Title:

Diffusion-ordered nuclear magnetic resonance spectroscopy (DOSY-NMR): a novel tool for identification of phosphorus compounds in soil extracts

Keywords: DOSY, NMR, phosphorus compounds, soil science

Authors:

Liming Wang[1 *], **Wulf Amelung**[2, 3], **Sabine Willbold**[1],

[1] Central Institute for Engineering, Electronics and Analytics, Analytics (ZEA-3), Jülich, Germany,

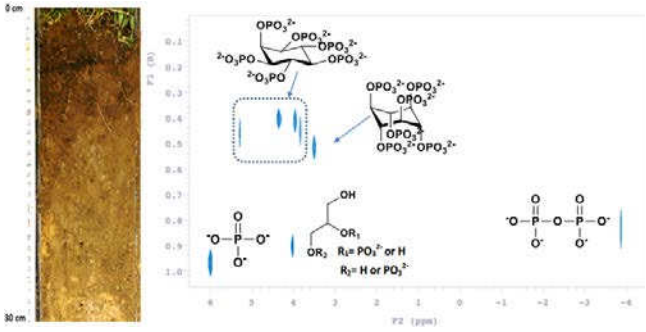
[2] Institute of Bio- and Geosciences, Agrosphere Institute (IBG-3), Forschungszentrum Jülich GmbH, Wilhelm Johnen Straße, 52425 Jülich, Germany,

[3] Institute of Crop Science and Resource Conservation, Soil Science and Soil Ecology, University of Bonn, Nußallee 13, D-53115 Bonn, Germany

Correspondence to: Liming Wang, lim.wang@fz-juelich.de, tel: +49 2461618730, fax: +49 2461612560

Abstract

Liquid-state, one-dimension ^{31}P nuclear magnetic resonance spectroscopy (NMR) has greatly advanced our understanding of the composition of organic phosphorous in the environment. However, the correct assignment of signals is complicated by overlapping and shifting signals in different types of soils. We applied therefore for the first time diffusion-ordered spectroscopy (DOSY) to soil extracts, allowing us to separate phosphorus components in the second domain based on their translational diffusion coefficients. After successful application to a mixture of fourteen model compounds, diffusion rates correlated closely with the molecular weight of the individual compound in aqueous solution ($R^2 = 0.97$). The method was then applied to NaOH/EDTA extracts of a grassland soil, of which paramagnetic contaminations were removed with sodium sulfide following high-velocity centrifugation (21500 g, 45 min) at 4°C. Diffusion rates in soil extracts were again closely related to molecular weight ($R^2 = 0.98$), varying from 163.9 to 923.8 Da. However, our DOSY application failed for a forest soil with low organic phosphorus content. Overall, DOSY did help to clearly identify specific NMR signals like *myo*- and *scyllo*-inositol hexakisphosphate. It thus provides a more confident signal assignment than 1D ^{31}P-NMR, although currently the ubiquitous use of this novel methodology is still limited to soil with high organic phosphorus content.



1. Introduction

Phosphorus (P) is one of the lesser available nutrients for primary production of forest and agricultural ecosystems.[1-4] Soils generally contain around 100-3000 mg P kg^{-1} soil, however ca. 15-80 % of which is present in organic forms.[5] As not all P compounds in soils are bioaccessible or bioavailable for the uptake by plants,[4] an enhanced release of P from these organic forms may be crucial for plant nutrition.[6] There are a wide range of organic P structures, including various phosphate diesters, monoesters and phosphonates reported on in soil;[7-9] yet, their exact origin and turnover is still hard to decipher. The limitations lie in the sample preparation and analytical method. At present, a rapid single step extraction with ethylenediaminetetraacetate (EDTA) and NaOH followed by ^{31}P nuclear magnetic resonance spectroscopy (NMR) analysis is the most widely used strategy for specification of soil organic P. This method is known to extract the largest amount and diversity of P species compared to other common methods,[10] but paramagnetic metal ions are simultaneously introduced into extract, leading to line broadening and undermining resolution of NMR spectra. Plenty of post-treatment procedures have been developed to remove paramagnetic ions, including anion exchange resins [11, 12] and dialysis membrane,[13, 14] the latter resulting in considerable risks of P losses.[15] Subsequently, Vestergren et al. (2012) [16] successfully treated forest soil extracts with Na_2S, therewith considerably increasing resolution of NMR spectra.

NMR signal assignment often relies on the direct comparison of chemical shift with previous studies. Some authors provided comprehensive libraries with a wide range of P model compounds for peak identification.[17, 18] Yet, this is also problematic because ^{31}P chemical shifts might depend on sample matrix properties, such as ionic strength and pH value. The solvent matrix used usually changes the spectral profile from the same soil sample.[19, 20] Especially, in crowded NMR spectra this may result in additional risks of signal misidentification, such as has been repeatedly reported for the monoester region of soil spectra. [21, 22] For correct sig-

nal identification, an increasing number of studies are verifying the NMR signals in monoester region by spiking with reference compounds.[23-25] However, more recently 2D-^{31}P-NMR approaches such as ^{31}P-^{1}H correlation spectroscopy have also found their way from biochemistry into soil science.[16, 26] The method allows for the unambiguous identification of many P signals according to chemical shift information from two (^{1}H and ^{31}P) domains and characteristic J-coupling interactions between P-C-H or P-O-C-H chemical bonds. Nevertheless, the interpretation of the corresponding 2D Spectra is also not straightforward, because there is as yet no direct separation of ^{31}P compounds of different size and mobility. [27] In contrast, diffusion-ordered spectroscopy (DOSY) gives a 2D plot with frequency (F) in horizontal domain, and diffusion coefficient (D) in another, which directly correlates with molecular weight (MW) under certain environment. [28] In this regard, DOSY enables to virtually separate the NMR signals of different species in a second domain according to their difference in D. Combining the additional information from the D domain with the chemical shifts in the F domain, confident peak identification can be achieved. To the best of our knowledge, DOSY is starting to gain popularity in the field of environmental research.[28, 29] In fact, ^{1}H or ^{19}F DOSY has already been employed to reveal aggregation behavior of natural organic matter,[30, 31] and the interaction between pollutant and humic matter in nature, [32, 33] thus offering also potentials for soil P research such as their distribution in different molecular weight fractions [34, 35] and the structural elucidation. The objectives of the study were : (i) to demonstrate the feasibility of DOSY experiment under ideal circumstances, i.e., using high concentrations of P model compounds in typical solvent (NaOD and D_2O) but without interference such as paramagnetic ions and the viscous soil solution matrix, (ii) to revise the Na_2S precipitation protocol for efficient and practical removal of paramagnetic ions needed for DOSY application, and (iii) to apply DOSY for soil P analysis and demonstrate its strengths and limitations.

