NMR methodology for complex mixture ‘separation’†

Nicholle G. A. Bell,a Lorna Murray,a Margaret C. Grahamb and Dušan Uhrin*a

Mixture ‘separation’ by NMR is demonstrated through the development of a pseudo 4D NMR experiment, 3D IPAP INEPT-INADEQUATE-HSQC, designed for the structural elucidation of 13C tagged compounds.

The structural elucidation of compounds contained within complex mixtures is a challenging task. Despite the advances in chromatography, some mixtures cannot be separated. Humic substances (HS), produced by the biodegradation of plant and animal residues, are the best known example of an ‘inseparable’ mixture consisting of thousands of organic compounds. Ubiquitous in nature, they make up a considerable proportion of the Earth’s carbon pool and are key players in many biogeochemical processes.

In order to comprehend the functional roles of HS on a molecular level, their structural composition needs to be deciphered. Humic substances, produced by the biodegradation of plant and animal residues, are the best known example of an ‘inseparable’ mixture consisting of thousands of organic compounds. Ubiquitous in nature, they make up a considerable proportion of the Earth’s carbon pool and are key players in many biogeochemical processes.

Unfortunately, these quickly fail when the complexity of mixtures increases. Inevitably, for chromatographically inseparable mixtures the “separation” must be done spectroscopically. One way of achieving this is by tagging the molecules with isotopically labelled nuclei. Once in place, the polarization transfer pathways are directed through these tags, reducing the complexity of spectra significantly. As illustrated here, this approach has a potential to elucidate molecular fragments of compounds contained in complex mixtures.

The bulk of HS is composed of oxygen rich H2C2O2 compounds with the Mw range of ~200–1000 g mol−1. Particularly prevalent functionalities decorating aromatic and aliphatic skeletons of HS are the hydroxyl and carboxyl groups. These groups are crucial to the interaction of HS with species such as heavy metals as well as the self-association of HS molecules. The methodology presented here aims at characterizing the aromatic moieties of HS carrying OH and COOH groups and relies on introducing 13C-enriched −O13CH3 and −COO13CH3 groups into HS compounds. HS have been methylated in the past and the inspection of −O13CH3 resonances yielded some rudimentary information about the nature of their COOH and OH groups. The novelty of our approach is that it uses labels to spy on their neighbourhood, obtaining the 1H and 13C chemical shifts and 1H−1H and 1H−13C coupling constants of the nuclei in their vicinity.

To achieve this aim we have designed a novel 3D NMR experiments, referred to here as 3D IPAP INEPT-INADEQUATE-HSQC (Fig. 1).

This NMR experiment can be viewed as a 3D extension of a 2D INEPT-INADEQUATE. It is also related to 3D HCCH experiments.3,4 3D IPAP INEPT-INADEQUATE-HSQC exploits polarization transfer pathways shown in Fig. 1, and correlates the double-quantum (DQ) coherences of long-range coupled carbons (F1) with corresponding single-quantum 13C chemical shifts (F2) and the 1H chemical shifts (F3). The polarizations transfer is tuned for 2JzCC (Fig. S1, ESI†) and correlates the chemical shifts of nuclei in 13CH4···13CH4 (y = 0, 1) fragments. It starts and ends concurrently on methoxy and aromatic protons located next to the methoxy groups. For CH3···CH fragments it provides chemical shifts of all four nuclei and can therefore be regarded as a pseudo 4D experiment; all three chemical shifts are obtained for the CH3···C9 moieties. Since the acquired NMR signal is filtered via isotopically enriched 13CH3 groups, the resulting spectra are significantly simplified.

The limited 1H and 13C chemical shifts range of methoxy groups (Fig. S2, ESI†) necessitates the use of high digital resolution in the

† Electronic supplementary information (ESI) available: Materials and methods, details of the optimization of NMR parameters, values of coupling constants, S/N ratios, examples of processing of IPAP spectra and a suppression of signals from aliphatic OMe groups. See DOI: 10.1039/c3cc48907h
The details of the pulse sequence are given in the ESI.†

**Fig. 1** The pulse sequence of the 3D IPAP INEPT-INADEQUATE-HSQC and its polarisation transfer pathways utilising various coupling constants.

The proposed method is illustrated on a model mixture of nine benzoic acid derivatives, 1a–9a, which were fully methylated using 13CH3I, yielding compounds 1–9 (Scheme 1, average concentration of 1.4 mM).

