-16.
711 Feakins S.J. and Sessions A.L. (2010) Controls on the D/H ratios of plant leaf waxes in an arid
712 ecosystem. *Geochim. Cosmochim. Acta* **74**, 2128-2141.

713 Freeman K. H., Boreham C.J., Summons R.E. and Hayes J. M. (1994) The effect of aromatization on
714 the isotopic compositions of hydrocarbons during early diagenesis. *Org. Geochem.* **21**, 1037-
715 1049.
716 Gleixner G. and Mügler I. (2007) Compound-specific hydrogen isotope ratios of biomarkers: tracing
717 climatic changes in the past. In *Stable Isotopes as Tracers for Ecological Change* (Eds. T.
718 Dawson and R. Siegwolf). Terrestrial Ecology Series, Academic Press, Amsterdam, pp. 249-
719 266.
720 Hou J., D'Andrea W.J., McDonald D. and Huang Y. (2007) Hydrogen isotopic variability in leaf
721 waxes among terrestrial and aquatic plants around Blood Pond, Massachussetts (USA). *Org.*
722 *Geochem.* **38**, 977-984.
723 Hyperchem™, Professional 6.1, Hypercube, Inc, 1115 NN 4th street, Gainsville, Florida, 32601,
724 USA.
725 Itô H. (1934) The chemical investigation of some gramineae oils. *J. Fac. of Agri., Hokkaido Imperial*
726 *University,* **XXXVII**, 1-40.
727 Jacob J., Disnar J.-R., Boussafir M., Sifeddine A., Albuquerque A.L.S. and Turcq B. (2005)
728 Pentacyclic triterpene methyl ethers in recent lacustrine sediments (Lagoa do Caco, Brazil).
729 *Org. Geochem.* **36**, 449-458.
730 Jacob J., Disnar J.R., Boussafir M., Albuquerque A.L.S., Sifeddine A. and Turcq B. (2007)
731 Contrasted distributions of triterpene derivatives in the sediments of Lake Caçó reflect
732 paleoenvironmental changes during the last 20,000 years in NE Brazil. *Org. Geochem.* **38**,
733 180-197.
734 Jacob J., Disnar J.-R., Arnaud F., Chapron E., Debret M., Lallier-Vergès E., Desmet M. and Revel-
735 Rolland M. (2008a) Millet cultivation history in the French Alps as evidenced by a
736 sedimentary molecule. *J. Archaeol. Sci.* **35**, 814-820.
737 Jacob J., Disnar J.-R. and Bardoux G. (2008b) Carbon isotope evidence for sedimentary miliacin as a
738 tracer of *Panicum miliaceum* (broomcorn millet) in the sediments of Lake le Bourget (French
739 Alps). *Org. Geochem.* **39**, 1077-1080.
740 Jacob J., Disnar J.-R., Arnaud F., Gauthier E., Billaud Y., Chapron E. and Bardoux G. (2009)
741 Impacts of new agricultural practices on soil erosion during the Bronze Age in French
742 Prealps. *The Holocene* **19**, 241-249.
743 Jacob J., LeMilbeau C., Bossard N., Disnar J.R. and Billaud Y. (2011) Combined $\delta^{13}C$ - δD analysis
744 of pentacyclic triterpenes and their derivatives. Studium Conference: "Hydrogen Isotopes as
745 Environmental Recorders", 15-16 September 2011, Orléans, France.
746 Killops S.D., Raine J.I., Woolhouse A.D. and Weston R.J. (1995) Chemostratigraphic evidence of
747 higher plant evolution in the Taranaki Basin, New Zealand. *Org. Geochem.* **23**, 429-445.