2. Materials and methods

2.1 Reference standards

A range of P compounds were selected to cover various range of MWs and each P category representatively existed in soil, including phosphonate, orthophosphate monoesters, orthophosphate diesters, pyrophosphate, and polyphosphate (**Table 1**). Compounds were purchased from Sigma-Aldrich, except for adenosine 5'-triphosphoric acid disodium acid (Pan-Reac AppliChem) and methylenediphosphonic acid (Alfa Aesar). The final concentration of each model compounds is 5mg/mL for NMR data acquisition. A mixture of NaOD (30%, w/v in D_2O) and D_2O was selected to dissolve model compounds and soil samples as well as to maintain a pH>13. The pKa of HPO_4^{2-} (12.67, at 25°C) is higher than any other common organic P compounds,[36] a pH of 13 is high enough to transform all P species into sodium salt hydrate form for accurate estimation of MW.

2.2 Soil collection and characterization

The methodological tests were performed with the surface A horizon from the Rollesbroich grassland test site, comprising a permanent temperate grassland that is part of the TERENO long-term field observatory network in Germany (50°37'26"N 6°18'15"E; located in North Rhine-Westphalia, Germany)[37]. According to the World Reference Base for Soil Resources (WRB) (IUSS Working Group WRB, 2015), the soil was classified as Cambisol (referred as **Roll**). Another sample was from the weathered Of horizon of an organic forest soil layer (Histosol) from Wettersteinwald in Bavarian limestone Alps (referred as **Wett**). After drying and sieving (2mm mesh size), the soils were then ground in a Retsch MM 400 ball mill (2 min, 400 RPM). For alkaline extraction of soil organic P structures, we followed a commonly applied procedure, [16, 21, 26, 38, 39] i.e., 10 g soil was shaken for 16 h with a mixture of 0.25 M NaOH and 50 mM Na_2EDTA (soil-to-solution mass ratio 1:20),[10] and centrifuged at 10000 g for 30 min. The supernatants were frozen at -18°C and subsequently lyophilized, yielding 5.22 g and 4.90 g solid material, respectively. The contents of P, Al, Ca, Mn and Fe were

determined by ICP-MS (see **Table S1 of supporting information**). Post-treatment of Roll soil was required to remove high content of Fe and Mn.

2.3 Na_2S treatment

The procedure of Na_2S treatment was established on the scheme proposed by Vestergren et al. (2012),[16] but with modifications using a prolonged high-speed centrifuge (215 00 g, 45 min) to get rid of fine colloids, as well as working at low temperatures (4°C) in order to improve the precipitation of excessive Na_2S by decreasing its saturated solubility. Three groups (in triplicate) of soil extracts were prepared to compare the resolution. In brief, 350 mg of soil extract was dissolved in 600 µL mixture of 30% NaOD and D_2O (pH 13) to get more concentrated P, in nonuplicate. The nine samples were separated into 3 groups (3 samples each). Group 1: centrifuged (7000 g, 30 min) at room temperature. Subsequently, 500mg of $Na_2S.9H_2O$ powder was added in each of the six remaining solutions and these were sonicated to produce a saturated solution, and then shaken for 16h at the horizontal mixer at ambient temperature. Group 2: centrifuged (7000 g, 30 min) at room temperature. Group 3: centrifuged (215 00 g, 45 min) at 4 °C. The decanted supernatant was then transferred to a 5-mm diameter NMR tube for data acquisition. The residues in the vials were weighted to compare the removal of particles and excessive Na_2S.

2.5 NMR parameters

The 1D NMR spectra and DOSY spectra were obtained on a Varian 600MHz spectrometer at a ^{31}P frequency of 242.81 MHz, equipped with 5mm broadband probe. The 1D spectra were acquired with the following parameters: 90° pulse calibrated at 10.59 µs, 0.680 s acquisition time, no spinning, 298K; proton inverse-gated decoupling; 15 s pulse delay was used for both 1D and DOSY acquisition, which is considerably long than other studies [19] and lead to a unusual long experimental time for all data acquisitions of the present study. The purpose is to cool down the conductive samples seriously heated by the decoupler coil of the probe

used (Varian high-field switchable broadband probe, part No. 01-908118-00).[40] For optimization of DOSY experimental parameters, the combination of diffusion gradient length (δ), diffusion delay (Δ), gradient stabilization delay, gradient strength (g) were adjusted to generally get >85 %[41], or ideally 90-95 % [42] total signal attenuation throughout the experiment. The DOSY parameters for model compounds were: 4.5 ms diffusion gradient length (*δ*), 100ms diffusion delay (*Δ*), 0.5ms gradient stabilization delay, 25 gradient increments with gradient strength (*g*) from 1.3 to 32.5 G.cm-^{1}, total acquisition time 14 hours. BPPSTE[43] (bipolar pulse pair stimulated echo) pulse sequence applied in this study enable the best quality of stacked pulsed field gradient NMR spectra. A detailed comparison of common pulse sequences is available[44]. Briefly, the anti-phase gradient pulse pair bracketing180^{o} pulse cancels the perturbation to B_0 field and deuterium field brought by gradient field. The phase error, baseline distortion and other systematic error are largely alleviated in the present setup.

DOSY was firstly introduced by Morris and Johnson in 1992 as facile alternative for mixture analysis. Molecules experience constantly translational motion commonly referred as diffusion. The diffusion coefficient (D) defines the root mean square distance traveled during a period of time for a group of certain molecule. D is an inherent characteristic parameter for individual compound. For a spherical molecule in certain environment, D value scales with molecular weight (MW) as

$$D \propto MW^{-1/3} \quad \textbf{(1)}$$.[28]

With the introduction of DOSY, D from the second domain may help with signal assignment. In addition, the MW value can be estimated based on equation **(1)**.

DOSY measurement for soil sample were acquired with an array of 12 gradient amplitudes ranging from 1.3 to 32.5 G.cm-^{1} in equal steps of gradient squared, 298 K, BPPSTE pulse sequence. The parameters for Roll soil are 2214 transients, 5099 complex data points, total

acquisition time 114 hours, a total diffusion encoding gradient of 5 ms, and a diffusion time of 150 ms. The parameters for Wett soil are 1227 transients, 5099 complex data points, total acquisition time 63 hours, a total diffusion encoding gradient of 5 ms, and a diffusion time of 70 ms. The NMR results were processed by the Vnmrj software (version 4.2, revision A), 1 Hz and 2 Hz line broadening was applied to P model compounds and the soil sample individually. Baseline correction and non-uniformity gradient correction features were applied to get a better curve fitting.

3. Results and Discussion

3.1 Model compounds

A mixture of fourteen P compounds, covering different categories, in NaOD and D_2O was analyzed for this study. This mixture included (12-phosphonododecyl) phosphonic acid (**1**), 2-aminoethylphosphonic acid (**2**), methylenediphosphonic acid (**MDP**, **3**), *myo*-phytic acid sodium salt hydrate (***myo*-IHP**, **4**), β-glycerophosphate disodium salt hydrate (**5**), guanosine 5'-monophosphate (**6**), cytidine 5'-monophosphate (**7**), α-D-glucose-1-phosphate (**8**), lipoteichoic acid from *Staphylococcus aureus* (**LTP**, **9**), desoxyribonucleic acid sodium salt from salmon testes (**DNA**, **10**), sodium pyrophosphate tetrabasic (**11**), adenosine 5'-Triphosphoric acid disodium acid (**ATP**, **12**), P^1P^5-di(adenosine-5') pentaphosphate (**13**), and ribonucleic acid from torula yeast Type VI (**RNA**, **14**). The DOSY spectrum shown in **Fig. 1** indicated RNA was degraded into 8 mononucleotides (signals in the dashed box **#, enlarged view of Fig. 1 was shown in Figure S1 of supporting information)** during the acquisition process. Unknown signals denoted as **&** are also shown in the spectrum. Signals were well-resolved in the D domain, signals from *myo*-IHP (signals in the dashed box **4,** 5.78, 4.86, 4.52, 4.37 ppm), ATP (signals in the dashed box **12**, from -4.47 to -20.22 ppm), and P^1P^5-Di(adenosine-5') pentaphosphate (**13**, from 10.15 to 21.40 ppm) could be readily identified at first glance. By contrast, in corresponding 1D spectrum, the resonances were too complex to be identified.