2D F2F3 (HSQC) (Fig. 2a and b) and 2D F1F2 (DQ) (Fig. 2c and d) planes of the 3D IPAP INEPT-INADEQUATE-HSQC spectrum of the model mixture illustrate the resolving power of this experiment. The spectrum was analysed as outlined in Fig. 2e–g. Shown here are the 2D DQ (F1F2) planes extracted at the aromatic and Me carbons (F3) indicated by red arrows in Fig. 2a and b. The obtained 13C and 1H chemical shifts allowed unambiguous identification of the fragment highlighted in the inset. The fragment was assigned to the correct molecule by considering the electronic effects of OMe and COOMe groups on the 13C and 1H chemical shifts of benzene and by analysing the proton–proton and long-range proton–carbon coupling constants determined in F3. The theoretical sensitivity of this experiment is comparable to that of a refocused gradient-selected 1H, 13C HMBC optimized for 3JCH couplings. We have determined the S/N ratios in 1D F3 traces extracted from F2F3 planes of the 3D spectrum obtained by the addition of the 3D IP and AP INEPT-INADEQUATE-HSQC spectra. The values were normalized (Fig. S7, ESI†) to the average concentration of the methylated compounds (1.4 mM). Due to their singlet character, the S/N is higher for –O13CH3 (60–1900 : 1) than for the aromatic resonances (57–600 : 1). Somewhat lower values were obtained for –COO13CH3 signals (40–240 : 1). These values clearly reflect the sizes of 2JCC couplings (Fig. S1, ESI†) involved in the polarisation transfer and can be increased for small coupling constants by optimizing the Δτ intervals for smaller JCC couplings (6 Hz at present). This will have to be balanced against faster relaxation in larger molecules. We note that the relaxation effects can be reduced by using 13CD2H in place of 13CH3 groups. This modification opens a route toward a design of more efficient 3D INADEQUATE-based experiments thus compensating partially for the loss of protons in 13CD3 groups.

Considering the lowest S/N observed (40 : 1), a 10-fold reduction of the concentration would still yield spectra of adequate quality. Thus assuming a 140 μM concentration of a 500 g mol⁻¹ Mw compound, 38.5 μg dissolved in 550 μl of CDCl₃ are required per compound. A 19.3 mg strong mixture of compounds containing 500 unique substitution patterns by –OH and –COOH groups on aromatic rings will therefore provide 3D spectra with sufficient S/N ratios to be analyzed. As the methylation procedure includes extraction into the organic phase, and thus selects a subset of HS compounds, simplification of these complex mixtures is inherent to this process. This could be further improved by employing partial fractionation of HS prior to methylation.

Obtaining 1H and 13C chemical shifts of the nuclei in the vicinity OH and COOH by itself does not lead to a structure, and
such data needs to be interpreted by considering the precursors of HS molecules and their chemical shifts. However, it is clear that this methodology supersedes information content of simple $^1$H, $^{13}$C HSQC spectra of HS.$^{16}$ The key is the increased number of nuclei that are identified as belonging to the same molecule. Additional valuable information can be obtained by hydrolysing the $-\text{COO}^{13}\text{CH}_3$ esters, while keeping the $-\text{O}^{13}\text{CH}_3$ groups untouched. The resulting site-specific changes of chemical shifts of aromatic spins can be quantified through recording the 3D IPAP INEPT-INADEQUATE-HSQC of a partially hydrolysed mixture (data not shown). Such information is very valuable for the identification of fragments carrying both OH and COOH groups.

In conclusion, we conjecture that tagging by $^{13}$C labelled methyl groups in combination with 3D NMR spectroscopy is a very promising approach to the analysis of complex mixtures such as HS. To the best of our knowledge, this work is the first example of a sophisticated NMR experiment designed to explore the chemical environment around the tags, rather than just the signals from the tags themselves. This approach is not limited to methylation, other tags, containing NMR active, fully abundant nuclei such as $^{15}$N or $^{19}$F can also be explored. Finally, many other mixtures can benefit from this approach, including those found in food, beverages, natural products, or biological samples, to name a few.

This work was supported by the NERC grant NE/L00044X/1 to MCG and DU. NGAB would like to acknowledge the support of the University of Edinburgh Principal’s Career Development Scholarship and NERC. We thank J. Bella for maintaining the NMR spectrometers.
Notes and references