748 Lafon O., Lesot P., Merlet D. and Courtieu J. (2004) Modified z-gradient filtering as a new mean to
749 obtain phased deuterium autocorrelation 2D-NMR spectra in oriented solvents. *J. Magn.*
750 *Reson.* **171**, 135-142.
751 Lavrieux M., Jacob J., Le Milbeau C., Zocatelli R., Masuda K., Bréheret J. G. and Disnar J. R. (2011)
752 Occurrence of triterpenyl acetates in soil and their potential as chemotaxonomical markers of
753 Asteraceae. *Org. Geochem.* **42**, 1315-1323.
754 Lesot P., Merlet D., Lowenstein A. and Courtieu J. (1998) Enantiomeric Visualisation using Proton-
755 Decoupled Natural Abundance Deuterium NMR in Poly(γ-benzyl-L-glutamate) Liquid
756 Crystalline Solutions. *Tetrahedron: Asymmetry* **9**, 1871-1881.
757 Lesot P., Baillif V. and Billault I. (2008) Combined analysis of four C-18 unsaturated fatty acids
758 using Natural Abundance Deuterium 2D-NMR spectroscopy in chiral oriented solvents. *Anal.*
759 *Chem.* **80**, 2963-2972.
760 Lesot P. and Lafon O. (2008) Enantiomeric Analysis using Natural Abundance Deuterium 3D-NMR
761 Spectroscopy in Polypeptide Chiral Oriented Media. *Chem. Phys. Lett.* **458**, 219-222.
762 Lesot P. and Courtieu J. (2009) Natural abundance deuterium NMR spectroscopy: Developments and
763 analytical applications in liquids, liquid crystals and solid phases. *Prog. Nucl. Magn. Reson.*
764 *Spectrosc.* **55**, 128-159.
765 Lesot, P., Serhan, Z. and Billault, I. (2011) Recent advances in the analysis of the site-specific
766 isotopic fractionation of metabolites such as fatty acids using anisotropic natural abundance
767 ^{2}H NMR spectroscopy: Application on conjugated linolenic methyl esters. *Anal. Bioanal.*
768 *Chem.* **399**, 1187-1200.
769 Lesot P., Serhan Z. Aroulanda C. and Billault I. (2012) Analytical contribution of NAD 2D-NMR
770 spectroscopy in polypeptide mesophases to the investigation of triglycerides *Magn. Reson.*
771 *in Chem.* **50**, S2-S11.
772 Lesot P., Aroulanda C., Zimmermann, H. and Luz Z. (2015) Enantiotopic Discrimination in the NMR
773 Spectrum of Prochiral Solutes in Chiral Liquid Crystals, *Chem. Soc. Rev.* **44**, 2330.
774 Massou S., Nicolas C., Letisse F., Portais J.-C., Nmr-Based Fluoxomics: Quantitative 2D-NMR
775 methods for isotopomers Analysis. *Phytochemistry*, 2007, **68**, 2330-2340.
776 Liu W., Yang H. and Li L. (2006) Hydrogen isotopic composition of *n*-alkanes from terrestrial plants
777 correlate with their ecological life form. *Oecologia* **150**, 330-338.
778 Lohmann F. (1988) Aromatisations microbiennes de triterpènes végétaux. Ph.D. thesis, Université
779 Louis Pasteur, Strasbourg, France.
780 Mahato S.B., Nandy A.K. and Roy G. (1992) Triterpenoids, *Phytochem.* **31**, 2199-2249.
781 Mahato S.B. and Sen S. (1997) Advances in triterpenoid research, 1990-1994. *Phytochem.* **44**, 1185-
782 1236.