In contrast to 2D-^{31}P-NMR, DOSY simply separated P species along the F1 domain according to their molecular sizes. The DOSY spectrum is thus easier to interpret than 2D-^{31}P-NMR spectra.

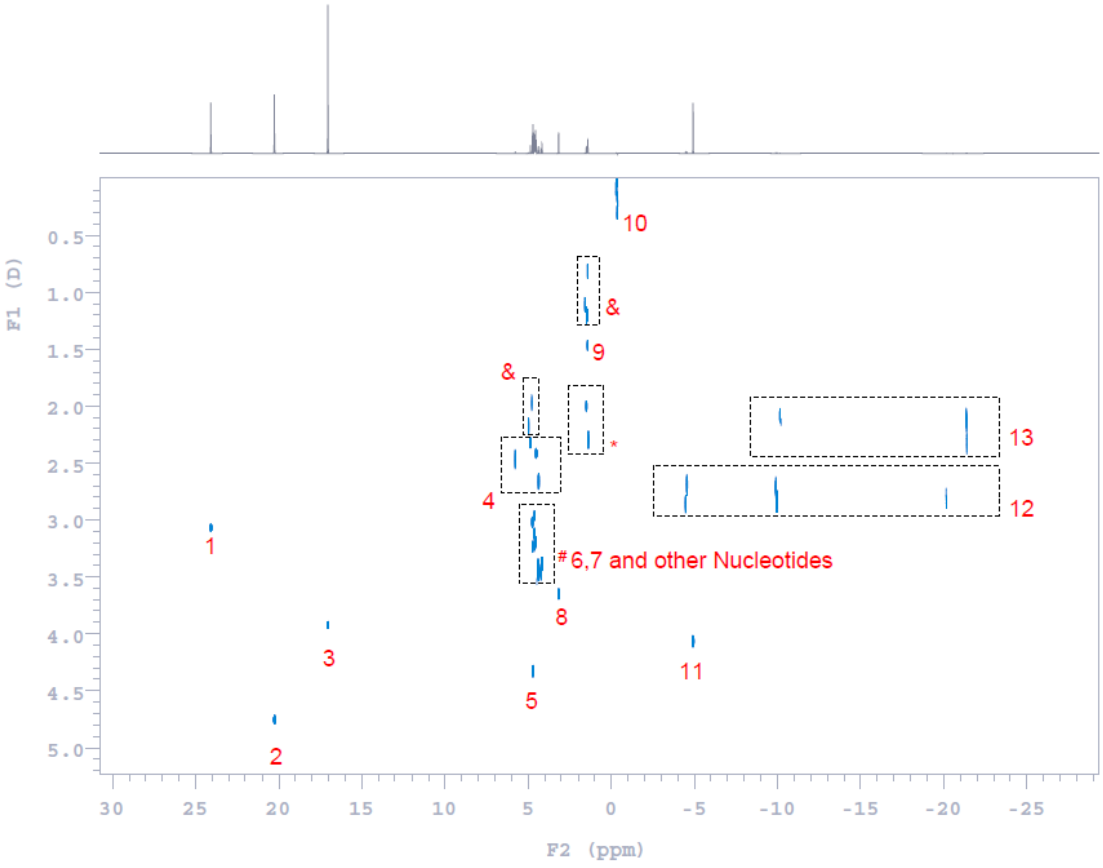


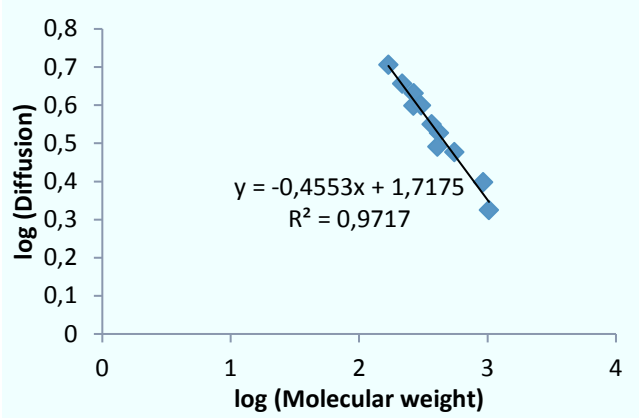
Figure 1. DOSY spectra of P compounds with some additional signals from degradation products and impurities. (*=impurities or degradation products of LTP, # = degradation products of RNA, &= unknown signal).

The equation **(1)** for relating the D value to MW is strictly valid for a spherical particle with a radius R_H. However, even though no molecule is truly spherical, we noticed that some researchers [42, 45-47] still got excellent fitting results between D and MW. This indicated that experimental R_H can also reflect changes in the conformation of a molecule or in its effective charge distribution. Theoretically, there may be a close connection between D and MW for

small molecules in D_2O and other common deuterated solvents.[48] For DOSY samples, the main concern is that different compounds may diffuse unexpectedly because they have different electronic distributions in aqueous solvents, which has a high dielectric constant (79.7 for water). To investigate the correlation between D and MW, we plotted the diffusion coefficients as obtained from the DOSY spectrum of our compounds against their MWs, except for DNA, RNA and LTP, whose precise MWs are unknown. The results showed that an excellent R^2 value (0.97) was achieved based on a linear best fit of log (D) = A*log (MW) + constant, with A < 0 being the slope of the regression (**Fig. 2**). The negative sign reflects that the diffusion rates declined as MW of the compounds increased. The fitting equation can be transformed as $D = 52.2\ MW^{-1/2.2}$ (shown in **Figure S3**), which agrees with Eq. 1. Then the MWs of LTP and DNA can be estimated as being 2.4 KDa and 75.7 KDa (data shown in **Table 1**).