783 Martin G.J. and Martin M.L., (1981) Deuterium labelling at the natural abundance level as studied by
784 high field quantitative ^{2}H NMR, *Tet. Lett.* **22**, 3525-3528.
785 Martin G.J., Zhang B.-L., Naulet N. and Martin M.L., (1986) Deuterium transfer in the bioconversion
786 of glucose to ethanol studied by speific labeling at the natural abundance level. *J. Am. Chem.*
787 *Soc.* **108**, 5116-5122.
788 McInerney F.A., Helliker B.R. and Freeman K.H. (2011) Hydrogen isotope ratios of leaf wax n-
789 alkanes in grasses are insensitive to transpiration. *Geochim. Cosmochim. Acta* **75**, 541-554.
790 Merlet D., Ancian B., Courtieu D. and Lesot P. (1999) Two-dimensional Deuterium NMR
791 spectroscopy of chiral molecules oriented in a polypeptide liquid crystal: application for the
792 enantiomeric analysis through natural abundance deuterium NMR. *J. Am. Chem. Soc.* **121**,
793 5249-5258.
794 Motuzaite-Matuzeviciute G., Jacob J., Telizhenko S., Jones M.K. (2013) Miliacin in paleosoils from
795 early Iron Age in Ukraine reveals paleofields of broomcorn millet. *Archaeological and*
796 *Anthropological Sciences*, 1-8.
797 Ohmoto T., Ikuse M. and Natori S. (1970) Triterpenoids of the Gramineae. *Phytochem.* **9**, 2137-
798 2148.
799 Regnery J., Püttmann W., Koutsodendris A., Mulch A. and Pross J. (2013) Comparison of the
800 paleoclimatic significance of higher land plant biomarker concentrations and pollen data: A
801 case study of lake sediments from the Holsteinian interglacial. *Org. Geochem.* **61**, 73-84.
802 Rullkötter J., Peakman T.M. and ten Haven, H.L. (1994) Early diagenesis of terrigenous triterpenoids
803 and its implications for petroleum geochemistry. *Org. Geochem.* **21**, 215-233.
804 Sachse D., Radke J. and Gleixner G. (2006) δD values of individual n-alkanes from terrestrial plants
805 along a climatic gradient – Implications for the sedimentary biomarker record. *Org. Geochem.*
806 **37**, 469-483.
807 Sarfati M., Lesot P., Merlet D. and Courtieu J. (2000) Theoretical and experimental aspects of
808 enantiomeric differentiation using natural abundance multinuclear NMR spectroscopy in
809 polypeptide liquid crystals. *Chem. Commun.* **21**, 2069-2081.
810 Sauer P.E., Eglinton T.I., Hayes J.M., Schiemmelmann A., Sessions A.L. (2001) Compound-specific
811 D/H ratios of lipid biomarkers from sediments as a proxy for environmental and climatic
812 conditions. *Geochim. Cosmochim. Acta* **65**, 213-222.
813 Schwab V.F., Garcin Y., Sachse S., Todou G., Séné O., Onana J.-M., Houndong G. and Gleixner G.
814 (2015) Effect of aridity on $\delta^{13}C$ and δ D values of C3 plant- and C4 graminoid-derived leaf
815 wax lipids from soils along an environmental gradient in Cameroon (Western Central Africa)
816 *Org. Geochem.* **78**, 99-109.

817 Serhan Z., Martel L., Billault I. and Lesot P. (2010) Complete determination of site specific bio-
818 enantiomeric excesses in linoleic acid using natural abundance deuterium 2D-NMR in
819 polypeptide mesophase. *Chem. Commun.* **46**, 6599-6601.
820 Sessions A.L., Burgoyne T.W., Schimmelmann A. and Hayes J.M. (1999) Fractionation of hydrogen
821 isotopes in lipid biosynthesis. *Org. Geochem.* **30**, 1193-1200.
822 Smetanina O.F., Kuznetzova T.A., Denisenko V.A., Pivkin M.V., Khudyakova Y.V., Gerasimenko
823 A.V., Popov D.Y., Il'in S.G. and Elyakov G.B. (2001) 3β-Methoxyolean-18-ene (miliacin)
824 from the marine fungus *Chaetomium olivaceum*. *Russ. Chem. Bull.* **50**, 2463-2465.
825 Smith F.A. and Freeman K.H. (2006) Influence of physiology and climate δD of leaf wax n-alkanes
826 from C3 and C4 grasses. *Geochim. Cosmochim. Acta* **70**, 1172-1187.
827 Terwilliger V.J., Eshetu Z., Disnar J.R., Jacob J., Adderley W.P., Huang Y., Alexandre M. and Fogel
828 M. (2013) Environmental changes and the rise and fall of civilizations in the northern Horn of
829 Africa: an approach combining δD analyses of land-plant derived fatty acids with multiple
830 proxies in soil. *Geochim. et Cosmochim. Acta* **111**, 140-161.
831 Terwilliger V.J. and Jacob J. (2013) Introduction: Hydrogen isotopes as environmental recorders.
832 *Geochim. et Cosmochim. Acta* **111**, 1-4.
833 Trendel J. (1985) Dégradation des triterpènes dans les sédiments. Aspects photochimiques et
834 microbiologiques. Ph.D. thesis, Université Louis Pasteur, Strasbourg, France, 126 p.
835 Wang Y., Sessions A.L., Nielsen R.J. and Goddard W.A. (2013) Equilibrium $^{2}H/^{1}H$ fractionation in
836 organic molecules: III. Cyclic ketones and hydrocarbons. *Geochim. et Cosmochim. Acta* **107**,
837 82-95.
838 Wang Y.V., Larsen T., Leduc G., Andersen N., Blanz T. and Schneider R.R. (2013) What does leaf
839 wax δD from a mixed C3/C4 vegetation region tell us? *Geochim. et Cosmochim. Acta* **111**,
840 128-139.
841 Wolff G.A., Trendel J.M. and Albrecht P. (1989) Novel Monoaromatic triterpenoids hydrocarbons
842 occuring in sediments. *Tetrahedron* **45**, 6721-6728.
843 Xu R., Fazio G.C. and Matsuda S.P.T. (2004) On the origins of triterpenoid skeletal diversity.
844 *Phytochem.* **65**, 261-291.
845