One must be aware that DNA, LTP, etc. are not monodisperse compounds but essentially a distribution of polymeric macromolecules with variable MWs. The MW value given here is simply a calculated value fitting average D value to the regression equation. Moreover, P nuclei in these polymeric molecules usually give rise to a broad envelope of signals as all P nuclei are not in the identical but a broadly similar chemical environment. [34] This kind of line shape was considered as "bad" registration of resonances, which leads to larger error of diffusion value. [42] But we can still differentiate LTP from DNA because their D value varied substantially in DOSY spectrum. Our error in predicting MW for most P compounds was less than 10% relative to their true MW, which corresponded well to the typical 10% limit of accuracy of such experiments. [45] The MW estimation error of *myo*-IHP and P^1P^5-di(adenosine-5') pentaphosphate was slightly higher above 10% because such slow-diffusion species are more susceptible to a calculation error of MW. The reason is that a slight fitting error of D for renders a higher calculation error of its MW because of the "steep" curve of D-MW exponential fitting (**Figure S3**). In addition, error of guanosine 5'-monophosphate was

224 considerably high as it overlapped with two other signals. Note, here that DOSY principally
225 delivers one D value per one frequency value. If one signal is not baseline resolved, the
226 measured D value will be a weighted average of all species under the signal. [49]



227

228 **Figure 2. Log-log-Relationship between the diffusion coefficients in DOSY ^{31}P-NMR spectra**
229 **and MW of organic P reference compounds**

230 Despite some limitations of DOSY mentioned above, we do consider DOSY methodology as
231 promising for both, predicting MW data from log (D) values, as well as for using DOSY for
232 separating the chemical shifts according to the MW of P compounds ranging from 169.0 to
233 1026.3 Da, which could already cover the majority of common P monoesters and polyphos-
234 phate found in soil.

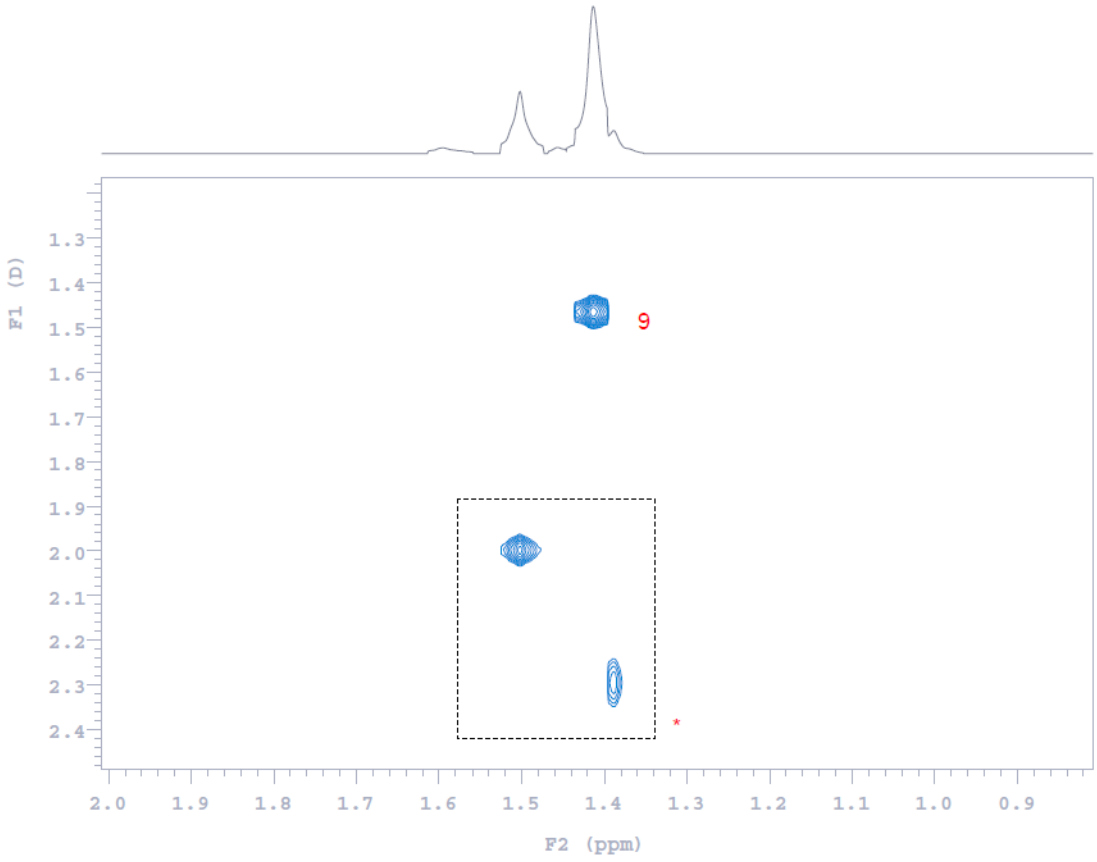
235 **Table 1. D-MW correlation analysis of DOSY data for reference compounds**

ID	10^{-10}D ($m^2.s^{-1}$)	True MW[1] (Da)	Fitted MW[2] (Da)	% Difference
(12-phosphonododecyl)phosphonic acid	3.3658	418.2	411.7	1.6
2-aminoethylphosphonic acid	5.0775	169.0	166.9	1.3
MDP	3.9657	263.9	287.2	8.8
myo-IHP	2.4977	923.8	792.7	14.2
β-glycerophosphate	4.5293	216.0	214.5	0.7
Guanosine 5'-monophosphate	3.0988	407.2	493.6	21.2

Cytidine 5' monophosphate	3.5447	367.2	367.4	0.1
α-D-glucose-1-phosphate	3.9771	304.1	285.3	6.2
ATP	2.9983	551.1	530.7	3.7
Sodium pyrophosphate	4.2782	265.9	243.1	8.6
P^1P^5-Di(adenosine-5') pentaphosphate	2.1128	1026.3	1144.8	11.5
DNA	0.3134		75706.9	
LTP	1.5034		2417.2	

[1] *theoretical MW of disodium salt form of each P compound*
[2] *Fitted MW of each P compound using DOSY NMR data*

Apart from the direct assessment of diffusion rates, enlarging specific chemical shift areas in the DOSY spectrum may also facilitate the direct identification of certain target compounds. Also of importance is that other compounds like *myo*-IHP (signal **4** in **Fig. 1**) are easily recognized in the DOSY spectra according to the similar D values, while in conventional 1D-^{31}P-NMR spectra these signals are sometimes mixed up with signal other monoesters, especially α and β –glycerophosphate.[21] In addition, the 8 points in the dotted rectangle **#** of **Fig. 1** are close to the signal from guanosine 5'-monophosphate (signal **6**, 4.60 ppm, enlarged view of **Fig. 1** was shown in **Figure S1** of supporting information) and cytidine 5'-monophosphate (signal **7**, 4.55 ppm), confirming the assignment of them to mononucleotides from RNA degradation. Another advantage for facilitated signal assignment is illustrated in **Fig. 3**. The signals at 1.50 and 1.41 ppm, possibly assigned to LTP [17, 18], are also indicated here (1D spectrum was shown in **Figure S2**). Yet, looking at the 2D spectrum (**Fig. 3**), the signal at 1.50 ppm and another signal at 1.39 ppm diffuse much slower than signal **9**. In this regard, we may discount the possibility that the two signals also belonged to LTP, and attribute this signal rather to impurities or LTP degradation products.



253

254 **Figure 3. Enlarged view of DOSY spectrum (full spectrum see Fig. 1) excluded the assign-**

255 **ment of signal * to LTP**

256 **3.2 Soil samples**

257 ***3.2.1 Na_2S treatment of Roll-2 sample***

258 The direct application of 1D-^{31}P-NMR to soil extracts suffered from low spectral resolution

259 (trace A in **Fig. 4**), mainly because of paramagnetic ions and that we had used an excessively

260 concentrated sample. The situation did improve first after Na_2S treatment (trace B in **Fig. 4**).

261 The resolution was then further enhanced again after high-speed ultracentrifugation (trace C

262 in **Fig. 4**). High speed centrifugation at low temperature removes more solids than low speed

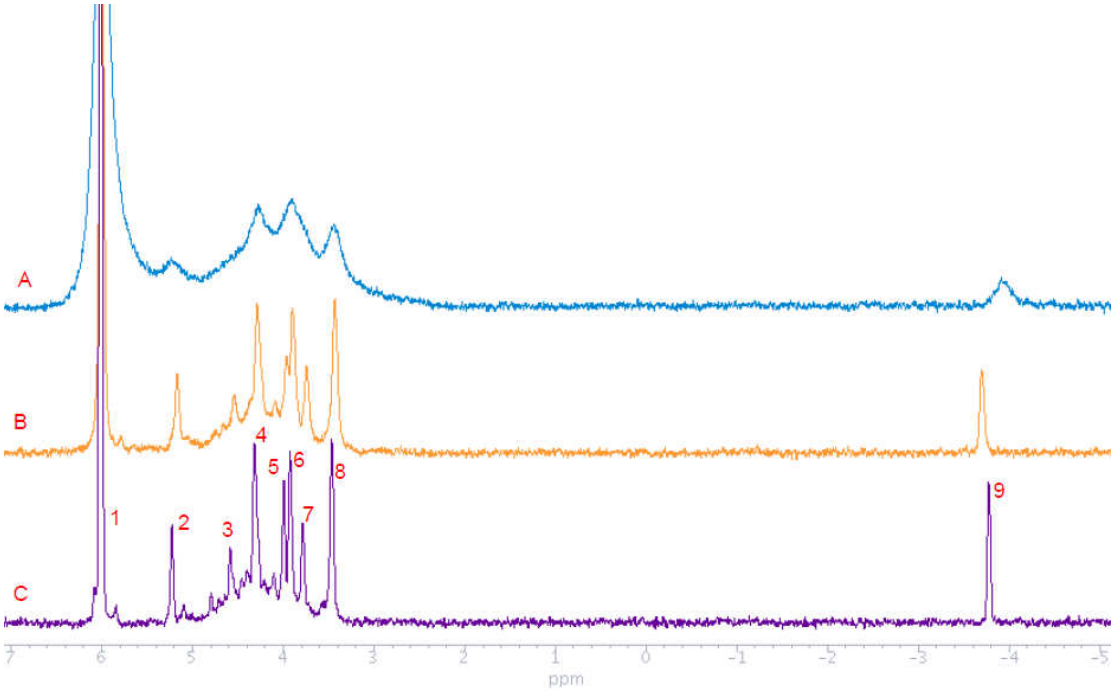
263 centrifugation at room temperature (**Table S2**). The reasons are twofold. Firstly, high-

264 velocity and longer centrifugation time eliminates fine precipitates of iron sulfide and other

265 colloids, thus improving the homogeneity of the magnetic field. As a result, we prolonged T2

266 relaxation and thereby reduced line-broadening. And secondly, the spectrum benefits from
267 the lower solubility of Na_2S at the lower temperature, facilitating the removal of excessive
268 Na_2S, thereby again lowering solution viscosity.

269 A non-viscous solution is required so that the different P components may undergo a non-
270 constructed or unrestricted diffusion. The recovery of P after this precipitation protocol was
271 72.8 ± 1.8 % of P relative to the bulk extract, slightly lower than that of the low-speed centri-
272 fuge scheme (84.5 ± 1.7 %). This is likely due to the absorption of P to Fe-containing colloids,
273 which were further removed by high-speed centrifugation. However, the proportional calcu-
274 lation of individual P species based on line-fitting trace B and C indicated that the high-speed
275 centrifuge protocol did not significantly alter the overall P composition. As a result, the pro-
276 posed post-treatment method enhanced the overall spectral resolution. Although the post-
277 treatment method proposed in current study did show an enhancement in resolution, but
278 further evaluation on other and more soil types are still required to validate the method.



279
280

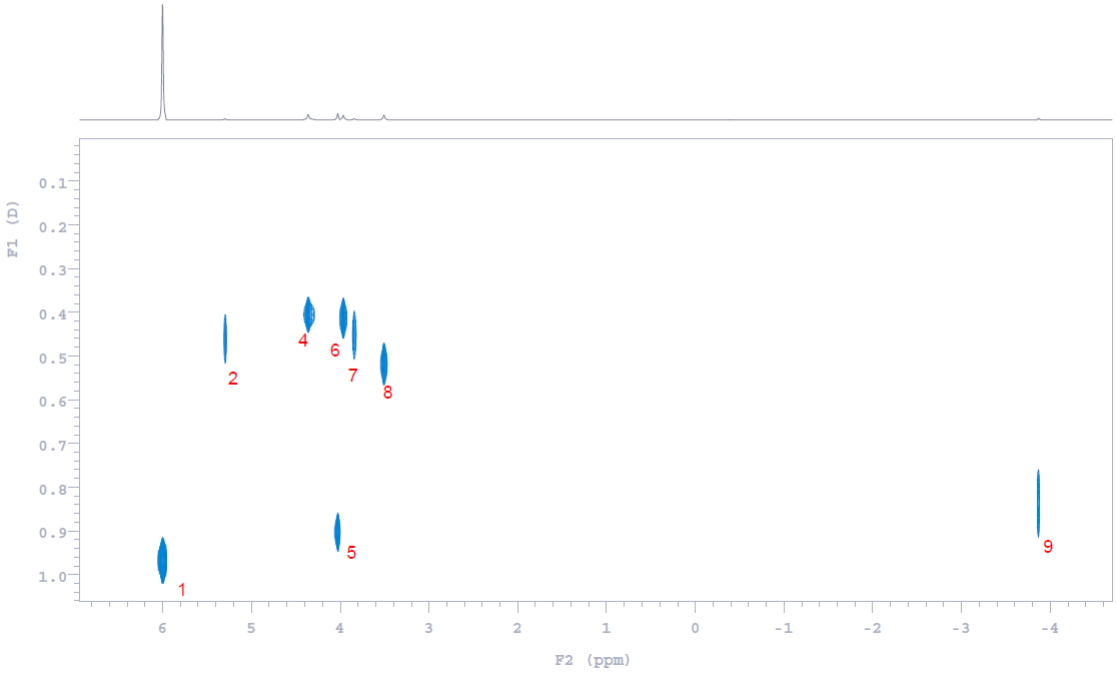
281 **Figure 4. Stacked spectra of soil sample with (A) no Na_2S precipitation and centrifuged at**
282 **7000g for 30min at ambient temperature, (B) Na_2S treatment and centrifuged at 7000g for**

30 min at ambient temperature, and (C) Na_2S treatment and centrifuged at 21500 g for 45 min at 4°C. All spectra were normalized to the same scale.