FIGURE CAPTIONS

Figure 1: Structure of miliacin where R_{31} = -OMe and gemanicol where R_{31} = -OH, together with the atomic numbering used and notation of rings (A to E) and methyl groups (23 to 31).

Figure 2: 600.1 MHz proton 1D-NMR spectrum of miliacin in $CDCl_3$ (295 K). The assignment of resonances of methyl groups is given in **Table S-1**, **S-2** and **S-3** in **Appendix A.I** of **Supp. Info.**). The peak marked by an asterisk (δ = 1.24 ppm) is an unidentified impurity.

Figure 3: (a) Methyl region (Me31 not shown) of the 92.1 MHz NAD tilted *Q*-resolved Fz 2D map of miliacin dissolved in PBLG/$CHCl_3$ and recorded at 295 K. (b) 600 MHz proton 1D spectrum of miliacin in $CHCl_3$ plot at the horizontal same scale (in ppm). A lorenzian filtering (LB = 1.5 Hz) has been applied in both spectral dimensions. On the F_1 and F_2 axes are shown the spectral 2D projections. Note the significant variations of peak intensity of ^{2}H doublets in F_1 dimension illustrating the enrichment/depletion in deuterons.

Figure 4: Series of four NAD 1D sub-spectra extracted from the tilted *Q*-resolved Fz 2D map showing the nine methyl groups of miliacin along with their assignments (see also **Fig. S-9**). Each sub-spectrum is plotted with the same amplitude of noise. Note the variation of peak intensity of ^{2}H doublets. The variations observed reveal the differences of (D/H) for each site.

Figure 5: 3D view of the minimal energy conformation of miliacin calculated with Hyperchem 3.0 software. A semi-empirical quantum chemistry methods based on the Hartree–Fock formalism (Austin Model 1 or AM1) has been used for modeling the molecule. The minimal energy was found at -8177.99 kJ/mol. Except for ring E distorted by the double bond, all other cycles have adopted a chair conformation. Note the difference of orientation axes of C-CH_3 vectors for Me23 and Me30 (close to the median horizontal molecular plan) compared to Me24 to Me29 (see also **Fig. S-9** in **Appendix A.II** of **Supp. Info.**).

Figure 6: Variation of (a) (D/H) ratios in ppm and (b) δD in ‰ for each methyl group of miliacin when $(D/H)_{Me31}$ is evaluated by the two-step NAD NMR approach. Excluding the exogenous methyl group 31, note the (inverted) quasi-symmetrical bell shape of both plots.

Figure 7: Synthetic pathway leading to miliacin. Modified from Sachse et al. (2012).

Figure 8: Potential diagenetic pathways of pentacyclic triterpenes (here starting from α-amyrin) illustrating the loss of hydrogen atoms related to methyl groups, methylene/methine. Modified from Wolff et al. (1989).