3.2.2 DOSY result of soil solution

Fig. 5 shows the DOSY result of the grassland soil sample treated with Na_2S following low-temperature centrifugation. In the F2 domain only, the signals **2**, **4**, **6**, **7** (5.2, 4.3, 3.9, 3.8 ppm) showed a 1:2:2:1 pattern, likely indicating contributions from *myo*-IHP. However, signal **5** (4.0 ppm) may be a potential source of misidentification for this assumption. From the F1 domain, signal **5** substantially diffused slower than signals **2**, **4**, **6**, **7**, indicating that it was composed of a smaller molecule, and, as a result, *myo*-IHP resonances were identified as signals **2**, **4**, **6**, **7**. Doolette et al. [21] showed that there have been misidentifications of glycerophosphates and *myo*-IHP in previous studies, because the resonances of the two strongest phytate signals and of the α- and β-glycerophosphate ones are deceptively similar. Our DOSY result differentiated them effortlessly.

As suggested by previous researchers, either *scyllo*-inositol hexakisphosphate (*scyllo*-IHP) or choline phosphate could possibly resonate as signal **8**.[21, 22] But in the DOSY spectrum signal **8** demonstrated a D value very similar to that of *myo*-IHP. Hence, we can rule out the presence of choline phosphate and attribute the signal to *scyllo*-IHP. As signal **5** showed a slightly lower D value than orthophosphate (theoretical MW 163.9 Da), it may be attributed to the glycerophosphate disodium salt (theoretical MW 216.0 Da). For the same reason, it was possible to assign signal **9** to sodium pyrophosphate tetrabasic compounds (theoretical MW 265.9 Da).



304

305 **Figure 5. DOSY spectra of soil sample treated with Na_2S then following high-velocity cen-**
306 **trifugation at 4 °C.**

307 The unambiguous identification of orthophosphate (signal **1**) and therein serves as a kind of
308 internal sample reference for plot of D against MW. When plotting the D value against the
309 MW of the three identified compounds, i.e., orthophosphate, *myo*-IHP, and *scyllo*-IHP, we
310 obtained a close correlation (R^2 =0.98; **Figure S3**). *Myo*-IHP gives rise to four signals, but
311 DOSY treated them separately. Yet, the accuracy of D calculation is dependent on data quali-
312 ty. Thus four various D values were provided here. However, C-5 signal of *myo*-IHP resonat-
313 ed distinctly downfield of most other monoesters as signal **2**,[20, 50] which represents a the
314 more reliable D value for line fitting. From this relationship, the experimental MW of signal **5**
315 and **9** could be derived as 195.2 and 236.8 Da (**Table 2**). Combining estimated MW and
316 chemical shift information, we confidently assign signal 5 and 9 to glycerophosphate disodi-
317 um salt (theoretical MW 216.0 Da) and sodium pyrophosphate tetrabasic (theoretical MW
318 265.9 Da). Again, large molecules feature with high calculation error, from 0.0 % (sodium
319 phosphate) to 15.9% (*myo*-IHP).

Table 2. D-MW correlation result of DOSY data for soil sample

Compound ID	10^{-10}D (m^2s^{-1})	True MW[1] (Da)	fitted MW[2] (Da)	% Difference
myo-IHP	0.4607	923.8	1071.0	15.9
scyllo-IHP	0.5181	923.8	795.9	13.8
Glycerophosphate	0.9032	216.0	195.2	9.6
Sodium phosphate	0.9680	163.9	163.9	0.0
Sodium pyrophosphate	0.8368	265.9	236.8	10.9

[1] *theoretical MW of disodium salt form of each P compound*

[2] *fitted MW of each P compound using DOSY NMR data*

In the DOSY spectrum, the width of the signal along the F1 (=D) axis is determined by the standard error of the D value as obtained from the fitting process.[49] It should be noted that the 2D plot (**Fig. 5**) of soil solution featured considerably broad signal widths in the D domain in comparison to that of model compounds (**Fig. 1**), indicating potentially larger statistical errors for the estimation of D values. In general, experimental errors of DOSY occur either statistically or systematically. The main reason of statistical errors is inadequate signal/noise ratio (S/N), therefore the contribution of noise considerably biases the exponential fitting. For example, signals **2** and **9** are relatively broader than the observed other stronger signals. Fitting results for weak signal **3,** as well as, for other smaller signals were even rejected by Vnmrj software because of the large statistical error (RSD> 10%). We performed DOSY analysis for Wett soil extract, but only D value of orthophosphate was given by DOSY (**Figure S5**). From our experience, we therefore now recommend to accumulate minimum 30 of S/N for individual signal in the first increment of DOSY dataset. On the other hand, another notable source of statistical error is the insufficient attenuation of signal intensity through the DOSY dataset, as caused by slow diffusion due to either large molecular size or highly viscous condition. For example, lower D value is obtained for soil extracts than for mixture of model

compounds. As a result, the signal attenuation of slow diffusion species through DOSY dataset becomes less obvious, causing larger statistical error.

By contrast, systematic artifacts refer to baseline distortion, phase distortion, and broad lineshape of DOSY dataset. These artifacts are primarily induced by three instrumental imperfections, including eddy currents, non-uniformity of the gradient field, and, finally convection.[42] But these artifacts are manageable with proper consideration in terms of experimental setup. For the current study, we applied the BPPSTE pulse sequence, which uses bipolar pulse pairs bracketing refocusing 180^{o} pulse to eliminate distortions from eddy currents and to get rid of gradient-dependent phase distortion.[44] However, these considerations in turn reduce S/N. The BPPSTE principally only allows only half of the signal to be detected, the effective signal is further reduced by signal attenuations due to T1, T2 relaxation. In this study, the 1st increment with the weakest gradient strength in the DOSY dataset resulted in less than 19% S/N ratio of normal 1D spectrum acquired with common simple '1PULSE' pulse sequence. All in all, inadequate S/N of dataset is the major problem for the application of DOSY for soil P research. The time consumption of DOSY could be substantially long to accumulate S/N.

As a first test of the applicability of DOSY to other soils, we also analyzed soil solution extracts that were prepared in similar manner from a Histosol. While resolution was sufficient for separating different signals in the organic soil, we failed to obtain an acceptable S/N (**Figure S5**). While the identification and comparison of different molecular size fractions in the different soils using DOSY-NMR warrants further attention, also in comparison to other methods which combining NMR with, e.g., molecular size fractionation, [34] applying DOSY to a wide range of different soils was beyond the scope of this study. However, the current data already show that the use of DOSY might currently be restricted to soils rich in organic P, while for other soils it might be needed to further concentrate organic P and improve

paramagnetic elimination steps, or to combine DOSY with other sophisticated NMR technology. [51]

In conclusion, DOSY was introduced as new facile and effortless method (compared with spiking experiment) to improve the identification of P compounds and to characterize their molecular weight for the very first time in soil extracts. Applying the technique to common P reference compounds represents an ideal circumstance, i.e., high S/N and low viscosity. Under such conditions DOSY provided a close relationship between diffusion coefficient and MW. This relationship helped to identify P containing degradation products and unknown compounds. However, a remaining concern of the universal application of DOSY for soil P study is inadequate S/N and insufficient signal attenuation of DOSY dataset due to the relatively low P abundance and high viscosity of such soil solutions. Future progress in NMR hardware and sample treatment methods may offer a solution here, e.g., the cryogenic cooling may improve the S/N ratio by a factor of 3-4 by suppressing thermal noises from the coil and preamplifier.[52] This enhancement in sensitivity could substantially cut down the time consumption of DOSY and decrease the statistical error.

Acknowledgements

The authors are thankful to the China Scholarship Council (CSC) for financial support. We thank Dr. Xiaoqian Jiang and Yajie Sun for providing the grassland soil sample as well as for advice during laboratory work. We also acknowledge Dr. Volker Nischwitz for ICP-OES measurement, Ms. Claudia Walraf for lyophilizing the samples, and Dr. Roland Bol for the improvement in English writing.

Supporting Information. Brief statement in a non-sentence format listing the contents of the material supplied as Supporting Information.

4. Reference

1. Bowman, R.; Moir, J., Basic EDTA as an extractant for soil organic phosphorus. *Soil Science Society of America Journal* **1993,** *57*, (6), 1516-1518.
2. Wei, K.; Chen, Z.; Zhu, A.; Zhang, J.; Chen, L., Application of 31 P NMR spectroscopy in determining phosphatase activities and P composition in soil aggregates influenced by tillage and residue management practices. *Soil and Tillage Research* **2014,** *138*, 35-43.
3. Chen, C. R.; Condron, L. M.; Xu, Z. H., Impacts of grassland afforestation with coniferous trees on soil phosphorus dynamics and associated microbial processes: A review. *Forest Ecology and Management* **2008,** *255*, (3–4), 396-409.
4. Ticconi, C. A.; Abel, S., Short on phosphate: plant surveillance and countermeasures. *Trends in plant science* **2004,** *9*, (11), 548-555.
5. Condron, L. M.; Turner, B. L.; Cade-Menun, B. J., Chemistry and dynamics of soil organic phosphorus. In *Phosphorus: Agriculture and the environment*, Sims, J. T., and Sharpley, A.N., Ed. Agron. Monogr. 46. ASA, CSSA, and SSSA: 2005; pp 87-121.
6. McDowell, R. W.; Koopmans, G. F., Assessing the bioavailability of dissolved organic phosphorus in pasture and cultivated soils treated with different rates of nitrogen fertiliser. *Soil Biology and Biochemistry* **2006,** *38*, (1), 61-70.
7. Condron, L.; Frossard, E.; Tiessen, H.; Newmans, R.; Stewart, J., Chemical nature of organic phosphorus in cultivated and uncultivated soils under different environmental conditions. *Journal of Soil Science* **1990,** *41*, (1), 41-50.
8. Sumann, M.; Amelung, W.; Haumaier, L.; Zech, W., Climatic effects on soil organic phosphorus in the North American Great Plains identified by phosphorus-31 nuclear magnetic resonance. *Soil Science Society of America Journal* **1998,** *62*, (6), 1580-1586.
9. Makarov, M. I.; Haumaier, L.; Zech, W.; Marfenina, O. E.; Lysak, L. V., Can 31P NMR spectroscopy be used to indicate the origins of soil organic phosphates? *Soil Biology and Biochemistry* **2005,** *37*, (1), 15-25.
10. Cade-Menun, B. J.; Preston, C. M., A comparison of soil extraction procedures for 31P-NMR spectroscopy. *Soil Science* **1996,** *166*, (770-85).
11. Turner, B., Soil organic phosphorus in tropical forests: an assessment of the NaOH–EDTA extraction procedure for quantitative analysis by solution 31P NMR spectroscopy. *European journal of soil science* **2008,** *59*, (3), 453-466.
12. Redel, Y.; Escudey, M.; Alvear, M.; Conrad, J.; Borie, F., Effects of tillage and crop rotation on chemical phosphorus forms and some related biological activities in a Chilean Ultisol. *Soil Use and Management* **2011,** *27*, (2), 221-228.
13. Lehmann, J.; Lan, Z.; Hyland, C.; Sato, S.; Solomon, D.; Ketterings, Q. M., Long-term dynamics of phosphorus forms and retention in manure-amended soils. *Environmental science & technology* **2005,** *39*, (17), 6672-6680.
14. Bol, R.; Amelung, W.; Haumaier, L., Phosphorus-31–nuclear magnetic–resonance spectroscopy to trace organic dung phosphorus in a temperate grassland soil. *Journal of Plant Nutrition and Soil Science* **2006,** *169*, (1), 69-75.
15. Amelung, W.; Rodionov, A.; Urusevskaja, I.; Haumaier, L.; Zech, W., Forms of organic phosphorus in zonal steppe soils of Russia assessed by 31 P NMR. *Geoderma* **2001,** *103*, (3), 335-350.
16. Vestergren, J.; Vincent, A. G.; Jansson, M.; Persson, P.; Ilstedt, U.; Gröbner, G.; Giesler, R.; Schleucher, J., High-Resolution Characterization of Organic Phosphorus in Soil Extracts Using 2D1H–31P NMR Correlation Spectroscopy. *Environmental Science & Technology* **2012,** *46*, (7), 3950-3956.
17. Cade-Menun, B. J., Improved peak identification in 31P-NMR spectra of environmental samples with a standardized method and peak library. *Geoderma* **2015,** *257–258*, 102-114.

18. Benjamin L. Turner; Nathalie Mahieu; Condron, L. M., Phosphorus-31 Nuclear Magnetic Resonance Spectral Assignments of Phosphorus Compounds in Soil NaOH–EDTA Extracts. *soil Science Society of America Journal* **2003,** *67*, 497-510.
19. Cade-Menun, B.; Liu, C. W., Solution Phosphorus-31 Nuclear Magnetic Resonance Spectroscopy of Soils from 2005 to 2013: A Review of Sample Preparation and Experimental Parameters. *Soil Science Society of America Journal* **2014,** *78*, (1), 19.
20. Smernik, R. J.; Dougherty, W. J., Identification of phytate in phosphorus-31 nuclear magnetic resonance spectra: The need for spiking. *Soil Science Society of America Journal* **2007,** *71*, (3), 1045-1050.
21. Doolette, A.; Smernik, R.; Dougherty, W., Spiking improved solution phosphorus-31 nuclear magnetic resonance identification of soil phosphorus compounds. *Soil Science Society of America Journal* **2009,** *73*, (3), 919-927.
22. Turner, B. L.; Richardson, A. E., Identification of *scyllo*-Inositol Phosphates in Soil by Solution Phosphorus-31 Nuclear Magnetic Resonance Spectroscopy. *Soil Science Society of America Journal* **2004,** *68*, (3), 802-808.
23. Liu, S.; Zhu, Y.; Meng, W.; He, Z.; Feng, W.; Zhang, C.; Giesy, J. P., Characteristics and degradation of carbon and phosphorus from aquatic macrophytes in lakes: Insights from solid-state 13C NMR and solution 31P NMR spectroscopy. *Science of The Total Environment* **2016,** *543, Part A*, 746-756.
24. Bünemann, E. K.; Smernik, R. J.; Doolette, A. L.; Marschner, P.; Stonor, R.; Wakelin, S. A.; McNeill, A. M., Forms of phosphorus in bacteria and fungi isolated from two Australian soils. *Soil Biology and Biochemistry* **2008,** *40*, (7), 1908-1915.
25. Turner, B. L.; Cheesman, A. W.; Godage, H. Y.; Riley, A. M.; Potter, B. V., Determination of neo-and D-chiro-inositol hexakisphosphate in soils by solution 31P NMR spectroscopy. *Environmental science & technology* **2012,** *46*, (9), 4994-5002.
26. Vincent, A. G.; Vestergren, J.; Gröbner, G.; Persson, P.; Schleucher, J.; Giesler, R., Soil organic phosphorus transformations in a boreal forest chronosequence. *Plant and soil* **2013,** *367*, (1-2), 149-162.
27. Morris, K. F.; Johnson Jr, C. S., Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society* **1992,** *114*, (8), 3139-3141.
28. Zheng, G.; Price, W. S., Environmental NMR: Diffusion Ordered Spectroscopy Methods. *eMagRes* **2014,** *2*, (4), 561-574.
29. Simpson, A. J.; McNally, D. J.; Simpson, M. J., NMR spectroscopy in environmental research: From molecular interactions to global processes. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2011,** *58*, (3–4), 97-175.
30. Lead, J. R.; Wilkinson, K. J.; Balnois, E.; Cutak, B. J.; Larive, C. K.; Assemi, S.; Beckett, R., Diffusion coefficients and polydispersities of the Suwannee River fulvic acid: comparison of fluorescence correlation spectroscopy, pulsed-field gradient nuclear magnetic resonance, and flow field-flow fractionation. *Environmental science & technology* **2000,** *34*, (16), 3508-3513.
31. Morris, K. F.; Cutak, B. J.; Dixon, A. M.; Larive, C. K., Analysis of diffusion coefficient distributions in humic and fulvic acids by means of diffusion ordered NMR spectroscopy. *Analytical chemistry* **1999,** *71*, (23), 5315-5321.
32. Šmejkalová, D.; Piccolo, A., Host-guest interactions between 2, 4-dichlorophenol and humic substances as evaluated by 1H NMR relaxation and diffusion ordered spectroscopy. *Environmental science & technology* **2008,** *42*, (22), 8440-8445.
33. Šmejkalová, D.; Spaccini, R.; Fontaine, B.; Piccolo, A., Binding of phenol and differently halogenated phenols to dissolved humic matter as measured by NMR spectroscopy. *Environmental science & technology* **2009,** *43*, (14), 5377-5382.

34. McLaren, T. I.; Smernik, R. J.; McLaughlin, M. J.; McBeath, T. M.; Kirby, J. K.; Simpson, R. J.; Guppy, C. N.; Doolette, A. L.; Richardson, A. E., Complex forms of soil organic phosphorus–a major component of soil phosphorus. *Environmental science & technology* **2015,** *49*, (22), 13238-13245.
35. Jarosch, K. A.; Doolette, A. L.; Smernik, R. J.; Tamburini, F.; Frossard, E.; Bünemann, E. K., Characterisation of soil organic phosphorus in NaOH-EDTA extracts: A comparison of 31 P NMR spectroscopy and enzyme addition assays. *Soil Biology and Biochemistry* **2015,** *91*, 298-309.
36. Jencks, W.; Regenstein, J., Ionization constants of acids and bases. *Handbook of biochemistry and molecular biology* **1976,** *4*, 595-635.
37. Zacharias, S.; Bogena, H.; Samaniego, L.; Mauder, M.; Fuß, R.; Pütz, T.; Frenzel, M.; Schwank, M.; Baessler, C.; Butterbach-Bahl, K., A network of terrestrial environmental observatories in Germany. *Vadose Zone Journal* **2011,** *10*, (3), 955-973.
38. Vincent, A.; Schleucher, J.; Gröbner, G.; Vestergren, J.; Persson, P.; Jansson, M.; Giesler, R., Changes in organic phosphorus composition in boreal forest humus soils: the role of iron and aluminium. *Biogeochemistry* **2012,** *108*, (1-3), 485-499.
39. Doolette, A.; Smernik, R.; Dougherty, W., A quantitative assessment of phosphorus forms in some Australian soils. *Soil Research* **2011,** *49*, (2), 152-165.
40. Alderman, D. W.; Grant, D. M., An efficient decoupler coil design which reduces heating in conductive samples in superconducting spectrometers. *Journal of Magnetic Resonance (1969)* **1979,** *36*, (3), 447-451.
41. Price, K. E.; Lucas, L. H.; Larive, C. K., Analytical applications of NMR diffusion measurements. *Analytical and bioanalytical chemistry* **2004,** *378*, (6), 1405-1407.
42. Antalek, B., Using pulsed gradient spin echo NMR for chemical mixture analysis: how to obtain optimum results. *Concepts in Magnetic Resonance* **2002,** *14*, (4), 225-258.
43. Wider, G.; Dotsch, V.; Wuthrich, K., Self-compensating pulsed magnetic-field gradients for short recovery times. *Journal of Magnetic Resonance, Series A* **1994,** *108*, (2), 255-258.
44. Pelta, M. D.; Barjat, H.; Morris, G. A.; Davis, A. L.; Hammond, S. J., Pulse sequences for high‐resolution diffusion‐ordered spectroscopy (HR‐DOSY). *Magnetic resonance in chemistry* **1998,** *36*, (10), 706-714.
45. Li, W.; Kagan, G.; Yang, H.; Cai, C.; Hopson, R.; Sweigart, D. A.; Williard, P. G., Physically Separated References for Diffusion Coefficient-Formula Weight (D-FW) Analysis of Diffusion-Ordered NMR Spectroscopy (DOSY) in Water. *Organic letters* **2010,** *12*, (12), 2698-2701.
46. Waggoner, R. A.; Blum, F. D.; Lang, J. C., Diffusion in aqueous solutions of poly (ethylene glycol) at low concentrations. *Macromolecules* **1995,** *28*, (8), 2658-2664.
47. Chari, K.; Antalek, B.; Minter, J., Diffusion and scaling behavior of polymer-surfactant aggregates. *Physical review letters* **1995,** *74*, (18), 3624.
48. Evans, R.; Deng, Z.; Rogerson, A. K.; McLachlan, A. S.; Richards, J. J.; Nilsson, M.; Morris, G. A., Quantitative Interpretation of Diffusion-Ordered NMR Spectra: Can We Rationalize Small Molecule Diffusion Coefficients? *Angewandte Chemie* **2013,** *125*, (11), 3281-3284.
49. Morris, G. A., Diffusion‐Ordered Spectroscopy. *eMagRes* **2009**. DOI: 10.1002/9780470034590.emrstm0119.pub2.
50. Turner, B. L., Optimizing phosphorus characterization in animal manures by solution phosphorus-31 nuclear magnetic resonance spectroscopy. *Journal of Environmental Quality* **2004,** *33*, (2), 757-766.
51. Stchedroff, M. J.; Kenwright, A. M.; Morris, G. A.; Nilsson, M.; Harris, R. K., 2D and 3D DOSY methods for studying mixtures of oligomeric dimethylsiloxanes. *Physical Chemistry Chemical Physics* **2004,** *6*, (13), 3221-3227.

538 52. Kovacs, H.; Moskau, D.; Spraul, M., Cryogenically cooled probes—a leap in NMR
539 technology. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2005,** *46*, (2-3), 131-155.

540

541